Statistical and Algorithmic Thresholds in Spin Glasses

by

Brice Huang

S.B., Massachusetts Institute of Technology, 2019
S.M., Massachusetts Institute of Technology, 2022

Submitted to the Department of Electrical Engineering and Computer Science
in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

at the

Massachusetts Institute of Technology

May 2025

Authored by: Brice Huang
Department of Electrical Engineering and Computer Science
May 16, 2025

Certified by: Guy Bresler
Associate Professor of Electrical Engineering and Computer Science
Thesis Supervisor

Certified by: Nike Sun
Professor of Mathematics
Thesis Supervisor

Accepted by: Leslie A. Kolodziejski
Professor of Electrical Engineering and Computer Science
Chair, Department Committee on Graduate Students

Dedicated to my family.

Statistical and Algorithmic Thresholds in Spin Glasses

by

Brice Huang

Submitted to the Department of Electrical Engineering and Computer Science
on May 16, 2025 in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

Abstract

This thesis studies *spin glasses*, disordered complex systems originating in statistical physics. Such systems model optimization, sampling, and inference problems from probability and statistics, which are of fundamental importance to modern data science. In particular, spin glasses provide natural examples of random, high-dimensional, and often highly non-convex cost or log-likelihood functions, making them an excellent testing ground for such questions.

Part I of this thesis studies statistical properties of these models. Chapter 2 identifies the storage capacity of the *Ising perceptron*, a simple model of a neural network, subject to a numerical condition. This gives a conditional proof of a 1989 conjecture of Krauth and Mézard. Chapter 3 gives a new proof of the celebrated *Parisi formula* for the free energy of the spherical mean-field spin glass, which was first proved by Talagrand and in more generality by Panchenko. Our proof takes a simpler modular approach, drawing on recent advances in spin glass free energy landscapes due to Subag. Chapter 4 characterizes the topology trivialization phase transition of multi-species spherical spin glasses and shows that low-temperature Langevin dynamics finds the ground state in the topologically trivial regime; the latter result is new even in the single-species setting.

Part II of this thesis concerns algorithms for optimization and sampling problems on spin glasses. Chapter 5 studies the problem of optimizing the Hamiltonian of a multi-species spherical spin glass. Our main result exactly characterizes the maximum value attainable by a class of algorithms that are suitably Lipschitz in the disorder. This class includes gradient-based algorithms and Langevin dynamics on constant time scales, and in particular includes the best algorithm known for this problem. This chapter is part of a series of works where we establish exact algorithmic thresholds using the branching overlap gap property (OGP), a landscape property introduced in our earlier work (which appears in our S.M. thesis). In this chapter, we develop a more robust way to establish the branching OGP that does not require Guerra's interpolation; this allows our method to be applied to models well beyond the (single-species) mean-field spin glass we previously considered.

Chapters 6 and 7 study sampling from the Gibbs measure of a spherical mean-field spin glass. Chapter 6 develops a sampling algorithm based on simulating Eldan's stochastic localization scheme, while Chapter 7 analyzes simulated annealing of Langevin dynamics. We prove both algorithms succeed for inverse temperatures up to a *stochastic localization threshold*. Chapter 6 gives the first stochastic localization-based sampler with a guarantee of vanishing total variation error, improving on earlier algorithms with vanishing Wasserstein error. Chapter 7 provides the first provable guarantees for a Markov chain in this model beyond the *uniqueness threshold*, where mixing from worst-case initialization is provably slow.

Acknowledgements

First and foremost, I would like to thank my advisors, Guy Bresler and Nike Sun, for their patient and generous mentorship. They have been exceptional role models — not only in mathematics, but also in mathematical communication, asking the right questions, and the more intangible aspects of research. They have been immensely supportive in every aspect of my graduate school experience, from deeply motivating research collaborations to invaluable career and life advice. I am lucky to have learned from both of them.

Equally significant has been the influence of my longtime collaborator, Mark Sellke. Three of the six papers in this thesis are coauthored with him, as is a comparable fraction of my papers in graduate school. Some of my fondest memories from this time are of afternoons working with Mark in the Math Department common room, bouncing ideas off of each other or exploring new topics together. This collaboration has continually pushed my probabilistic intuition, and I have learned a great deal of math as a result.

I would also like to thank all the other collaborators I have coauthored papers with in graduate school, of which not all is yet published: Ryan Alweiss, Matthew Brennan, Sitan Chen, Aukosh Jagannath, Jerry Li, Allen Liu, Sidhanth Mohanty, Ankur Moitra, Andrea Montanari, Elchanan Mossel, Huy Tuan Pham, Amit Rajaraman, David X. Wu, Claire Zhang, and Alexis Zhou. I also thank Hang Du, David Gamarnik, Eren Kızıldağ, Frederic Koehler, Kuikui Liu, Ben McKenna, Ankur Moitra, Cristopher Moore, Dheeraj Nagaraj, Will Perkins, Francisco Pernice, Yury Polyanskiy, Mehtaab Sawhney, Tselil Schramm, Youngtak Sohn, Alex Wein, Jiaming Xu, and Ilias Zadik for many enlightening and thought-provoking mathematical discussions.

I spent two stimulating semesters visiting the Simons Institute at UC Berkeley, for the programs "Probability, Geometry, and Computation in High Dimensions" in Fall 2020 and "Computational Complexity of Statistical Inference" in Fall 2021. These were among the most formative research experiences during my graduate studies, and I thank everyone I met at these programs.

In addition to those named above, I am grateful to numerous academic siblings, fellow students, and friends, who were an important part of my time in graduate school: Mitali Bafna, Kiril Bangachev, Enric Boix-Adserà, Anna Brandenberger, Josh Brunner, Evan Chen, Justin Chen, Holly Cheng, Byron Chin, Dingding Dong, Yilun Du, Melihcan Erol, Max Fishelson, Jason Gaitonde, Kavish Gandhi, Noah Golowich, Chenghao Guo, Meghal Gupta, Alina Harbuzova, Sunny He, Matthew Hong, Jun-Ting Hsieh, Henry Hu, Jane Lange, John Kuszmaul, Timothy Kwa, Daniel Lee, Saba Lepsveridze, Yang Liu, Henry Ma, Mahbod Majid, Varun Mangalick, Nitya Mani, Surya Mathialagan, Abhijit Mudigonda, Shivam Nadimpalli, Shyam Narayanan, Ted Pyne, Dhruv Rohatgi, Ashwin Sah, Sandeep Silwal, Oriol Solé Pi, Zoey Song, Austin Stromme, Norah Tan, Arsen Vasilyan, Tanay Wakhare, Tony Wang, Rachel Wu, Zixuan Xu, Dmitrii Zakharov, Rachel Zhang, Sylvia Zhang, Wei Zhang, and Kai Zheng. Special thanks to the Theory of Computation group, where I (unofficially) sat for the past several years, for being such a welcoming home.

I would like to express my deep gratitude to the childhood teachers who nurtured my interest in mathematics: J. Timothy Cornell, Alyce Doehner, Rick Heckman, and Chuen-Chin Hsuchen. I am grateful to Alex Postnikov, Mustazee Rahman, and Wuttisak Trongsiriwat for mentoring my first research projects and guiding me on the research path. I am also grateful to Joe Gallian, whose mentorship in the 2018 Duluth Research Experience for Undergraduates (REU) convinced me to pursue a Ph.D., and to my advisors and classmates at Duluth that inspiring summer: Niven Achenjang, Levent Alpöge, Aaron Berger, Amanda Burcroff, Colin Defant, Marisa Gaetz, Caleb Ji, Noah Kravitz, Erik Metz, Ashwin Sah, and Danielle Wang.

I would like to thank Panpan Gao, Min-Min Liang, and Kang Zhou from the MIT Global Languages department, with whom I took three semesters of Chinese during graduate school. These classes were a delightful way to reconnect with my heritage language and culture, and a welcome break that brought balance and meaning to each day. I am also grateful to my classmates for making this experience so rewarding, and to Feli Xiao for several inspiring conversations.

I am grateful to the members of `CelloWorld();`, MIT's cello club, including Aidan Chambers, Valerie Chen, Joy Feng, Hee Jae Hong, Elizabeth Lee, Janelle Sands, Maya Sankar, Eren Shin, Archer Wang, Jett Wang, David Xiong, and many more. I am lucky to have been a part of this community, which was a constant source of happiness and sanity throughout graduate school — from sight-reading sessions and BSO outings to caroling in Lobby 10 for the annual Holiday Cellobration. During the COVID years, we stayed connected through multitrack projects where we digitally stitched together parts that we recorded individually, and this was one of the high points of my pandemic experience.

Finally, I am deeply indebted to my family for their consistent and unconditional support. To my parents, Suzanne and Kevin, thank you for encouraging my mathematical pursuits from the beginning, for teaching me the importance of being curious, and for always giving the right life advice. To my sister Catherine, our FaceTime calls are a constant source of positivity at the end of each day. I would not be who I am today without my family, and this thesis is dedicated to them.

Funding acknowledgment

I gratefully acknowledge the following sources of funding over the course of my graduate studies: a NSF Graduate Research Fellowship; a Siebel Scholarship; a Google Ph.D. Fellowship; NSF award CCF-1940205; NSF award DMS-1940092; NSF-Simons collaboration award DMS-2031883; and the Solomon Buchsbaum Research Fund at MIT.

Contents

Chapter 1

Introduction

This thesis studies *spin glasses*, models of random complex systems originating in statistical physics. These models were first introduced in the 1970s to model certain disordered magnetic materials [EA75, SK75]. Since then, the mathematical study of these models has led to wide-ranging applications in computer science, statistics, and beyond.

Concretely, a spin glass is described by a random *Hamiltonian* $H : \Omega \to \mathbb{R}$ over a high-dimensional state space $\Omega \subseteq \mathbb{R}^N$ (common choices are $S_N = \sqrt{N}\mathbb{S}^{N-1}$ or $\Sigma_N = \{\pm 1\}^N$), which can be interpreted as a high-dimensional objective function. The associated *Gibbs measure* $\mu(\boldsymbol{x}) \propto e^{H(\boldsymbol{x})}$ is likewise a natural random probability distribution. These notions of random function and distribution are useful to many applications:

- In Bayesian inference, one aims to estimate an unknown parameter $\boldsymbol{x}$ given some noisy observations $\boldsymbol{A}$. In such settings, μ models the **posterior distribution** $\mathcal{L}(\boldsymbol{x}|\boldsymbol{A})$, which is a random distribution because $\mathcal{A}$ is random. Applications of this perspective include community detection [DKMZ11, Moo17, Abb18], error-correcting codes [RU08], and compressed sensing [ZK16]; see [ABDM23] for a survey treatment.

- The Hamiltonian H models an **objective function generated from random data**. These arise in many applications in statistics and machine learning, for example as the objective in variational inference or the loss function of a neural network [CHM$^+$15, GGLZ23].

- In a complex system, small elementary components interact in a local manner, giving rise to interesting emergent global behaviors. Here, H models an **energy function** whose terms describe which local configurations are favored or disfavored. Applications include protein folding [BW87, RMS$^+$21], flocking birds [BCG$^+$12], ecology [MGA23], and economics [BMN23].

All in all, spin glasses are a natural testing ground for **random, high-dimensional optimization, sampling, and inference problems** at the heart of modern data science. This thesis will study these models through the following questions.

- (Statics) What are the values of key statistical properties of a spin glass, such as the free energy or satisfiability threshold? What is the structure of the energy landscape?

- (Algorithms) What are the optimal efficient algorithms for optimization and sampling problems over random, highly non-convex landscapes? On the other hand, when are such problems intractable?

A recurring theme of this thesis is that these questions are deeply intertwined: algorithms can give proofs of statistical properties, and statistical insights on solution geometry can inform algorithm design. In the rest of this introductory chapter, we summarize the key ideas from each work appearing in this thesis.

1.1 Statics

Part I of this thesis studies three problems in statics. A common thread in these works is to use modern algorithmic ideas to make progress on statics problems.

1.1.1 Capacity threshold of Ising perceptron

Chapter 2 is based on the paper [Hua24], which appeared in FOCS 2024 as the Best Student Paper.

The Ising perceptron is a defined as the intersection of a high-dimensional discrete cube with random half-spaces. Formally, for $\Sigma_N = \{\pm 1\}^N$ and i.i.d. gaussian vectors $\boldsymbol{g}^1, \ldots, \boldsymbol{g}^M \sim \mathcal{N}(0, \boldsymbol{I}_N)$, the solution set of this model is defined by

$$S = \{\boldsymbol{x} \in \Sigma_N : \langle \boldsymbol{g}^a, \boldsymbol{x} \rangle \geq 0 \quad \forall 1 \leq a \leq M\} . \tag{1.1}$$

This is a simple model of a neural network: S represents the set of synaptic weights that memorize M random patterns $\boldsymbol{g}^1, \ldots, \boldsymbol{g}^M$. We let $M = \lfloor \alpha N \rfloor$ for *constraint density* α, and study the following problem.

Problem 1.1.1 (Capacity problem)**.** Determine (if it exists) the critical constraint density $\alpha_\star$ such that

$$\lim_{N\to\infty} \mathbb{P}(S \neq \emptyset) = \begin{cases} 1 & \alpha < \alpha_\star, \\ 0 & \alpha > \alpha_\star. \end{cases}$$

This critical $\alpha_\star$ is known as the capacity threshold, and models the maximum number of patterns per synapse that this neural network can memorize.

Using non-rigorous statistical physics techniques, Krauth and Mézard [KM89] conjectured an explicit value of $\alpha_\star$, which is approximately 0.833. Ding and Sun [DS18] showed that this value is a lower bound on the capacity, under a numerical condition that an explicit univariate function is nonpositive.[1] In Chapter 2, we show the matching upper bound, under a numerical condition that an explicit bivariate function (plotted in Figure 2.1) is nonpositive. This gives a conditional proof of Krauth and Mézard's conjecture.

Furthermore, one may define a more general perceptron model with margin parameter $\kappa \in \mathbb{R}$ by

$$S = \left\{ \boldsymbol{x} \in \Sigma_N : \frac{\langle \boldsymbol{g}^a, \boldsymbol{x} \rangle}{\sqrt{N}} \geq \kappa \quad \forall 1 \leq a \leq M \right\} .$$

Both our result and that of [DS18] hold for this model as well, and together confirm an analogous capacity prediction $\alpha_\star(\kappa)$ under several further numerical conditions depending on κ.

Ideas of the proof. One strategy to try to locate the capacity threshold is the first and second moment method: if $\mathbb{E}[|S|] \ll 1$, then S is empty with high probability, and if $\mathbb{E}[|S|^2] \asymp \mathbb{E}[|S|]^2$, then S is nonempty with positive probability. If these estimates hold for respectively all $\alpha > \alpha_\star$ and all $\alpha < \alpha_\star$, this shows the capacity is $\alpha_\star$. This strategy identifies the capacity threshold in the *symmetric Ising perceptron* [APZ19, PX21, ALS22b], a closely related model where the constraints take the form $|\langle \boldsymbol{g}^a, \boldsymbol{x} \rangle|/\sqrt{N} \leq \kappa$. Unfortunately, this direct approach does not work in our model (1.1), as $\mathbb{E}[|S|] = 2^{N-M}$ only is vanishing for $\alpha > 1$, which does not locate the true threshold $\alpha_\star \approx 0.833$. This demonstrates that the first moment is dominated by rare events and does not capture the model's typical behavior.

The main idea of both the paper of Ding and Sun [DS18] and our work is to work with a distribution similar to the true one, on which the first and second moment method succeeds. The choice of distribution is motivated by the TAP heuristic [TAP77] from physics: let $\boldsymbol{G} \in \mathbb{R}^{M \times N}$ be the matrix whose rows are $\boldsymbol{g}^1, \ldots, \boldsymbol{g}^M$, and for $\boldsymbol{m} \in \mathbb{R}^N$, $\boldsymbol{n} \in \mathbb{R}^M$, define the *TAP equation*

$$\boldsymbol{n} = \dot{F}\left(\frac{\boldsymbol{G}\boldsymbol{m}}{\sqrt{N}} - b(\boldsymbol{m})\boldsymbol{n}\right), \qquad \boldsymbol{m} = \widehat{F}\left(\frac{\boldsymbol{G}^\top \boldsymbol{n}}{\sqrt{N}} - d(\boldsymbol{m}, \boldsymbol{n})\boldsymbol{n}\right). \tag{1.2}$$

Here, $\dot{F}, \widehat{F} : \mathbb{R} \to \mathbb{R}$ are explicit nonlinearities (applied coordinate-wise) and b, d are explicit scalar-valued functions. The TAP heuristic predicts that for a typical realization of $\boldsymbol{G}$, there is a unique solution $(\boldsymbol{m}, \boldsymbol{n})$, and furthermore provides a predicted distribution for $(\boldsymbol{m}, \boldsymbol{n})$.

This leads to a wishful calculation: consider a *planted model* where we sample $(\boldsymbol{m}, \boldsymbol{n})$ from its predicted distribution, and then $\boldsymbol{G}$ conditional on satisfying (1.2). If the TAP heuristic is true, the distribution of $\boldsymbol{G}$ sampled from the planted model should approximate the true model. Indeed, we may think of $(\boldsymbol{m}, \boldsymbol{n})$ as

[1] The result of [DS18] states that for $\alpha < \alpha_\star$, $S \neq \emptyset$ with positive probability, i.e. $\liminf_{N\to\infty} \mathbb{P}(S \neq \emptyset) > 0$. The subsequent works [Xu21, NS23] showed that this model has a sharp threshold sequence, which improves this guarantee to high probability.

a function of $\boldsymbol{G}$, and interpret the planted model as first sampling $(\boldsymbol{m},\boldsymbol{n})$ from its distribution, and then sampling $\boldsymbol{G}$ to be consistent with $(\boldsymbol{m},\boldsymbol{n})$. Moreover, one can calculate that in the planted model, the first and second moment method **conditional on a typical realization of** $(\boldsymbol{m},\boldsymbol{n})$ locates the exact capacity threshold $\alpha_\star$ (assuming our and [DS18]'s numerical conditions, which encode these moment computations). That is, conditioning on $(\boldsymbol{m},\boldsymbol{n})$ removes the rare events that dominated $\mathbb{E}[|S|]$ in the direct moment method, and allows the moment method to succeed.

However, this wishful calculation does not mean anything unless we can rigorously link the planted model to the true model, and this is the main difficulty of this problem. In [DS18], this calculation instead enters in a motivational role; their rigorous argument constructs a truncation of $|S|$ that weaves in the TAP heuristic, and carries out the first and second moment method on this random variable. This truncation is the reason their approach gives a lower bound. Our work shows that the true and planted models are contiguous in a suitable sense, so that the wishful calculation implies the capacity threshold in the true model.

We show this contiguity by proving a version of the TAP heuristic: for a typical realization of $\boldsymbol{G}$ from the true model, there exists a unique solution $(\boldsymbol{m},\boldsymbol{n})$ to (1.2) in a certain region $\mathcal{S}\subseteq\mathbb{R}^N\times\mathbb{R}^M$ (on which the predicted distribution of $(\boldsymbol{m},\boldsymbol{n})$ concentrates) with high probability. Existence is proven algorithmically, by showing that the *approximate message passing* (AMP) iteration [Bol14, BM11] finds a TAP fixed point. Uniqueness is shown by a new double-counting argument, which is the crux of this work. We show that in the planted model, the same AMP iteration with high probability finds **the planted point** $(\boldsymbol{m},\boldsymbol{n})$. Since any realization of $\boldsymbol{G}$ with multiple TAP fixed points can arise in the planted model with any of these TAP fixed points as the planted one, such $\boldsymbol{G}$ cannot occur in the true model with non-vanishing probability.

Finally, we mention that this strategy of showing contiguity with a model with planted TAP fixed point also appears as a key technical step in Chapters 6 and 7, on sampling from a spin glass's Gibbs measure. This strategy first appeared in the work in Chapter 6, though the contiguity is proved in a different way and used for a different purpose. The connections between these works are explained in detail in Subsections 1.2.2, 2.2.2 and 7.7.1.

1.1.2 Constructive proof of spherical Parisi formula

Chapter 3 is based on the paper [HS23b], which is joint work with Mark Sellke and is submitted for publication.

Let $\gamma_1,\dots,\gamma_P\ge 0$ be a fixed sequence of model parameters, which we encode in the *mixture function* $\xi(s)=\sum_{p=1}^P\gamma_p^2 s^p$. The mean-field spin glass Hamiltonian is the random function $H_N:\mathbb{R}^N\to\mathbb{R}$ defined by

$$H_N(\boldsymbol{\sigma})=\sum_{p=1}^P\frac{\gamma_p}{N^{(p-1)/2}}\sum_{i_1,\dots,i_p=1}^N g_{i_1,\dots,i_p}\sigma_{i_1}\cdots\sigma_{i_p},\qquad g_{i_1,\dots,i_p}\stackrel{i.i.d.}{\sim}\mathcal{N}(0,1). \tag{1.3}$$

Equivalently, this is the gaussian process on $\mathbb{R}^N$ with covariance

$$\mathbb{E}H_N(\boldsymbol{\sigma})H_N(\boldsymbol{\rho})=N\xi(\langle\boldsymbol{\sigma},\boldsymbol{\rho}\rangle/N)\,.$$

We consider the domain of H_N to be either $S_N=\sqrt{N}\mathbb{S}^{N-1}$ or $\Sigma_N=\{\pm1\}^N$ (*spherical* or *Ising* spins). In the spherical case, we define the *free energy* of this model to be

$$F_N=\log\int e^{H_N(\boldsymbol{\sigma})}\,\mathsf{d}\mu_0(\boldsymbol{\sigma})\,,$$

where μ_0 is the uniform Haar measure on S_N. In the Ising case the only difference is that μ_0 is the counting measure on Σ_N. A central problem in spin glass theory is the following.

Problem 1.1.2. Determine the in-probability limit $F=\text{p-lim}_{N\to\infty}F_N/N$.

In [Par79], Parisi gave a prediction for the limiting free energy based on his groundbreaking *replica symmetry breaking* ansatz. This formula takes the form of an infinite-dimensional variational problem:

$$\underset{N\to\infty}{\text{p-lim}}\frac{F_N}{N}=\mathsf{P}(\xi)\equiv\inf_{\zeta\in\mathcal{P}([0,1])}\mathsf{P}(\zeta;\xi)\,. \tag{1.4}$$

Here $\mathcal{P}([0,1])$ denotes the set of Borel probability measures on $[0,1]$, and P is a certain *Parisi functional.*

After decades of progress in the probability and statistical physics communities [MPS+84a, MPV87, Rue87, GG98, ASS03, Gue03], this formula was famously proved by Talagrand [Tal06b], and in more generality by Panchenko [Pan13a]. In Chapter 3, we give a new and simpler proof of the (more difficult) lower bound in the Parisi formula for spherical models.

Proof ideas. Our proof takes a new view of the *ultrametricity* prediction [MPS+84a] at the heart of physicists' understanding of this model, which states that the Gibbs measure $\mu(\mathsf{d}\boldsymbol{\sigma}) \propto e^{H(\boldsymbol{x})}\mu_0(\mathsf{d}\boldsymbol{\sigma})$ concentrates on a hierarchy of clusters within clusters (and therefore the dominant contribution to F_N comes from such a set). For the spherical spin glass, this means that the Gibbs measure concentrates on a union of spherical caps ("pure states"), whose centers are the leaves of an ultrametric, orthogonally-branching tree in $\mathbb{R}^N$. The tree's ancestor nodes represent recursive cluster centers located inside the sphere.

Moreover, it is predicted that this tree branches precisely at radii $r \in [0, \sqrt{N}]$ whose corresponding self-overlap $R = r^2/N$ lies in the support of the minimizer ζ_* to the Parisi formula (1.4). In other words, ζ_* specifies the tree's shape, by specifying the set of radii where this tree branches; see Figure 1.1 for an illustration. In general ζ_* can be discrete ("finite RSB"), continuous ("full RSB"), or an interleaving of these behaviors, and the ultrametric tree is correspondingly complex.

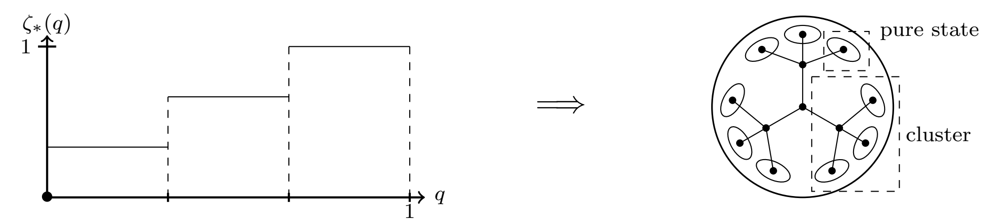


Figure 1.1: Cumulative distribution function of a Parisi measure ζ_* and its corresponding ultrametric tree. If ζ_* has a continuous part, the tree "branches continuously" in the limit.

Panchenko's celebrated work [Pan13a] proved the Parisi ultrametricity conjecture, which leads to one proof of the Parisi formula. In our proof, ultrametricity is not a derived consequence, but a starting point: we construct a set of spherical caps arranged according to Parisi's ansatz, each with free energy given by the Parisi formula. This witnesses the lower bound.

Our approach begins with the minimizer ζ_* of the Parisi formula; note that existence and uniqueness of this measure is a fact about the Parisi variational problem, which a priori says nothing about the free energy. We then construct a tree that branches precisely at the radii specified by ζ_*, starting at the origin and proceeding outwards "layer by layer." A key ingredient of this proof strategy is the *uniform concentration* lemma introduced in [Sub24]. This lemma implies that even though any node $\boldsymbol{x} \in \mathbb{R}^N$ of the tree we construct depends on H_N in a complicated way, the restriction of H_N to a codimension-1 orthogonal band passing through $\boldsymbol{x}$ resembles a (rescaled) spherical spin glass in one fewer dimension. As a result, constructing the children of an internal node $\boldsymbol{x}$ of the tree amounts to understanding a smaller spin glass's ground state energy. Similarly, computing the free energy of the pure state around a leaf $\boldsymbol{x}$ amounts to determining a smaller spin glass's free energy.

All in all, this results in a modular proof strategy, where the main task is to understand the free energy or ground state energy of several smaller spin glasses. The benefit of this decomposition is that each of the smaller spin glasses is simple: whereas the ζ_* of the original model (and thus, the corresponding ultrametric tree) can be arbitrarily complicated, each of the sub-models in the decomposition has one of four basic behaviors corresponding to a single layer of the tree. We call these behaviors "fundamental model types," and they are depicted in Figure 1.2.

The remaining task is to lower bound the ground state energy of the *topologically trivial*, *1-step RSB (1RSB)* , and *full RSB (FRSB)* types, and the free energy of the *replica symmetric (RS)* type. Out of these, two types are already understood: the ground state of the topologically trivial type is determined by works [FLD14, BČNS22] (described further in Subsection 1.1.3) studying the energy landscape via the Kac–Rice formula, and that of the FRSB type is witnessed by an algorithm due to Subag [Sub21a]. Our work develops

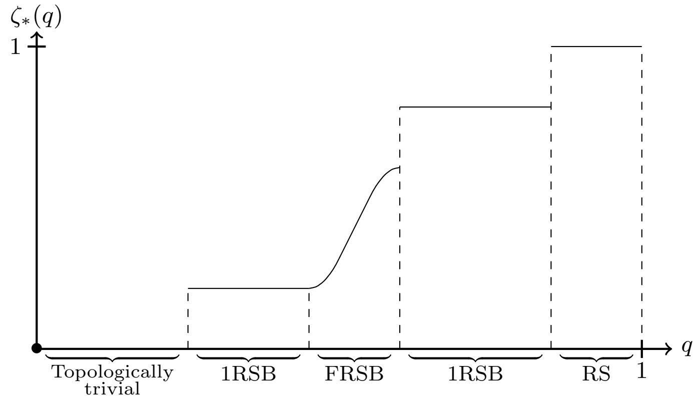


Figure 1.2: Decomposition of an Parisi measure ζ_* into fundamental parts.

a new truncated second moment argument, which determines the free energy of the RS type and the ground state energy of the 1RSB type. This completes the proof.

1.1.3 Strong topological trivialization

Chapter 4 is based on the paper [HS23c], which is joint work with Mark Sellke. This work will appear in the Annals of Probability.

A fundamental direction of inquiry in spin glass theory concerns the geometry of the landscape, the surface plotted by the random Hamiltonian. An influential line of work, pioneered by [Fyo04, ABČ13], proposed to quantify the complexity of the landscape by the number of critical points — that is, a higher number of critical points is indicative of a more rugged landscape. They study the number of critical points (henceforth, the *complexity*) via the Kac–Rice formula [Kac48, Ric44], which provides access to the expectation of this random variable (the *annealed complexity*).

One notable phenomenon studied in this line of work is *topology trivialization*. This is a phase transition separating a "complex" regime, where the Hamiltonian has exponentially many critical points, from a "simple" one where it has $O(1)$ critical points. For the spherical mean-field model with Hamiltonian (1.3), [FLD14, Fyo15, BČNS22] locate the phase boundary of *annealed topological trivialization* in the following sense. If the parameter γ_1 (which controls the strength of the external field term $\gamma_1 \sum_{i=1}^N g_i\sigma_i$ in H_N) is small, the annealed complexity of H_N is exponentially large. As γ_1 increases beyond a critical value, the annealed complexity drops to $2 + o(1)$. Since any differentiable function on S_N has at least two critical points, namely the global maximum and minimum, in the latter regime these are the only critical points with high probability.

In Chapter 4, we introduce and study several questions, which probe aspects of the topology trivialization phenomenon that are not yet understood; see (1), (2), and (3). In this introductory chapter, we focus the discussion on one of these questions, which links topology trivialization to algorithmic guarantees.

Problem 1.1.3. It has long been expected that in the topologically trivial phase, algorithms such as gradient ascent or low-temperature Langevin dynamics efficiently find the global maximum. However, this implication is not known rigorously, as regions with small gradient can in principle cause slow convergence. Does topological trivialization imply the rapid convergence of such optimization algorithms?

We answer this question affirmatively, by establishing a structural property in the topologically trivial phase that implies algorithmic tractability. Namely, we show that all $O(1)$ critical points of H_N have well-conditioned (Riemannian) Hessian, and all points where H_N has small gradient are near a true critical point, a condition we call *strong topological trivialization*. We then show that under this condition, low-temperature Langevin dynamics find the global maximum of H_N efficiently, in $O(\log N)$ time.

In fact, we show these results for the *multi-species spherical spin glass*, where the Hamiltonian is a generalization of (1.3) and the domain is a product of $r = O(1)$ high-dimensional spheres (see (1.6) and (1.7) below). This model includes the $r = 2$ bipartite SK model [KC75, KS85], and has been the subject of much recent work [BCMT15, Pan15, Mou21, Mou23, BL20, BS22a, Sub21b, Sub23b, Kiv23]. We identify the annealed topology trivialization phase boundary for this model. In the trivialized regime, we show there are exactly 2^r critical points with high probability, which is the minimum possible for a Morse function on a product of r spheres. Going further, we show strong topological trivialization also holds in this regime, and therefore low-temperature Langevin dynamics efficiently find the global maximum.

Working with the multi-species spin glass introduces several new challenges not present in the single-species setting. First, the main term one must evaluate in the Kac–Rice formula is the expected determinant of a large random matrix. In the single-species setting, this matrix is a sample from the Gaussian Orthogonal Ensemble (GOE), and the determinant has an explicit formula. In the multi-species setting, this matrix is a gaussian block matrix, and the access we get to the determinant is much less explicit: it is a functional of the random matrix's limiting spectrum, which is defined through a vector Dyson equation. Our work develops an understanding of the vector Dyson equation's solution in order to complete this calculation.

Another challenge is that tools for evaluating these more general random determinants are only accurate to a multiplicative factor of $e^{o(N)}$. As a result, evaluating the Kac–Rice formula only shows an upper bound of $e^{o(N)}$ on the annealed complexity in the topologically trivial regime, as opposed to $2 + o(1)$ in the single-species setting. One of the key innovations of our work is a method that, using only this weak form of annealed topological trivialization as input, deduces that there are exactly 2^r critical points with high probability.

1.2 Algorithms

Part II of this thesis studies algorithms for optimization and sampling problems on spin glasses. A recurring theme of this part is to identify the geometric structure in the problem that allows an efficient algorithm to succeed, thereby linking algorithmic tractability to structural phase transitions in the solution landscape.

1.2.1 Optimization of spin glass Hamiltonians

Chapter 5 is based on the paper [HS23a], which is joint work with Mark Sellke and is submitted for publication. It studies the question: how well can an efficient algorithm optimize a random objective?

Background: optimizing the mean-field model. This chapter builds on our earlier work [HS25] with Sellke, which appears in our S.M. thesis and studies the following problem.

Problem 1.2.1. Let H_N be the mean-field spin glass Hamiltonian (1.3). Devise a polynomial-time algorithm $\mathcal{A}$ which takes H_N as input and outputs $\mathcal{A}(H_N) = \boldsymbol{\sigma}^{\mathsf{alg}}$ in $S_N = \sqrt{N}\mathbb{S}^{N-1}$ or $\Sigma_N = \{\pm 1\}^N$ with $H_N(\boldsymbol{\sigma}^{\mathsf{alg}})$ as large as possible. What is the largest value of $H_N(\boldsymbol{\sigma}^{\mathsf{alg}})$ that can be achieved?

It is known that the in-probability limiting maximum values

$$\mathsf{OPT}^{\mathsf{Is}} = \operatorname*{p-lim}_{N\to\infty} \frac{1}{N} \max_{\boldsymbol{\sigma}\in\Sigma_N} H_N(\boldsymbol{\sigma}), \qquad \mathsf{OPT}^{\mathsf{Sp}} = \operatorname*{p-lim}_{N\to\infty} \frac{1}{N} \max_{\boldsymbol{\sigma}\in S_N} H_N(\boldsymbol{\sigma})$$

are given by a zero-temperature version of the Parisi formula [Par79, Tal06b, Pan13a, AC17]. Optimization algorithms for this problem were developed in [Sub21a] for the spherical setting, and [Mon21, AMS21, Sel24a] for the Ising setting. These algorithms achieve values $\mathsf{ALG}^{\mathsf{Is}}$ and $\mathsf{ALG}^{\mathsf{Sp}}$ which are explicit but generally smaller than $\mathsf{OPT}^{\mathsf{Is}}$ and $\mathsf{OPT}^{\mathsf{Sp}}$. That is, for any $\varepsilon > 0$ independent of N,

$$\mathbb{P}(H_N(\boldsymbol{\sigma}^{\mathsf{alg}})/N \geq \mathsf{ALG} - \varepsilon) = 1 - o(1). \tag{1.5}$$

Here $\mathsf{ALG} = \mathsf{ALG}^{\mathsf{Is}}$ or $\mathsf{ALG}^{\mathsf{Sp}}$ respectively, and in fact the probability is at least $1 - e^{-cN}$ for some $c = c(\varepsilon) > 0$.

The main result of [HS25] is a matching hardness result for a class of Lipschitz algorithms, which we now introduce. Let $\mathscr{H} = \mathbb{R}^{N^2+\cdots+N^P}$, and let $\boldsymbol{g} \in \mathscr{H}$ be the concatenation of the gaussians $g_{i_1,\ldots,i_p}$ appearing in H_N. We may identify H_N with $\boldsymbol{g}$, and view $\mathcal{A}$ as a map from $\boldsymbol{g} \in \mathscr{H}$ to $\boldsymbol{\sigma}^{\mathsf{alg}} \in \mathbb{R}^N$. We say $\mathcal{A}$ is L-Lipschitz if for all $\boldsymbol{g}, \boldsymbol{g}' \in \mathscr{H}$,

$$\|\mathcal{A}(\boldsymbol{g}) - \mathcal{A}(\boldsymbol{g}')\| \leq L\|\boldsymbol{g} - \boldsymbol{g}'\|.$$

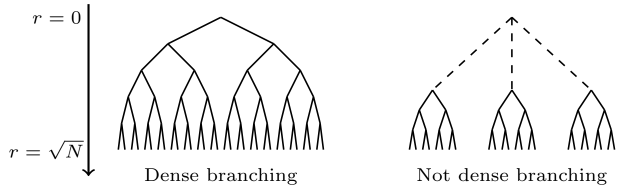


Figure 1.3: Schematics of a densely branching tree and a tree with discrete jumps. Nodes represent points in $\mathbb{R}^N$, and the vertical axis represents the points' distance to the origin. The tree of points with value ALG must branch densely, whereas the one of points with value OPT may have either behavior.

We also relax the domains S_N and Σ_N to $B_N = \{\boldsymbol{\sigma} \in \mathbb{R}^N : \|\boldsymbol{\sigma}\| \le \sqrt{N}\}$ and $C_N = [-1,1]^N$; in the Ising case this is necessary for Lipschitz algorithms mapping to this domain to exist. As explained in [HS25], the class of $O(1)$-Lipschitz algorithms includes natural algorithms such as gradient descent, Langevin dynamics, and approximate message passing (AMP) for $O(1)$ time. In particular, this includes the above algorithms that achieve ALG.[2]

Say H_N is *even* if $\gamma_p = 0$ for all odd p. [HS25] shows the following hardness result.

Theorem 1.2.2. *Suppose H_N is even and let $\mathsf{ALG} = \mathsf{ALG}^{\mathsf{Sp}}$ (resp. $\mathsf{ALG}^{\mathsf{Is}}$). For any $\varepsilon > 0$ and $L > 0$, there exists $c = c(\varepsilon, L) > 0$ such that the following holds for sufficiently large N. For any L-Lipschitz $\mathcal{A} : H_N \to \boldsymbol{\sigma}^{\mathsf{alg}} \in B_N$ (resp. C_N),*

$$\mathbb{P}(H_N(\boldsymbol{\sigma}^{\mathsf{alg}})/N \ge \mathsf{ALG} + \varepsilon) \le e^{-cN}.$$

The proof is based on a landscape property that we introduce, which we call the *branching overlap gap property.* This is a generalization of the overlap gap property (OGP) introduced by Gamarnik and Sudan [GS17a]; see [Gam21, Mon23a, AMS23c] for survey treatments. The basic idea is that a Lipschitz algorithm, run on a suitable correlated family of problem instances H_N, can be made to output a constellation of solutions arranged as the leaves of a densely branching ultrametric tree; see e.g. (5.31) for a formal definition. We show that for any $\varepsilon > 0$, with probability $1 - e^{-cN}$ the solution space (of this correlated family of problems) does not have a set of points with this geometry, each with value at least $\mathsf{ALG} + \varepsilon$. This implies that a Lipschitz algorithm cannot reach $\mathsf{ALG} + \varepsilon$.

Densely branching ultrametric trees also play an important role in the algorithms of [Sub21a, Mon21, AMS21, Sel24a], and this is why our method is able to locate a sharp algorithmic threshold. The strategy of these algorithms is to trace a root-to-leaf path of a densely branching ultrametric tree of points, whose leaves all have value ALG. In other words, **we can characterize ALG geometrically as the supremal value whose super-level set typically contains a densely branching ultrametric tree**. Algorithms can reach values up to ALG by following a densely branching ultrametric, and cannot do better due to the branching OGP.

Remark 1.2.3. As a consequence of the ultrametricity of Gibbs measures predicted by Parisi [Par79, Par83] and proved by Panchenko [Pan13a], the set of points with value OPT also forms the leaves of an ultrametric tree. The main difference between this tree and the tree of points with value ALG described above is that the latter tree must branch densely, i.e. at a dense set of radii. In contrast, the tree of points with value OPT may have jumps, corresponding to gaps in the overlap support of the model (1.3); see Figure 1.3 for an illustration. In particular, $\mathsf{ALG} = \mathsf{OPT}$ if and only if the tree of points with value OPT is also densely branching, i.e. if the model (1.3) has full overlap support $[0,1]$ at zero temperature.

Our results and proof ideas. The results of [HS25] motivate the following question.

Problem 1.2.4. Can we establish exact algorithmic thresholds for optimizing spin glass Hamiltonians beyond the mean-field model? How general is the above geometric description of ALG?

[2]The algorithms of [Mon21, AMS21, Sel24a] are AMP iterations run for $O(1)$ time. While the algorithm of [Sub21a] does not have this form, an analogous AMP iteration also achieves ALG in the spherical setting, as explained in [AMS21].

A key challenge is that the method used to establish the branching OGP in [HS25] is somewhat brittle. The main technical input to the hardness proof therein is to show that the leaves of any densely branching ultrametric tree have average energy at most ALG. In [HS25], this is proved using Guerra's interpolation [Gue03], which applies only to models whose covariance structure is convex in a certain sense. This limits the set of models on which this method succeeds.

In Chapter 5, we study this problem for multi-species spherical spin glass Hamiltonian, defined as follows. Let $\mathcal{I}_1 \cup \cdots \cup \mathcal{I}_r$ be a partition of $[N] = \{1, \ldots, N\}$, with $\lim_{N\to\infty} |\mathcal{I}_s|/N = \lambda_s \in (0,1)$ for all $s \in [r]$. For $i \in [N]$, let $s(i) \in [r]$ be such that $i \in \mathcal{I}_{s(i)}$. For fixed model parameters $\gamma_{s_1,\ldots,s_p} \ge 0$, where $1 \le p \le P$ and $s_1, \ldots, s_p \in [r]$, consider the Hamiltonian

$$H_N(\boldsymbol{\sigma}) = \sum_{p=1}^{P} \frac{1}{N^{(p-1)/2}} \sum_{i_1,\ldots,i_p=1}^{N} \gamma_{s(i_1),\ldots,s(i_p)} g_{i_1,\ldots,i_p} \sigma_{i_1} \cdots \sigma_{i_p}, \qquad g_{i_1,\ldots,i_p} \overset{i.i.d.}{\sim} \mathcal{N}(0,1). \tag{1.6}$$

We study the problem of optimizing this Hamiltonian over the product-of-spheres domain

$$\mathcal{S}_N = \left\{ \boldsymbol{x} \in \mathbb{R}^N : \|\boldsymbol{x}_s\|^2 = \lambda_s N\,,\ \forall s \in \mathscr{S} \right\}\,, \tag{1.7}$$

where $\boldsymbol{x}_s \in \mathbb{R}^{\mathcal{I}_s}$ denotes the restriction of $\boldsymbol{x}$ to coordinates $\mathcal{I}_s$.

We emphasize that the interpolation method does not apply to this model, and for this reason the value of OPT remains open. Nonetheless, we are able to characterize an explicit algorithmic threshold ALG. Our main result is as follows. Let

$$\mathcal{B}_N = \left\{ \boldsymbol{x} \in \mathbb{R}^N : \|\boldsymbol{x}_s\|^2 \le \lambda_s N\,,\ \forall s \in \mathscr{S} \right\}$$

be the product of balls which relaxes the product of spheres $\mathcal{S}_N$.

Theorem 1.2.5. *For any $\varepsilon > 0$ and $L > 0$, there exists $c = c(\varepsilon, L) > 0$ such that the following holds for sufficiently large N. For any L-Lipschitz $\mathcal{A} : H_N \to \boldsymbol{\sigma}^{\mathsf{alg}} \in \mathcal{B}_N$,*

$$\mathbb{P}(H_N(\boldsymbol{\sigma}^{\mathsf{alg}})/N \ge \mathsf{ALG} + \varepsilon) \le e^{-cN}.$$

Our companion work [HS24] provides an AMP algorithm similar to those of [Mon21, AMS21, Sel24a], which outputs $\boldsymbol{\sigma}^{\mathsf{alg}} \in \mathcal{S}_N$ achieving value ALG in the sense of (1.5). Together, these results show that ALG is the optimal value attained by Lipschitz algorithms.

The main tool in the proof is a new method to establish the branching OGP, which does not use the interpolation method. Instead, we upper bound the average value of the leaves of any dense ultrametric tree by adapting the uniform concentration idea introduced in [Sub24]. Uniform concentration upper bounds the increment between the value of any node in the tree and the average value of its children. This gives a simpler and more robust proof of the branching OGP based on only gaussian concentration of measure. Specializing to the one-species setting, our method also gives a new proof of Theorem 1.2.2 in spherical models without the evenness assumption, which was previously required to apply the interpolation method.

This new method of showing the branching OGP is quite general. In work in progress with Sellke and Sun [HSS25], we demonstrate that this approach also establishes an exact algorithmic threshold for the following generalized random perceptron model. Let $M = \lfloor \alpha N \rfloor$ for fixed constraint density $\alpha > 0$. Let $\phi \in \mathcal{C}_b(\mathbb{R})$ be an arbitrary activation function and $\boldsymbol{g}^1, \ldots, \boldsymbol{g}^M \sim \mathcal{N}(0, \boldsymbol{I}_N)$. We consider the problem of optimizing the Hamiltonian

$$H_N(\boldsymbol{\sigma}) = \sum_{a=1}^{M} \phi\left(\frac{\langle \boldsymbol{g}^a, \boldsymbol{\sigma}\rangle}{\sqrt{N}}\right).$$

A significant amount of recent work has been devoted to studying algorithmic properties of this and related models [BS20, ALS22a, AS22, GKPX22, GKPX23, BAKZ24, MZZ24, LSZ25]. In [HSS25], we give a matching optimization algorithm and hardness result for Lipschitz algorithms via the branching OGP.

Altogether, these results demonstrate that the geometric description of the algorithmic threshold put forward in [HS25] is very general. This provides a unified picture of optimal algorithms in a broad class of random optimization problems.

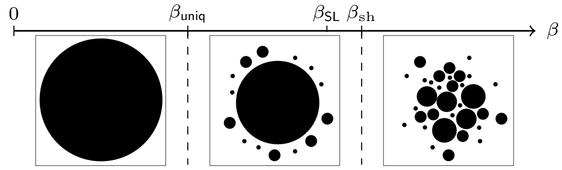


Figure 1.4: Structural phases of the Gibbs measure; images adapted from [KMRT+07].

1.2.2 Sampling from Gibbs measures

Chapter 6 is based on the paper [HMP24], which is joint work with Andrea Montanari and Huy Tuan Pham and is submitted for publication. Chapter 7 is based on the paper [HMRW25], which is joint work with Sidhanth Mohanty, Amit Rajaraman, and David X. Wu and will appear in STOC 2025.

Recall the mixed p-spin Hamiltonian H_N defined by (1.3). Let μ_0 denote the uniform Haar measure on $S_N = \sqrt{N}\mathbb{S}^{N-1}$. For *inverse temperature* $\beta \geq 0$, define the *Gibbs measure* μ_β by

$$\mu_\beta(\mathsf{d}\boldsymbol{\sigma}) = \frac{e^{\beta H_N(\boldsymbol{\sigma})}}{Z_\beta}\mu_0(\mathsf{d}\boldsymbol{\sigma})\,,$$

where Z_β is a normalizing constant. This is a random, highly non-log-concave probability distribution. Chapters 6 and 7 study the following problem.

Problem 1.2.6. Devise a polynomial-time randomized algorithm which takes (H_N, β) as input and outputs $\boldsymbol{\sigma}^{\mathsf{alg}} \in S_N$, such that the the law μ^{alg} of $\boldsymbol{\sigma}^{\mathsf{alg}}$ satisfies $\mathrm{TV}(\mu^{\mathsf{alg}}, \mu_\beta) = o(1)$. For which β is this possible?

Physics heuristics suggest that as one varies β, one encounters a series of phase transitions in the structure of μ_β which govern the tractability of this problem; see Figure 1.4 for an illustration. For small β, μ_β is expected to be well-connected, in the sense of e.g. satisfying a Poincaré inequality, and thus the Langevin dynamics from any initialization mix rapidly to μ_β. For β beyond the *uniqueness threshold* β_{uniq}, it is expected that nearly all of the mass of μ_β remains in a well-connected cluster, but small clusters with poor connectivity to the bulk of the measure (*metastable states*) appear. In this regime, it is expected that the Langevin dynamics with *random* initialization will succeed by sampling from the main cluster, even though these dynamics with *worst-case* initialization can get trapped in a metastable state. Beyond the *shattering threshold* β_{sh}, μ_β shatters into exponentially many small clusters. In this regime, the Langevin dynamics fail to efficiently sample from μ_β, and it is expected that no efficient algorithm succeeds [CHS93].

However, much less is rigorously known about sampling algorithms for this problem. At sufficiently high temperature, it is known that a Poincaré inequality holds and the Langevin dynamics mix rapidly to μ_β [GJ19]. Relatedly, recent lines of work on *spectral independence* and *localization schemes* [ALO21, CE22] have developed powerful tools to establish functional inequalities that imply rapid mixing of many Markov chains. This notably includes the Glauber dynamics for the analogous Gibbs measure on the cube $\Sigma_N = \{\pm 1\}^N$ [EKZ22, AJK+21a, ABXY24] for suitably small β. However, all of these techniques can only apply for $\beta < \beta_{\mathsf{uniq}}$, as they show rapid mixing from a worst-case initialization, which is false in the presence of metastable states.

In order to sample at larger values of β, [AMS22] introduced a different sampling algorithm based on simulating Eldan's stochastic localization scheme [Eld13, Eld20b]. In particular, for the Sherrington–Kirkpatrick model on the (more difficult) cube Σ_N, their algorithm succeeds in the entire high-temperature regime $\beta < 1$. This approach has since been applied to Bayesian posterior sampling [MW23] and is equivalent to denoising diffusions from machine learning [HJA20]; see [Mon23b] for the connection. However, until now a key limitation of this approach is that it comes with a guarantee of vanishing *Wasserstein* error, rather than total variation error. That is, with high probability over H_N, there exists a coupling of $\boldsymbol{\sigma}^{\mathsf{alg}} \sim \mu^{\mathsf{alg}}$ and $\boldsymbol{\sigma} \sim \mu_\beta$ such that

$$\mathbb{E}\|\boldsymbol{\sigma}^{\mathsf{alg}} - \boldsymbol{\sigma}\|^2 = o(N)\,. \tag{1.8}$$

We also mention that several other aspects of the physics picture above have been rigorously proven for the spherical pure p-spin models. The presence of metastable states for all $\beta > \beta_{\mathsf{uniq}}$ was shown in [BJ24].

[AMS25] shows that at an absolute constant factor above $\beta_{\rm sh}$, μ_β is shattered, and a class of stable algorithms fails to sample from μ_β. This provides evidence that the problem is computationally hard in the shattered regime.

Our work improves on both approaches to sampling. In Chapter 6, we develop a stochastic localization algorithm that samples from the spherical spin glass with vanishing total variation error. In Chapter 7, we show that a *simulated annealing* variant of the Langevin dynamics, where we start at $\beta = 0$ and gradually increase β over time, also samples from μ_β. Both algorithms succeed up to a *stochastic localization threshold* $\beta_{\mathsf{SL}} \in (\beta_{\mathsf{uniq}}, \beta_{\rm sh})$. In particular, Chapter 7 provides the first guarantee for a Markov chain in this problem beyond β_{uniq}, where mixing from a worst-case initialization is provably slow. For the pure p-spin models, β_{SL} is also within an absolute constant factor of the conjectured computational threshold $\beta_{\rm sh}$.

Proof ideas: algorithmic stochastic localization. The stochastic localization process can be described by the Itô diffusion

$$\mathsf{d}\boldsymbol{y}_t = \boldsymbol{m}(\boldsymbol{y}_t, t)\mathsf{d}t + \mathsf{d}\boldsymbol{B}_t\,, \qquad \boldsymbol{y}_0 = \boldsymbol{0}\,. \tag{1.9}$$

Here $\boldsymbol{B}_t$ is a $\mathbb{R}^N$-valued standard Brownian motion, and $\boldsymbol{m}(\boldsymbol{y}_t, t)$ is the mean of the tilted distribution

$$\mu_{\beta,t}(\mathsf{d}\boldsymbol{\sigma}) \propto e^{\langle \boldsymbol{y}_t, \boldsymbol{\sigma}\rangle} \mu_\beta(\mathsf{d}\boldsymbol{\sigma})\,.$$

As explained in e.g. [Eld20b], $\mu_{\beta,t}$ is a measure-valued martingale that, as $t \to \infty$, localizes to a point mass (whose location is thus distributed as μ_β). Hence, algorithmically simulating the process (1.9) via Euler discretization provides a way to sample from μ_β.

The main challenge in implementing this strategy is to algorithmically estimate the means $\boldsymbol{m}(\boldsymbol{y}_t, t)$ of the measures $\mu_{\beta,t}$. In [AMS22], this estimator $\widetilde{\boldsymbol{m}}_t$ is defined as the fixed point of a TAP equation to which a certain AMP iteration converges. Using the *state evolution* [Bol14, BM11] analysis of AMP, it is shown that this estimator has error satsisfying

$$\mathbb{E}\|\widetilde{\boldsymbol{m}}_t - \boldsymbol{m}(\boldsymbol{y}_t, t)\|^2 = o(N)\,. \tag{1.10}$$

This is enough to imply the Wasserstein error guarantee (1.8). The main contribution of our work in Chapter 6 is to develop and analyze a more accurate mean estimator, which upgrades the error guarantee to total variation. Our work shows that the error in (1.10) is actually $O(1)$. Then, we identify a correction term to $\widetilde{\boldsymbol{m}}_t$, resulting in an estimator $\widehat{\boldsymbol{m}}_t$ such that

$$\mathbb{E}\|\widehat{\boldsymbol{m}}_t - \boldsymbol{m}(\boldsymbol{y}_t, t)\|^2 = o(1)\,. \tag{1.11}$$

Adapting an analysis from [CCL+23] shows that an algorithmic simulation of (1.9) using mean estimator $\widehat{\boldsymbol{m}}_t$ has vanishing total variation error to the true process. This implies the result.

We discuss some ideas of the proof of the main estimate (1.11) at the end of this subsection.

Proof ideas: simulated annealing. A common method to show that a Markov chain mixes rapidly is to establish a functional inequality, such as a Poincaré inequality. For simplicity, we restrict the discussion here to measures $\mu \in \mathcal{P}(S_N)$ with densities with respect to μ_0, with the spherical Langevin dynamics as the Markov semigroup, though these facts hold much more generally. For test functions $f, g \in \mathcal{C}^\infty(S_N)$, define the *Dirichlet form*

$$\mathcal{E}_\mu(f, g) = \mathbb{E}_{\boldsymbol{\sigma}\sim\mu}\langle \nabla f(\boldsymbol{\sigma}), \nabla g(\boldsymbol{\sigma})\rangle\,.$$

We say μ satisfies a *Poincaré inequality* with constant C if for any $f \in \mathcal{C}^\infty(S_N)$,

$$\mathcal{E}_\mu(f, f) \ge C\,\mathsf{Var}_\mu(f)\,. \tag{1.12}$$

Let ν_t be the distribution obtained by running the spherical Langevin dynamics for time t, from initial distribution $\nu_0 \in \mathcal{P}(S_N)$. The inequality (1.12) implies the exponential contraction of χ^2 divergence

$$\chi^2(\nu_t\|\mu) \le e^{-Ct}\chi^2(\nu_0\|\mu)\,. \tag{1.13}$$

If C is at least inverse polynomially large in N, this implies polynomial-time mixing of the spherical Langevin dynamics from worst-case initialization.

However, as discussed above, such an inequality can hold for μ_β only if $\beta < \beta_{\mathsf{uniq}}$, as rapid mixing from worst-case initialization cannot hold in the presence of metastable states. In order to analyze Markov chain dynamics for larger β, in Chapter 7 we study *weak Poincaré inequalities* of the form

$$\mathcal{E}_\mu(f,f) \ge C\mathsf{Var}_\mu(f) - \varepsilon\|f - \mathbb{E}_\mu f\|_\infty^2 \,. \tag{1.14}$$

Similarly to (1.13), this implies a mixing guarantee of the form

$$\chi^2(\nu_t\|\mu) \le e^{-Ct}\chi^2(\nu_0\|\mu) + \varepsilon\left\|\frac{\mathsf{d}\nu_0}{\mathsf{d}\mu} - 1\right\|_\infty^2 \,. \tag{1.15}$$

In our application, we will have $C = \Omega(1)$ and $\varepsilon = e^{-\Omega(N^{1/5})}$. Then (1.15) implies mixing (up to χ^2 error $e^{-\Omega(N^{1/5})}$, and thus total variation error $e^{-\Omega(N^{1/5})}$) **from sufficiently warm starts**.

This mixing guarantee combines well with the *simulated annealing* algorithm, which proceeds in $T = \mathsf{poly}(N)$ stages as follows. For $1 \le i \le T$, let $\beta_i = \beta \cdot i/T$. In the i-th stage, we run the spherical Langevin dynamics corresponding to μ_{β_i} for $\mathsf{poly}(N)$ time, initialized at the output of the previous stage. (In the first stage, we initialize at a uniformly random point in S_N.) The main idea of our approach is that for suitable T, each $\mu_{\beta_{i-1}}$ is a sufficiently warm start for μ_{β_i}. Thus, if all the μ_{β_i} satisfy the weak Poincaré inequality (1.14), we can inductively argue that the output of the i-th stage is an approximate sample from μ_{β_i}. Hence simulated annealing approximately samples from μ_β.

Our proof that the μ_{β_i} satisfy a weak Poincaré inequality parallels recent developments in the spectral independence and localization schemes [ALO21, CE22] lines of work. For $\boldsymbol{h} \in \mathbb{R}^N$ and $\mu \in \mathcal{P}(S_N)$, define the tilted measure

$$\mu_{\boldsymbol{h}}(\mathsf{d}\boldsymbol{\sigma}) \propto e^{\langle \boldsymbol{h},\boldsymbol{\sigma}\rangle}\mu(\mathsf{d}\boldsymbol{\sigma}) \,.$$

A central message of [ALO21, CE22] is that, roughly speaking, if $\|\mathsf{Cov}(\mu_{\boldsymbol{h}})\|_{\mathsf{op}} = O(1)$ for all $\boldsymbol{h} \in \mathbb{R}^N$, then μ satisfies a Poincaré inequality with constant $\Omega(1)$. This provides a powerful method for establishing a Poincaré inequality, which has been useful in numerous applications.

Our work shows that analogously, if $\|\mathsf{Cov}(\mu_{\boldsymbol{h}})\|_{\mathsf{op}} = O(1)$ with high probability for *random* $\boldsymbol{h}$, then μ satisfies a weak Poincaré inequality. More precisely, let $t \ge 0$, and let $\boldsymbol{h} = t\boldsymbol{\sigma} + \sqrt{t}\boldsymbol{g}$, for $(\boldsymbol{\sigma},\boldsymbol{g}) \sim \mu \otimes \mathcal{N}(0,\boldsymbol{I}_N)$. If for all t, $\|\mathsf{Cov}(\mu_{\boldsymbol{h}})\|_{\mathsf{op}} = O(1)$ with probability $1 - e^{-\Omega(N^{1/5})}$, then ν satisfies (1.14) with $C = \Omega(1)$, $\varepsilon = e^{-\Omega(N^{1/5})}$. This provides a framework for proving weak Poincaré inequalities, where the main technical input is to show $\|\mathsf{Cov}(\mu_{\boldsymbol{h}})\|_{\mathsf{op}} = O(1)$ holds with high probability for the aforementioned random tilts $\boldsymbol{h}$. We discuss this input further below.

Proof ideas: conditioning on TAP fixed point. In the algorithmic stochastic localization approach, we must prove that the mean estimator $\widehat{\boldsymbol{m}}_t$ of $\mu_{\beta,t}$ satisfies (1.11). In the weak Poincaré inequalities approach, we must show $\|\mathsf{Cov}(\mu_{\boldsymbol{h}})\|_{\mathsf{op}} = O(1)$ with high probability. One of the main steps of both proofs will be to show that the law of H_N is contiguous to a model with a planted TAP fixed point, similarly to the strategy described in Subsection 1.1.1 and carried out in Chapter 2. We next explain why passing to a planted model is useful.

It can be shown that in the true model, the measures $\mu_{\beta,t}$ and $\mu_{\boldsymbol{h}}$ concentrate near the (random) codimension-1 band orthogonal to a suitable TAP fixed point $\boldsymbol{m}^{\mathsf{TAP}}$ (which equals $\widetilde{\boldsymbol{m}}_t$ in the algorithmic stochastic localization approach), and passing through it:

$$B = \left\{\boldsymbol{\sigma} \in S_N : \langle \boldsymbol{\sigma} - \boldsymbol{m}^{\mathsf{TAP}}, \boldsymbol{m}^{\mathsf{TAP}}\rangle = 0\right\} \,.$$

So, for either estimating the mean or bounding the covariance, the region near this band is the key part of the Gibbs measure we must understand.

However, it is difficult to study the Gibbs measure on this band in the true model, as we do not explicitly know the joint law of $(\boldsymbol{m}^{\mathsf{TAP}}, H_N)$. This is the problem that the planted model solves: in the planted model, we can explicitly compute the law of H_N conditional on $\boldsymbol{m}^{\mathsf{TAP}}$. This opens the way to calculations that prove the mean estimate in Chapter 6 and covariance bound in Chapter 7.

Part I

Statics

Chapter 2

Capacity threshold for the Ising perceptron

Abstract – We show that the capacity of the Ising perceptron is with high probability upper bounded by the constant $\alpha_\star \approx 0.833$ conjectured by Krauth and Mézard, under the condition that an explicit two-variable function $\mathscr{S}_\star(\lambda_1, \lambda_2)$ is maximized at $(1, 0)$. The earlier work of Ding and Sun [DS18] proves the matching lower bound subject to a similar numerical condition, and together these results give a conditional proof of the conjecture of Krauth and Mézard.

2.1 Introduction

The Ising perceptron was introduced in [Wen62, Cov65] as a simple model of a neural network. Mathematically, it is an intersection of a high-dimensional discrete cube with random half-spaces, defined as follows. Fix any $\kappa \in \mathbb{R}$ (our main result is for $\kappa = 0$). For $N \geq 1$, let $\Sigma_N = \{\pm 1\}^N$, and let $\boldsymbol{g}^1, \boldsymbol{g}^2, \dots$ be a sequence of i.i.d. samples from $\mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_N)$. For $M \geq 1$, the Ising perceptron is the random set

$$S_N^M = \left\{ \boldsymbol{x} \in \Sigma_N : \frac{\langle \boldsymbol{g}^a, \boldsymbol{x} \rangle}{\sqrt{N}} \geq \kappa \quad \forall 1 \leq a \leq M \right\}. \tag{2.1}$$

As explained in [Gar87], S_N^M models the set of configurations of synaptic weights in a single-layer neural network that memorize all M patterns $\boldsymbol{g}^1, \dots, \boldsymbol{g}^M$. Define the random variable $M_N = M_N(\kappa)$ as the largest M such that $S_N^M \neq \emptyset$. Then, the **capacity** of this model is defined as the ratio M_N/N, and models the maximum number of patterns this network can memorize per synapse.

Krauth and Mézard [KM89] analyzed this model using the (non-rigorous) replica method from statistical physics. They conjectured that as $N \to \infty$, the capacity concentrates around an explicit constant $\alpha_\star = \alpha_\star(\kappa)$, which is approximately 0.833 for $\kappa = 0$ and is formally defined in Proposition 2.3.2 below.[1] This was part of a series of works in the statistical physics literature [Gar87, GD88, Gar88, KM89, Méz89] which analyzed various perceptron models using the replica or cavity methods and put forward detailed predictions for their behavior. In particular, [KM89] provided a conjecture for the limiting capacity of the Ising perceptron, while [GD88] gave an analogous conjecture for the spherical perceptron, where the spins $\boldsymbol{x}$ belong to the sphere $\{\boldsymbol{x} \in \mathbb{R}^N : \|\boldsymbol{x}\| = \sqrt{N}\}$ instead of Σ_N.

Ding and Sun [DS18] proved that $\alpha_\star$ is a rigorous lower bound for the capacity, subject to a numerical condition that an explicit univariate function is maximized at 0.

Theorem 2.1.1. *[DS18, Theorem 1.1] Under Condition 1.2 therein, the following holds for the $\kappa = 0$ Ising perceptron. For any $\alpha < \alpha_\star$, $\liminf_{N\to\infty} \mathbb{P}(M_N/N \geq \alpha) > 0$.*

[1] [KM89] studied a model with Bernoulli disorder, i.e. where the g_i^a are i.i.d. samples from $\mathsf{unif}(\pm 1)$ rather than $\mathcal{N}(0, 1)$. As [NS23] shows this model's sharp threshold sequence is universal with respect to any subgaussian disorder, we may work with gaussian disorder for convenience.

Furthermore, [Xu21, NS23] showed that the capacity has a sharp threshold sequence, thereby improving the positive probability guarantee of Theorem 2.1.1 to high probability. Our main result is a matching upper bound for the capacity, subject to a similar numerical condition.

Theorem 2.1.2. *Under Condition 2.1.3 below, the following holds for the $\kappa = 0$ Ising perceptron. For any $\alpha > \alpha_\star$, $\lim_{N\to\infty} \mathbb{P}(M_N/N \geq \alpha) = 0$.*

Condition 2.1.3. The function $\mathscr{S}_\star(\lambda_1, \lambda_2)$ defined in (2.8) satisfies $\mathscr{S}_\star(\lambda_1, \lambda_2) \leq 0$ for all $\lambda_1, \lambda_2 \in \mathbb{R}$.

See Subsection 2.2.6 for a discussion of this condition. In particular $\mathscr{S}_\star(1, 0) = 0$ is a local maximum, and numerical plots suggest it is the unique global maximum.

Theorem 2.1.2 is a consequence of the more general Theorem 2.3.6, which states that $\alpha_\star(\kappa)$ upper bounds the capacity for general κ, under a number of numerical conditions depending on κ. The most complicated of these is Condition 2.1.3, and we derive Theorem 2.1.2 by verifying the remaining conditions when $\kappa = 0$. This computer-assisted verification is described in Appendix 2.B and carried out in the accompanying Python 3 file[2] using `python-flint`, a rigorous library for interval arithmetic.

2.1.1 Related work

For the spherical perceptron, the capacity threshold of [GD88] has been proved rigorously for all $\kappa \geq 0$ [ST03, Sto13a]. (See also [Sto13b] for some work on the $\kappa < 0$ case.) These works exploit the fact that the spherical perceptron with $\kappa \geq 0$ is a convex optimization problem. The Ising perceptron does not have this property, and our understanding of it is comparatively less complete. The replica heuristic also gives a prediction for the free energy of a positive-temperature version of this model [GD88, KM89], which was verified by [Tal00] at sufficiently high temperature using a rigorous version of the cavity method. The works [KR98, Tal99] showed that for the $\kappa = 0$ perceptron, there exists $\varepsilon > 0$ such that $\varepsilon \leq M_N/N \leq 1 - \varepsilon$ with high probability. The breakthrough work of Ding and Sun [DS18] showed that $\alpha_\star$ lower bounds the capacity for the $\kappa = 0$ perceptron, conditional on a numerical assumption. Very recently, [AT24] showed that 0.847 is a rigorous upper bound for the capacity in this model. Recent works have also shown the replica-symmetric formula for the free energy at low constraint density in generalized perceptron models [BNSX22], existence of a sharp threshold sequence [Xu21, NS23], and universality in the disorder [NS23]. We also mention the works [AS22, MZZ24] on algorithms for the negative spherical perceptron.

Another recent line of work originating with [APZ19] studied the **symmetric binary perceptron**, where the constraints in (2.1) are replaced by $|\langle \boldsymbol{g}^a, \boldsymbol{x}\rangle|/\sqrt{N} \leq \kappa$. Symmetry makes this model significantly more tractable (see Subsection 2.2.1 for more discussion); a series of remarkable works have established the limiting capacity [PX21, ALS22b], "frozen 1-RSB" structure [PX21], lognormal limit of partition function [ALS22b], and critical window [Alt23, SS23], and shed light on the performance of algorithms [ALS22a, GKPX22, GKPX23, BAKZ24].

2.1.2 Notation

While we introduce other parameters over the course of the proof, unless stated otherwise we send $N \to \infty$ first, treating the remaining parameters as small or large constants. Thus, we use $o_N(1)$ to denote a quantity vanishing with N, while notations like $o_\varepsilon(1)$ denote quantities independent of N tending to zero as the subscripted parameter tends to 0 or ∞ (which will be clear from context). We say an event occurs with high probability if it occurs with probability $1 - o_N(1)$. Further notations will be introduced in Subsection 2.4.1, before the main body of proofs.

Acknowledgements

I would like to thank Mehtaab Sawhney for pointing me to the reference [GZ00], and Will Perkins, Mehtaab Sawhney, Mark Sellke, and Nike Sun for helpful feedback on the manuscript. I am also grateful to Andrea Montanari and Huy Tuan Pham for a collaboration that inspired parts of this work. Thanks to Saba Lepsveridze for a helpful and motivating conversation.

[2] This file can be found here: `https://github.com/bricehuang/perceptron-interval-arithmetic/blob/main/perceptron-appendix-b-interval-arithmetic.py`

2.2 Further background and proof outline

This section contains a technical overview of the paper, and is organized as follows. In Subsection 2.2.1, we review the AMP-conditioned moment method used in [DS18] to prove the capacity lower bound and discuss the main difficulties of proving the upper bound. In Subsection 2.2.2, we outline a new approach based on reducing to a planted model and argue that if three primary inputs (R1), (R2), (R3) hold, then the upper bound reduces to a tractable moment computation. Subsection 2.2.3 discusses the most difficult input (R1), and Subsection 2.2.4 discusses the more straightforward inputs (R2) and (R3). Subsection 2.2.5 discusses related work involving planted models. Finally, Subsection 2.2.6 heuristically carries out the aforementioned moment computation, explains how Condition 2.1.3 emerges from it, and gives numerical evidence for Condition 2.1.3 when $\kappa = 0$.

2.2.1 AMP-conditioned moment method

A natural approach to studying the limiting capacity is the moment method. Let $M = \alpha N$, and let $\boldsymbol{G} \in \mathbb{R}^{M\times N}$ have rows $\boldsymbol{g}^1, \ldots, \boldsymbol{g}^M$. Then let $S_N(\boldsymbol{G}) = S_N^M$ (recall (2.1)) and $Z_N(\boldsymbol{G}) = |S_N(\boldsymbol{G})|$. If $\mathbb{E}[Z_N(\boldsymbol{G})] \ll 1$, then $S_N(\boldsymbol{G})$ is w.h.p. empty, and if $\mathbb{E}[Z_N(\boldsymbol{G})^2]/\mathbb{E}[Z_N(\boldsymbol{G})]^2$ is bounded, then $S_N(\boldsymbol{G})$ is nonempty with positive probability. If these two estimates hold for (respectively) $\alpha = \alpha_\star + \varepsilon$ and $\alpha = \alpha_\star - \varepsilon$, for any $\varepsilon > 0$, this shows the limiting capacity is $\alpha_\star$.

Let $\boldsymbol{m}_\star(\boldsymbol{G}) = \frac{1}{|S_N(\boldsymbol{G})|}\sum_{\boldsymbol{x}\in S_N(\boldsymbol{G})} \boldsymbol{x}$ denote the barycenter of the solution set $S_N(\boldsymbol{G})$. For models where $\boldsymbol{m}_\star(\boldsymbol{G}) = \boldsymbol{0}$, such as the symmetric binary perceptron [APZ19, PX21, ALS22b], this two-moment analysis often suffices to determine the limiting capacity. However, due to the asymmetry of the activation in the present model, $\boldsymbol{m}_\star(\boldsymbol{G})$ is typically macroscopic and random. It is expected that for any $\alpha > 0$, large-deviations events in the location of $\boldsymbol{m}_\star(\boldsymbol{G})$ dominate the first and second moments. Thus $Z_N(\boldsymbol{G})$ is typically exponentially smaller than $\mathbb{E}[Z_N(\boldsymbol{G})]$, and $\mathbb{E}[Z_N(\boldsymbol{G})]^2$ exponentially smaller than $\mathbb{E}[Z_N(\boldsymbol{G})^2]$, which causes the moment method to fail. For example, for the $\kappa = 0$ perceptron, $\frac{1}{N}\log \mathbb{E}[Z_N(\boldsymbol{G})]$ crosses zero at $\alpha = 1$, larger than $\alpha_\star(0) \approx 0.833$.

To overcome this difficulty, [DS18] and [Bol19] (the latter for the Sherrington–Kirkpatrick model) concurrently developed a conditional moment method, in which one conditions on a suitable proxy for $\boldsymbol{m}_\star(\boldsymbol{G})$ before computing moments. The conditioning step effectively recenters spins around $\boldsymbol{m}_\star(\boldsymbol{G})$, after which the moment method can potentially succeed.

The choice of conditioning is motivated by the TAP heuristic [TAP77] from statistical physics, which provides a powerful but non-rigorous framework to study this and other mean-field models. The central object in this framework is a **TAP free energy** $\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m}, \boldsymbol{n})$, which is defined in (2.15) and can be thought of as a mean-field (dense graph) limit of the Bethe free energy of an appropriate message-passing system. It is expected that $\mathcal{F}_{\mathsf{TAP}}$ has a unique stationary point $(\boldsymbol{m}, \boldsymbol{n}) \in [-1, 1]^N \times \mathbb{R}^M$, with the following interpretation: $\boldsymbol{m}$ approximates the barycenter $\boldsymbol{m}_\star(\boldsymbol{G})$ of $S_N(\boldsymbol{G})$, and for each $a \in [M]$, n_a approximates a function of the average slack of the constraint $\langle \boldsymbol{g}^a, \boldsymbol{x}\rangle/\sqrt{N} \ge \kappa$ over solutions $\boldsymbol{x} \in S_N(\boldsymbol{G})$.[3] It is also predicted that $\boldsymbol{m}$ and $\boldsymbol{n}$ have specific coordinate profiles: for $(q_\star, \psi_\star)$ defined as the fixed point of a scalar recursion (see Condition 2.3.1) and $F = F_{1-q_\star}$ as in (2.13), the prediction is that the coordinates of $\dot{\boldsymbol{h}} = \mathrm{th}^{-1}(\boldsymbol{m})$ and $\widehat{\boldsymbol{h}} = F^{-1}(\widehat{\boldsymbol{h}})$ have empirical distribution approximating $\mathcal{N}(0, \psi_\star)$ and $\mathcal{N}(0, q_\star)$.[4]

An important fact we will exploit is that for fixed $(\boldsymbol{m}, \boldsymbol{n})$, the stationarity condition $\nabla \mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m}, \boldsymbol{n}) = \boldsymbol{0}$ can be written as two **linear** equations in $\boldsymbol{G}$. These are the **TAP equations**, defined in (2.16). Using this fact, we can define a **planted model**, which plays an important motivational role in [DS18, Bol19]: we first chooose $(\boldsymbol{m}, \boldsymbol{n})$ with aforementioned coordinate profile, and then sample $\boldsymbol{G}$ conditional on $\nabla \mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m}, \boldsymbol{n}) = \boldsymbol{0}$. (This is different from the more well-known notion of planted model introduced in [AC08], in that we are planting a TAP fixed point rather than a satisfying assignment; see Subsection 2.2.5 for further discussion.)

If we imagine for a moment that $\boldsymbol{G}$ were sampled from this planted model, then the moment method becomes tractable. In this model, the law of $\boldsymbol{G}$ conditional on $(\boldsymbol{m}, \boldsymbol{n})$ remains gaussian because the TAP equa-

[3] More generally, the statistical physics literature predicts that the Gibbs measure — here, the uniform measure on $S_N(\boldsymbol{G})$ — decomposes as a convex combination of well-concentrated "pure states," whose barycenters each approximate a stationary point of the TAP free energy [MPV87]. The present model is expected to be replica symmetric, meaning the entire Gibbs measure is one pure state.

[4] Here and throughout, nonlinearities such as th^{-1} and F^{-1} are applied coordinate-wise.

tions are linear in $\boldsymbol{G}$, and the conditional first and second moments of $Z_N(\boldsymbol{G})$ can be computed. They amount to tractable $O(1)$-dimensional optimization problems: for example, computing $\mathbb{E}[Z_N(\boldsymbol{G})|\boldsymbol{m},\boldsymbol{n}]$ amounts to optimizing the exponential-order contribution to the first moment from subsets of Σ_N defined by their inner products with $\boldsymbol{m}$ and $\dot{\boldsymbol{h}}$ (see Subsection 2.2.6 for details). The planted model removes the main difficulty of the macroscopically-fluctuating barycenter, giving the moment method a chance to succeed.

However, this planted model is different from the true model, in which the TAP solution $(\boldsymbol{m},\boldsymbol{n})$ depends on $\boldsymbol{G}$ in a complicated way. It is a priori unclear that these can be rigorously linked, because in the true model both existence and uniqueness of the TAP solution are not known. To carry out this approach, [DS18, Bol19] instead condition on a sequence of **approximate message passing** (AMP) iterates $(\boldsymbol{m}^0,\boldsymbol{n}^0,\ldots,\boldsymbol{m}^k,\boldsymbol{n}^k)$ whose dependence on $\boldsymbol{G}$ is explicit. The AMP iteration was introduced in [Bol14, BM11], and is defined (roughly speaking, see (2.17)) by iterating the TAP equations. Its behavior can be understood through the powerful state evolution description of [Bol14, BM11, JM13, BMN20]: for any k not growing with N, state evolution exactly characterizes the limiting overlap structure of $(\boldsymbol{m}^0,\ldots,\boldsymbol{m}^k)$ and $(\boldsymbol{n}^0,\ldots,\boldsymbol{n}^k)$. Using this description, it can be shown that the AMP iterates converge to an approximate stationary point of $\mathcal{F}_{\mathsf{TAP}}$:

$$\lim_{k_1,k_2\to\infty} \operatorname*{p-lim}_{N\to\infty} N^{-1/2}\|(\boldsymbol{m}^{k_1},\boldsymbol{n}^{k_1})-(\boldsymbol{m}^{k_2},\boldsymbol{n}^{k_2})\| = \lim_{k\to\infty} \operatorname*{p-lim}_{N\to\infty} N^{-1/2}\|\nabla\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m}^k,\boldsymbol{n}^k)\| = 0. \tag{2.2}$$

Here p-lim denotes limit in probability. It is in this sense that the AMP iterates are a proxy for $(\boldsymbol{m},\boldsymbol{n})$.

While the main advantages of conditioning on the AMP filtration are explicit dependence on $\boldsymbol{G}$ and state evolution, the main disadvantage is the greater complexity of the resulting moment calculation. Although the law of $\boldsymbol{G}$ conditional on $(\boldsymbol{m}^0,\boldsymbol{n}^0,\ldots,\boldsymbol{m}^k,\boldsymbol{n}^k)$ remains gaussian, the conditional first and second moments of $Z_N(\boldsymbol{G})$ are now $O(k)$-dimensional optimization problems, in which one optimizes over subsets of Σ_N defined by their inner products with $\boldsymbol{m}^0,\ldots,\boldsymbol{m}^k$ and related vectors. These problems are not in general tractable. We note that [Bol19, BNSX22] successfully carry out this optimization in their respective settings, but only at sufficiently high temperature or low constraint density.

An important insight of [DS18] is that this approach still gives a tractable proof of the capacity lower bound, because — to show a lower bound for $Z_N(\boldsymbol{G})$ — one may truncate $Z_N(\boldsymbol{G})$ before computing moments. They construct a truncation $\widetilde{Z}_N(\boldsymbol{G})$ of $Z_N(\boldsymbol{G})$, restricting (among other conditions) to $\boldsymbol{x}\in\Sigma_N$ with prescribed inner products with $\boldsymbol{m}^0,\ldots,\boldsymbol{m}^k$. The conditional first moment of $\widetilde{Z}_N(\boldsymbol{G})$ is then explicit, while the conditional second moment becomes a 1-dimensional optimization. [DS18] shows that (under the aforementioned numerical condition) $\mathbb{E}[\widetilde{Z}_N(\boldsymbol{G})^2]/\,\mathbb{E}[\widetilde{Z}_N(\boldsymbol{G})]^2$ is bounded for any $\alpha<\alpha_\star$, which implies the capacity lower bound.

We mention that [BY22, BNSX22] carry out similar truncated second moment arguments in their respective settings, and the former improves the parameter regime where the method of [Bol19] obtains the replica symmetric free energy lower bound for the Sherrington–Kirkpatrick model.

The main difficulty of the capacity upper bound is that truncation is no longer available. Without it, proving the capacity upper bound within the AMP-conditioned moment method would require solving the above $O(k)$-dimensional optimization problem, which does not appear to be tractable.

2.2.2 Approximate contiguity with planted model

Our proof revisits and justifies the planted model heuristic described above, where we select $(\boldsymbol{m},\boldsymbol{n})$ with appropriate coordinate profile and generate $\boldsymbol{G}$ conditional on $\nabla\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n})=\boldsymbol{0}$. We will show that the true model is approximately contiguous to the planted model, in the sense of (2.3) below. So, rather than conditioning on the AMP filtration, we can condition directly on $(\boldsymbol{m},\boldsymbol{n})$ after all. The conditional first moment of $Z_N(\boldsymbol{G})$ then reverts to a simple optimization in two, rather than $O(k)$, dimensions. This makes the capacity upper bound tractable.

The idea of passing by contiguity to a model with a planted TAP solution is also used in simultaneous joint work with A. Montanari and H. T. Pham [HMP24], on sampling from the Gibbs measure of a spherical mixed p-spin glass in total variation by an algorithmic implementation of stochastic localization [Eld20b, AMS22]. A similar inequality to (2.3) appears as Proposition 4.4(d) therein. However, these two papers differ in both how this reduction is used, and how it is proved. While [HMP24] develops a reduction similar to (2.3), its main focus is to compute a high-precision estimate for the mean of a Gibbs measure, and the reduction to a planted model arises as a step in the analysis of this estimator. In the present paper, the reduction (2.3) is itself the main technical step, but the proof of it is also more challenging. Most notably, a key ingredient in the proof of (2.3), in both the present paper and [HMP24], is the uniqueness of the TAP fixed point in a certain region, see (R1) below. Whereas this ingredient is available in the spin glass setting of [HMP24] from known results, showing it in our setting requires new ideas, described in detail in Subsection 2.2.3.

We now state the approximate contiguity estimate. For small $\upsilon > 0$, let $\mathcal{S}_\upsilon$ denote the set of $(\boldsymbol{m}, \boldsymbol{n})$ whose coordinate profile is υ-close (in a suitable metric, see (2.27)) to that predicted by the TAP heuristic. We will show, roughly speaking, that there exists $C = O(1)$ such that for any $\boldsymbol{G}$-measurable event $\mathscr{E}$,

$$\mathbb{P}(\mathscr{E}) \le C \sup_{(\boldsymbol{m},\boldsymbol{n})\in\mathcal{S}_\upsilon} \mathbb{P}(\mathscr{E} | \nabla \mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n}) = \mathbf{0})^{1/2} + o_N(1). \tag{2.3}$$

Remark 2.2.1. For reasons described below, we actually prove (2.3) for perturbations $\mathcal{F}^{\varepsilon}_{\mathsf{TAP}}$, $\mathcal{S}_{\varepsilon,\upsilon}$ of $\mathcal{F}_{\mathsf{TAP}}$, $\mathcal{S}_\upsilon$, and this qualification holds for the entire discussion below, even where not stated. These perturbations are defined in (2.24) and (2.27), and the formal version of (2.3) is given in Lemma 2.3.8.

We then take $\mathscr{E} = \{S_N(\boldsymbol{G}) \neq \emptyset\}$. The first moment bound will show that (under Condition 2.1.3) this event has vanishing probability in the planted model for any $\alpha > \alpha_\star$. Then (2.3) implies the conclusion.

Next, we discuss the proof of (2.3). The following two central ingredients establish uniqueness and existence of the critical point of $\mathcal{F}_{\mathsf{TAP}}$ within the set $\mathcal{S}_\upsilon$, with high probability in the true model.

(R1) The expected number of critical points of $\mathcal{F}_{\mathsf{TAP}}$ in $\mathcal{S}_\upsilon$ is $1 + o(1)$.

(R2) With high probability, there exists a critical point of $\mathcal{F}_{\mathsf{TAP}}$ in $\mathcal{S}_\upsilon$.

Remark 2.2.2. Although the TAP perspective predicts $\mathcal{F}_{\mathsf{TAP}}$ has a unique critical point in the full input space, uniqueness in $\mathcal{S}_\upsilon$ (and for the perturbed $\mathcal{F}^{\varepsilon}_{\mathsf{TAP}}$) suffices for our proof.

A short argument based on the Kac–Rice formula [Kac48, Ric44] (see [AT09, Theorem 11.2.1] for a textbook treatment) shows that (2.3) follows from (R1), (R2), and the following additional input, which is a concentration condition on the change of volume term $|\det \nabla^2 \mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n})|$ in the Kac–Rice formula. This argument is carried out in the proof of Lemma 2.3.8, see (2.33).

(R3) There exists $C' = O(1)$ such that uniformly over $(\boldsymbol{m},\boldsymbol{n}) \in \mathcal{S}_\upsilon$,

$$\mathbb{E}[|\det \nabla^2 \mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n})|^2 \big| \nabla \mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n}) = \mathbf{0}]^{1/2} \le C' \, \mathbb{E}[|\det \nabla^2 \mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n})| \big| \nabla \mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n}) = \mathbf{0}].$$

Remark 2.2.3. Since the probability in (2.3) is exponentially small, the proof can be carried out with $e^{o(N)}$ in place of C in (2.3). Consequently, showing (R1) and (R3) with $e^{o(N)}$ in place of $1+o(1)$, $O(1)$ also suffices.

Input (R2) is proved constructively, by showing that AMP finds a critical point in the following sense.

(R4) There exists $r_k = o_k(1)$ such that with high probability, $\mathcal{F}_{\mathsf{TAP}}$ has a unique critical point in a $r_k\sqrt{N}$-neighborhood of the AMP iterate $(\boldsymbol{m}^k, \boldsymbol{n}^k)$ (which lies in $\mathcal{S}_\upsilon$ by state evolution), for each sufficiently large k.

Input (R3) will follow from a classic spectral concentration argument of [GZ00]. We next discuss the proofs of (R1), (R4) and (R3), in that order.

2.2.3 Topological trivialization of TAP free energy

Condition (R1) is the most important input to the proof of (2.3). It is related to a remarkable line of work pioneered by [Fyo04, ABČ13], on the landscapes of random high-dimensional functions. This line of work has obtained expected critical point counts in a variety of settings, including spherical p-spin glasses [AB13, ABČ13] (see [Sub17a, AG20, SZ21, BSZ20, HS23b] for matching second moment estimates in certain cases) spiked tensor models [BMMN19, ABL22], the TAP free energy for $\mathbb{Z}_2$-synchronization [FMM21, CFM23], bipartite spin glasses [Kiv23, McK24], the elastic manifold [BBM24], and generalized linear models [MBB20]. We also refer the reader to earlier non-rigorous work on this topic from the statistical physics literature [BM80, PP95, CLR05].

One phenomenon studied in these works is **topological trivialization** [FLD14, Fyo15, BČNS22, HS23c], a phase transition where the number of critical points drops from e^{cN} to $e^{o(N)}$, or often $O(1)$. Proving (R1) amounts to showing **annealed topological trivialization** for $\mathcal{F}^{\varepsilon}_{\mathsf{TAP}}$ on $\mathcal{S}_{\varepsilon,\upsilon}$.

The strategy of these works is to calculate the expected number of critical points using the Kac–Rice formula, evaluating the integrand using random matrix theory. Usually, the most complicated term in the integrand is the expected absolute value of the determinant of a random matrix. The most well-understood application is where the landscape is a spherical mixed p-spin glass, in which case this random matrix is a GOE shifted by a scalar multiple of the identity. For this case, an exact formula for this expected absolute determinant is known, see [ABČ13, Lemma 3.3]. This makes the Kac–Rice calculation explicit and tractable. In particular, [Fyo15, BČNS22] use this approach to determine the topologically trivial phase of spherical mixed p-spin glasses, and [HMP24] uses these results to establish (R1) for its application. However, for other models, results on topological trivialization are not as readily available.

It may still be possible to show (R1) for our model in this way, by evaluating the more general random determinant that appears in the Kac–Rice formula. This is the approach taken by [FMM21] which, for $\mathbb{Z}_2$-synchronization at sufficiently large signal, shows annealed trivialization of suitably low-energy TAP solutions. Their method bounds the random determinant in the Kac–Rice formula using free probability [Voi91]. Furthermore, [BBM23] introduced a general tool for studying random determinants, showing that under mild conditions, their exponential order is the integral of $\log|\lambda|$ against the random matrix's limiting spectral measure. The spectral measure can then be studied using free probability.

Using this approach, one can often express the exponential order of the expected number of critical points as a variational formula, in which one term is an implicitly-defined function arising from free probability [Kiv23, HS23c, BBM24, McK24]. This yields a plausible way to show (R1): if we can show the variational formula for our model has value zero, annealed trivialization follows (in the sense of $e^{o(N)}$ expected critical points, which suffices by Remark 2.2.3). Recently, [HS23c] showed that this method can be carried out for multi-species spherical spin glasses, and it in fact characterizes the topologically trivial phase. Nonetheless, the variational formula is highly model-dependent — the proof in [HS23c] relies on a detailed understanding of a vector Dyson equation — and it is unclear if this method can be carried out for our model.

We instead show annealed topological trivialization by a different, and arguably more conceptual, approach. We will show that (R1) follows from the following variant of (R4):

(R5) In a model where we plant a stationary point $(\boldsymbol{m},\boldsymbol{n}) \in \mathcal{S}_{\varepsilon,\upsilon}$ of $\mathcal{F}^{\varepsilon}_{\mathsf{TAP}}$ (i.e. condition on $\nabla\mathcal{F}^{\varepsilon}_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n}) = \boldsymbol{0}$), the same AMP iteration finds $(\boldsymbol{m},\boldsymbol{n})$, in the sense of (R4), with high probability.

This implication is proved in Lemma 2.4.15. Heuristically, the reason (R5) implies (R1) is that any realization of the disorder where $\mathcal{F}^{\varepsilon}_{\mathsf{TAP}}$ has $T > 1$ stationary points in $\mathcal{S}_{\varepsilon,\upsilon}$ can arise in T different planted models, and the event in (R5) can hold in only one of these T realizations. If the expected number of critical points is too large, (R5) cannot occur with the stated probability. The input (R5) can be proved by similar methods as (R4), as described in the next subsection. This method yields the first proof of topological trivialization that does not directly evaluate the Kac–Rice formula. We believe this is interesting in its own right.

2.2.4 Critical point near late AMP iterates and determinant concentration

This subsection discusses inputs (R4), (R5), and (R3), in that order. As state evolution ensures $\|\nabla\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m}^k,\boldsymbol{n}^k)\| = o_k(1)\sqrt{N}$ (recall (2.2)), (R4) holds if, for example, $\mathcal{F}_{\mathsf{TAP}}$ is C-strongly concave in a neighborhood of late AMP iterates for $C > 0$ independent of k. Recent works in the variational inference literature [CFM23, CFLM23, Cel24] develop tools to establish this local concavity, and using them prove analogs of (R4) in several models.

In our setting, the fact that $\mathcal{F}_{\mathsf{TAP}}$ is **not** strongly concave near late AMP iterates introduces some complications. In fact, $\mathcal{F}_{\mathsf{TAP}}$ is strongly concave in $\boldsymbol{m}$, but convex — and problematically, not strongly convex — in $\boldsymbol{n}$. This issue is one reason we carry out the argument on a perturbation $\mathcal{F}^{\varepsilon}_{\mathsf{TAP}}$ of $\mathcal{F}_{\mathsf{TAP}}$, and a similarly perturbed AMP iteration and set $\mathcal{S}_{\varepsilon,\upsilon}$. (This perturbation serves several other purposes as well, described in Remark 2.4.5.) We will show that near late AMP iterates, $\mathcal{F}^{\varepsilon}_{\mathsf{TAP}}$ is strongly convex in $\boldsymbol{n}$ and $\mathcal{G}^{\varepsilon}_{\mathsf{TAP}}(\boldsymbol{m}) \equiv \inf_{\boldsymbol{n}} \mathcal{F}^{\varepsilon}_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n})$ is strongly concave, which is enough to imply (R4). Strong convexity of $\mathcal{F}^{\varepsilon}_{\mathsf{TAP}}$ in $\boldsymbol{n}$ holds (deterministically) essentially by construction.

Our proof of local strong concavity of $\mathcal{G}^{\varepsilon}_{\mathsf{TAP}}$ uses an idea introduced in [Cel24], to bound the Hessian at a late AMP iterate by applying a gaussian comparison inequality conditionally on the AMP iterates. [Cel24] considers a setting where AMP is performed on disorder $\boldsymbol{W} \sim \mathsf{GOE}(N)$ and the relevant Hessian is of the form $\boldsymbol{A} + \boldsymbol{W}$, where $\boldsymbol{A}$ is a function of a late AMP iterate. He develops a method to upper bound the top eigenvalue of this matrix by applying the Sudakov–Fernique inequality [Sud71, Fer75, Sud79] to the part of $\boldsymbol{W}$ that remains random after observing the AMP iterates. For us, the Hessian takes the form

$$\nabla^2 \mathcal{G}^{\varepsilon}_{\mathsf{TAP}}(\boldsymbol{m}, \boldsymbol{n}) = \boldsymbol{A}_1 + \frac{1}{N} \boldsymbol{G}^\top \boldsymbol{A}_2 \boldsymbol{G} + \boldsymbol{\Delta}, \tag{2.4}$$

where $\boldsymbol{A}_1, \boldsymbol{A}_2$ are functions of $(\boldsymbol{m}, \boldsymbol{n})$, and $\boldsymbol{\Delta}$ is a low-rank term depending on both $\boldsymbol{G}$ and $(\boldsymbol{m}, \boldsymbol{n})$. We can arrange $\mathcal{F}^{\varepsilon}_{\mathsf{TAP}}$ so that $\boldsymbol{\Delta}$ does not contribute to the top eigenvalue. However, the post-AMP Sudakov–Fernique inequality does not apply to the remaining part, because — unlike for a GOE matrix — the quadratic form induced by $\boldsymbol{G}^\top \boldsymbol{A}_2 \boldsymbol{G}$ is not a gaussian process. We instead recast the top eigenvalue as a minimax program, via the identity (for $\boldsymbol{A}_2 \prec 0$)

$$\lambda_{\max}\left(\boldsymbol{A}_1 + \frac{1}{N} \boldsymbol{G}^\top \boldsymbol{A}_2 \boldsymbol{G}\right) = \sup_{\|\dot{\boldsymbol{v}}\|=1} \inf_{\widehat{\boldsymbol{v}} \in \mathbb{R}^M} \left\{ \langle \dot{\boldsymbol{v}}, \boldsymbol{A}_1 \dot{\boldsymbol{v}} \rangle - \langle \widehat{\boldsymbol{v}}, \boldsymbol{A}_2^{-1} \widehat{\boldsymbol{v}} \rangle + \frac{2}{\sqrt{N}} \langle \widehat{\boldsymbol{v}}, \boldsymbol{G} \dot{\boldsymbol{v}} \rangle \right\}.$$

This can be bounded by Gordon's inequality [Gor85, Gor88] conditional on the AMP iterates. Interestingly, the bound obtained in this way is sharp, matching a lower bound for the top eigenvalue obtained by free probability (see Remark 2.6.15).

The input (R5) follows similarly to (R4). We will show that with high probability over the planted model, late AMP iterates are approximate critical points of $\mathcal{F}^{\varepsilon}_{\mathsf{TAP}}$, near which $\mathcal{F}^{\varepsilon}_{\mathsf{TAP}}(\boldsymbol{m}, \cdot)$ is strongly convex and $\mathcal{G}^{\varepsilon}_{\mathsf{TAP}}$ is strongly concave. While the law of the disorder is different under the planted model, it remains gaussian and a similar analysis can be carried out.

We turn to (R3). An argument of [GZ00] implies that if a symmetric $\boldsymbol{X} \in \mathbb{R}^{N \times N}$ has independent (not necessarily centered or identically distributed) entries on and above the diagonal with uniformly bounded log-Sobolev constant, then $\frac{1}{\sqrt{N}} \boldsymbol{X}$ enjoys a strong spectral concentration property: any 1-Lipschitz spectral trace has $O(1)$-scale subgaussian fluctuations. We will see that conditional on $\nabla \mathcal{F}^{\varepsilon}_{\mathsf{TAP}}(\boldsymbol{m}, \boldsymbol{n}) = \boldsymbol{0}$, $\det \nabla^2 \mathcal{F}^{\varepsilon}_{\mathsf{TAP}}(\boldsymbol{m}, \boldsymbol{n})$ is a nonrandom multiple of $\det \nabla^2 \mathcal{G}^{\varepsilon}_{\mathsf{TAP}}(\boldsymbol{m}, \boldsymbol{n})$, which has form (2.4). The entries of this matrix are not independent, but we can rewrite it via the classical trick

$$\det\left(\boldsymbol{A}_1 + \frac{1}{N} \boldsymbol{G}^\top \boldsymbol{A}_2 \boldsymbol{G}\right) = \det \boldsymbol{X}, \qquad \boldsymbol{X} = \begin{bmatrix} \boldsymbol{A}_1 & \frac{1}{\sqrt{N}} \boldsymbol{G}^\top \\ \frac{1}{\sqrt{N}} \boldsymbol{G} & -\boldsymbol{A}_2^{-1} \end{bmatrix}. \tag{2.5}$$

Conditional on $\nabla \mathcal{F}^{\varepsilon}_{\mathsf{TAP}}(\boldsymbol{m}, \boldsymbol{n}) = \boldsymbol{0}$, the matrices $\boldsymbol{A}_1, \boldsymbol{A}_2$ are nonrandom while $\boldsymbol{G}$ has a (noncentered) gaussian law. Thus the result of [GZ00] applies to $\boldsymbol{X}$. (A slightly more elaborate version of (2.5) also accounts for the random low-rank spike $\boldsymbol{\Delta}$ in (2.4), see (2.76).)

From the above discussion, conditional on $\nabla \mathcal{F}^{\varepsilon}_{\mathsf{TAP}}(\boldsymbol{m}, \boldsymbol{n}) = \boldsymbol{0}$, $\mathcal{F}^{\varepsilon}_{\mathsf{TAP}}(\boldsymbol{m}, \cdot)$ is strongly convex near $\boldsymbol{n}$ and $\mathcal{G}^{\varepsilon}_{\mathsf{TAP}}$ is w.h.p. strongly concave near $\boldsymbol{m}$. This implies that the spectrum of $\nabla^2 \mathcal{F}^{\varepsilon}_{\mathsf{TAP}}(\boldsymbol{m}, \boldsymbol{n})$, and thus $\boldsymbol{X}$, is bounded away from zero, and provides the final ingredient to prove (R3): since $x \mapsto \log|x|$ is $O(1)$-Lipschitz away from zero, $\log|\det \boldsymbol{X}|$ is approximately a $O(1)$-Lipschitz spectral trace, which has $O(1)$-scale subgaussian fluctuations by [GZ00].

Remark 2.2.4. The fact that this log determinant has $O(1)$-scale fluctuations is only possible because the spectrum is bounded away from zero. For Wigner or Ginibre matrices, two examples of random matrices whose limiting bulk spectrum does include zero, the log determinant is known to have $\Theta(\sqrt{\log N})$ fluctuations [TV12, NV14], which diverges with N.

2.2.5 On planted models

Reducing to a planted model is a powerful tool in the analysis of random functions. This technique was introduced in the seminal work [AC08] and has seen a wide range of applications in the past decade. The underlying idea is to show contiguity of the original model with a planted version, defined as the null model conditioned on having a particular (randomly chosen) solution. If this holds, properties of the null model can be deduced from the planted version, which is often easier to analyze.

A frequent application of this method is to probe the local landscape around a typical solution. This is the original application of [AC08]: contiguity implies that the landscape around a typical solution to the null model can be approximated by the landscape around the planted solution in the planted model. From this, [AC08] shows the existence of a shattering transition in several random constraint satisfaction problems. This approach has since also been used to show "frozen 1RSB" structure in the symmetric binary perceptron [PX21, ALS22b] and shattering in the Gibbs measures of spherical spin glasses [AMS25]. In a similar spirit, [HMP24] passes to a model with a planted TAP solution to obtain a high-precision estimate of the magnetization of a spherical spin glass.

In other applications, including the present work, the object of interest is not the local landscape, but the planted model is nonetheless simpler to analyze than the null model. Such applications include the RS free energy of random constraint satisfaction problems [BC16, BCH+16, CKPZ17, CEJ+18, CKM20], the 1RSB free energy of random regular NAE-SAT [SSZ22], and the Parisi formula for spherical spin glasses in the RS and zero-temperature 1RSB phases [HS23b]. Passage to a planted model is also a crucial tool in the analysis of sampling algorithms based on stochastic localization [AMS22, AMS23b].

2.2.6 First moment in planted model

In this subsection, we give a heuristic calculation of the first moment of $Z_N(\boldsymbol{G})$ in the planted model. The function $\mathscr{S}_\star(\lambda_1,\lambda_2)$ appearing in Condition 2.1.3 arises from this calculation, and under this condition the first moment method succeeds. At the end of this subsection, we also give numerical evidence for Condition 2.1.3 when $\kappa = 0$.

We work at constraint density $\alpha_\star$, setting $M = \lfloor \alpha_\star N \rfloor$ and $\boldsymbol{G}, S_N(\boldsymbol{G}), Z_N(\boldsymbol{G})$ as above with this M. Let $\mathbb{P}^{\boldsymbol{m},\boldsymbol{n}}_{\mathsf{Pl}}$ and $\mathbb{E}^{\boldsymbol{m},\boldsymbol{n}}_{\mathsf{Pl}}$ denote probability and expectation w.r.t. the model conditional on $\nabla \mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n}) = \boldsymbol{0}$. We will argue that under Condition 2.1.3, $\mathbb{E}^{\boldsymbol{m},\boldsymbol{n}}_{\mathsf{Pl}} Z_N(\boldsymbol{G}) = e^{o(N)}$. Then, at any constraint density $\alpha > \alpha_\star$, the $(\alpha-\alpha_\star)N$ additional constraints will make this moment exponentially small.

This argument will be made rigorous in Section 2.7. Per the above discussion, the rigorous version of this argument will plant a critical point of $\mathcal{F}^\varepsilon_{\mathsf{TAP}}$ rather than $\mathcal{F}_{\mathsf{TAP}}$.

We first define the function $\mathscr{S}_\star$. Let $(q_0,\psi_0) = (q_\star(\alpha_\star,\kappa), \psi_\star(\alpha_\star,\kappa))$ be defined by Condition 2.3.1. As discussed in Subsection 2.2.1, these are the variances of the (gaussian) coordinate empirical measures of $\widehat{\boldsymbol{h}}$, $\dot{\boldsymbol{h}}$ predicted by the TAP heuristic, at constraint density $\alpha_\star$. Let $\dot{\boldsymbol{H}} \sim \mathcal{N}(0,\psi_0)$ and $\hat{\boldsymbol{H}} \sim \mathcal{N}(0,q_0)$. These two random variables may be defined on different probability spaces, as all expectations in the below formulas will involve random variables from only one space. Let $\boldsymbol{M} = \mathrm{th}(\dot{\boldsymbol{H}})$ and $\boldsymbol{N} = F_{1-q_0}(\hat{\boldsymbol{H}})$. For any measurable $\boldsymbol{\Lambda} : \mathbb{R} \to [-1,1]$, define

$$\mathsf{ent}(\boldsymbol{\Lambda}) = \mathbb{E}\,\mathcal{H}\left(\frac{1+\boldsymbol{\Lambda}(\dot{\boldsymbol{H}})}{2}\right), \tag{2.6}$$

where $\mathcal{H}(x) = -x\log x - (1-x)\log(1-x)$ is the binary entropy function. Let Ψ be the complementary gaussian cumulative density function defined in (2.12). For $s \ge 0$, define

$$\mathscr{S}_\star(\boldsymbol{\Lambda},s) = \frac{1}{2}s^2\psi_0 + \mathsf{ent}(\boldsymbol{\Lambda}) + \alpha_\star\,\mathbb{E}\log\Psi\left\{\frac{\kappa - \frac{\mathbb{E}[\boldsymbol{M}\boldsymbol{\Lambda}(\dot{\boldsymbol{H}})]}{q_0}\hat{\boldsymbol{H}} - \frac{\mathbb{E}[\dot{\boldsymbol{H}}\boldsymbol{\Lambda}(\dot{\boldsymbol{H}})]}{\psi_0}\boldsymbol{N}}{\sqrt{1-\frac{\mathbb{E}[\boldsymbol{M}\boldsymbol{\Lambda}(\dot{\boldsymbol{H}})]^2}{q_0}}} + s\boldsymbol{N}\right\}. \tag{2.7}$$

Finally, let $\boldsymbol{\Lambda}_{\lambda_1,\lambda_2}(x) = \mathrm{th}(\lambda_1 x + \lambda_2 \mathrm{th}(x))$ and define

$$\mathscr{S}_\star(\boldsymbol{\Lambda}) = \inf_{s\ge 0}\mathscr{S}_\star(\boldsymbol{\Lambda},s), \qquad\qquad \mathscr{S}_\star(\lambda_1,\lambda_2) = \mathscr{S}_\star(\boldsymbol{\Lambda}_{\lambda_1,\lambda_2}). \tag{2.8}$$

These quantities have the following physical meanings. $\dot{\boldsymbol{H}}, \hat{\boldsymbol{H}}, \boldsymbol{M}, \boldsymbol{N}$ are the coordinate distributions of $\dot{\boldsymbol{h}}, \widehat{\boldsymbol{h}}, \boldsymbol{m}, \boldsymbol{n}$. $\boldsymbol{\Lambda}$ specifies a set $\Sigma_N(\boldsymbol{\Lambda}) \subseteq \Sigma_N$ of points $\boldsymbol{x}$ where x_i has "conditional average" $\boldsymbol{\Lambda}(\dot{h}_i)$, in the sense that (informally, see (2.81))

$$\frac{1}{\#\{i\in[N] : \dot{h}_i \approx \dot{h}\}}\sum_{i\in[N]:\dot{h}_i\approx\dot{h}} x_i \approx \boldsymbol{\Lambda}(\dot{h}), \qquad \forall \dot{h}\in\mathbb{R}. \tag{2.9}$$

Note that $\mathsf{ent}(\mathbf{\Lambda})$ is the entropy of this set, that is (see Lemma 2.7.6)

$$\frac{1}{N}\log|\Sigma_N(\mathbf{\Lambda})| \simeq \mathsf{ent}(\mathbf{\Lambda}). \tag{2.10}$$

Here and throughout, $\simeq$ denotes equality up to additive $o_N(1)$.

Let $Z_N(\boldsymbol{G},\mathbf{\Lambda}) = |S_N(\boldsymbol{G})\cap\Sigma_N(\mathbf{\Lambda})|$ denote the number of solutions with profile $\mathbf{\Lambda}$. We will see that for all $s\ge 0$, $\mathscr{S}_\star(\mathbf{\Lambda},s)$ upper bounds the exponential order of $\mathbb{E}^{\boldsymbol{m},\boldsymbol{n}}_{\mathsf{Pl}} Z_N(\boldsymbol{G},\mathbf{\Lambda})$. Thus $\mathscr{S}_\star(\mathbf{\Lambda})$ also upper bounds this quantity, and $\mathbb{E}^{\boldsymbol{m},\boldsymbol{n}}_{\mathsf{Pl}} Z_N(\boldsymbol{G})$ is bounded (heuristically) by Laplace's principle:

$$\frac{1}{N}\log\mathbb{E}^{\boldsymbol{m},\boldsymbol{n}}_{\mathsf{Pl}} Z_N(\boldsymbol{G}) \simeq \sup_{\mathbf{\Lambda}}\left\{\frac{1}{N}\log\mathbb{E}^{\boldsymbol{m},\boldsymbol{n}}_{\mathsf{Pl}} Z_N(\boldsymbol{G},\mathbf{\Lambda})\right\} \le \sup_{\mathbf{\Lambda}}\mathscr{S}_\star(\mathbf{\Lambda}) + o_N(1).$$

While this supremum is a priori an infinite-dimensional optimization problem, the following observation reduces it to two dimensions. For any a_1,a_2, a Lagrange multipliers calculation (see Lemma 2.7.10) shows that the maximum of $\mathsf{ent}(\mathbf{\Lambda})$ subject to $\mathbb{E}[\dot{\boldsymbol{H}}\mathbf{\Lambda}(\dot{\boldsymbol{H}})] = a_1$, $\mathbb{E}[\boldsymbol{M}\mathbf{\Lambda}(\dot{\boldsymbol{H}})] = a_2$ is attained by $\mathbf{\Lambda}$ of the form $\mathbf{\Lambda}_{\lambda_1,\lambda_2}$. As the remaining terms in $\mathscr{S}_\star(\mathbf{\Lambda},s)$ depend on $\mathbf{\Lambda}$ only through $\mathbb{E}[\dot{\boldsymbol{H}}\mathbf{\Lambda}(\dot{\boldsymbol{H}})]$ and $\mathbb{E}[\boldsymbol{M}\mathbf{\Lambda}(\dot{\boldsymbol{H}})]$, we may restrict attention to $\mathbf{\Lambda}$ of this form. Thus

$$\frac{1}{N}\log\mathbb{E}^{\boldsymbol{m},\boldsymbol{n}}_{\mathsf{Pl}} Z_N(\boldsymbol{G}) \le \sup_{\lambda_1,\lambda_2}\mathscr{S}_\star(\lambda_1,\lambda_2) + o_N(1).$$

This implies $\mathbb{E}^{\boldsymbol{m},\boldsymbol{n}}_{\mathsf{Pl}} Z_N(\boldsymbol{G}) = e^{o(N)}$ under Condition 2.1.3.

We next argue that $\mathscr{S}_\star(\mathbf{\Lambda},s)$ upper bounds the exponential order of $\mathbb{E}^{\boldsymbol{m},\boldsymbol{n}}_{\mathsf{Pl}} Z_N(\boldsymbol{G},\mathbf{\Lambda})$, as claimed above. Due to (2.10), it suffices to bound the probability that some $\boldsymbol{x}\in\Sigma_N(\mathbf{\Lambda})$ satisfies all constraints. The planted model has the following law. Let $\dot{\boldsymbol{h}}\in\mathbb{R}^N$, $\widehat{\boldsymbol{h}}\in\mathbb{R}^M$ have coordinate distributions approximating $\mathcal{N}(0,\psi_0)$, $\mathcal{N}(0,q_0)$, and let $\boldsymbol{m} = \mathrm{th}(\dot{\boldsymbol{h}})$, $\boldsymbol{n} = F_{1-q_0}(\widehat{\boldsymbol{h}})$. A gaussian conditioning calculation (see Corollary 2.4.18) shows that conditional on $\nabla\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n}) = \boldsymbol{0}$,

$$\frac{\boldsymbol{G}}{\sqrt{N}} \overset{d}{=} \frac{\widehat{\boldsymbol{h}}\boldsymbol{m}^\top}{Nq_0} + \frac{\boldsymbol{n}\dot{\boldsymbol{h}}^\top}{N\psi_0} + \frac{P^\perp_{\boldsymbol{n}}\widetilde{\boldsymbol{G}}P^\perp_{\boldsymbol{m}}}{\sqrt{N}} + o_N(1).$$

Here $\widetilde{\boldsymbol{G}}$ is an i.i.d. copy of $\boldsymbol{G}$, $P^\perp_{\boldsymbol{m}}$ denotes the projection operator to the orthogonal complement of $\boldsymbol{m}$, and $o_N(1)$ is a matrix of operator norm $o_N(1)$. For any $\boldsymbol{x}\in\Sigma_N(\mathbf{\Lambda})$, we have $\frac{1}{N}\langle\boldsymbol{m},\boldsymbol{x}\rangle \simeq \mathbb{E}[\boldsymbol{M}\mathbf{\Lambda}(\dot{\boldsymbol{H}})]$ and $\frac{1}{N}\langle\dot{\boldsymbol{h}},\boldsymbol{x}\rangle \simeq \mathbb{E}[\dot{\boldsymbol{H}}\mathbf{\Lambda}(\dot{\boldsymbol{H}})]$. So,

$$\frac{\boldsymbol{G}\boldsymbol{x}}{\sqrt{N}} \overset{d}{=} \frac{\mathbb{E}[\boldsymbol{M}\mathbf{\Lambda}(\dot{\boldsymbol{H}})]}{q_0}\widehat{\boldsymbol{h}} + \frac{\mathbb{E}[\dot{\boldsymbol{H}}\mathbf{\Lambda}(\dot{\boldsymbol{H}})]}{\psi_0}\boldsymbol{n} + \sqrt{1-\frac{\mathbb{E}[\boldsymbol{M}\mathbf{\Lambda}(\dot{\boldsymbol{H}})]^2}{q_0}}\,\widetilde{\boldsymbol{g}} + o(\sqrt{N}),$$

where $\widetilde{\boldsymbol{g}}\sim\mathcal{N}(0,P^\perp_{\boldsymbol{n}})$ and $o(\sqrt{N})$ denotes a vector of norm $o(\sqrt{N})$. Thus

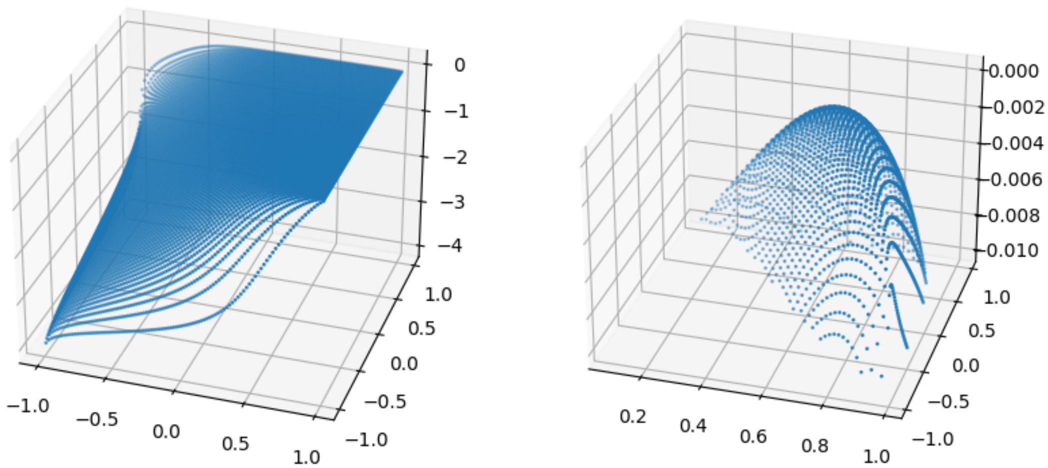
$$\frac{1}{N}\log\mathbb{P}^{\boldsymbol{m},\boldsymbol{n}}_{\mathsf{Pl}}\left(\frac{\boldsymbol{G}\boldsymbol{x}}{\sqrt{N}}\ge\kappa\boldsymbol{1}\right) \simeq \frac{1}{N}\log\mathbb{P}\left\{\widetilde{\boldsymbol{g}} \ge \frac{\kappa\boldsymbol{1} - \frac{\mathbb{E}[\boldsymbol{M}\mathbf{\Lambda}(\dot{\boldsymbol{H}})]}{q_0}\widehat{\boldsymbol{h}} - \frac{\mathbb{E}[\dot{\boldsymbol{H}}\mathbf{\Lambda}(\dot{\boldsymbol{H}})]}{\psi_0}\boldsymbol{n}}{\sqrt{1-\frac{\mathbb{E}[\boldsymbol{M}\mathbf{\Lambda}(\dot{\boldsymbol{H}})]^2}{q_0}}}\right\}. \tag{2.11}$$

This can be bounded by a change of measure calculation also used in [DS18]. Let $\widehat{\boldsymbol{g}}\sim\mathcal{N}(s\boldsymbol{n},\boldsymbol{I}_N)$ for any $s\ge 0$. Note that conditional on $\langle\widehat{\boldsymbol{g}},\boldsymbol{n}\rangle = 0$, we have $\widehat{\boldsymbol{g}} =_d \widetilde{\boldsymbol{g}}$. So, if S denotes the event in (2.11), then

$$\mathbb{P}(\widetilde{\boldsymbol{g}}\in S) \le \frac{\mathbb{P}(\widehat{\boldsymbol{g}}\in S)}{\mathbb{P}(\langle\widehat{\boldsymbol{g}},\boldsymbol{n}\rangle\approx 0)} \approx \exp\left(\frac{1}{2}s^2\psi_0 N\right)\mathbb{P}(\widehat{\boldsymbol{g}}\in S).$$

Since $\widehat{\boldsymbol{h}}$ has coordinate distribution $\hat{\boldsymbol{H}}$, this implies (see Lemma 2.7.8 for formal statement) that (2.11) is bounded by

$$\frac{1}{2}s^2\psi_0 + \alpha_\star\,\mathbb{E}\log\Psi\left(\frac{\kappa - \frac{\mathbb{E}[\boldsymbol{M}\mathbf{\Lambda}(\dot{\boldsymbol{H}})]}{q_0}\hat{\boldsymbol{H}} - \frac{\mathbb{E}[\dot{\boldsymbol{H}}\mathbf{\Lambda}(\dot{\boldsymbol{H}})]}{\psi_0}\boldsymbol{N}}{\sqrt{1-\frac{\mathbb{E}[\boldsymbol{M}\mathbf{\Lambda}(\dot{\boldsymbol{H}})]^2}{q_0}}} + s\boldsymbol{N}\right).$$



(a) $x, y \in [-1,1]^2$

(b) $\overline{\mathscr{S}}_\star(\operatorname{th}^{-1}(x), \operatorname{th}^{-1}(y)) \ge -0.01$

Figure 2.1: Plots of $(x,y) \mapsto \overline{\mathscr{S}}_\star(\operatorname{th}^{-1}(x), \operatorname{th}^{-1}(y))$ for $\kappa = 0$. Figure 2.1a plots over $x, y \in [-1,1]^2$, while Figure 2.1b restricts to inputs with $\overline{\mathscr{S}}_\star(\operatorname{th}^{-1}(x), \operatorname{th}^{-1}(y)) \ge -0.01$. The plots lie below 0, and from Figure 2.1b it appears the unique maximizer is $(x,y) = (\operatorname{th}(1), 0)$, corresponding to $(\lambda_1, \lambda_2) = (1, 0)$.

Combining with (2.10) shows that $\frac{1}{N} \log \mathbb{E}_{\mathsf{Pl}}^{\boldsymbol{m},\boldsymbol{n}} Z_N(\boldsymbol{G}, \boldsymbol{\Lambda}) \le \mathscr{S}_\star(\boldsymbol{\Lambda}, s) + o_N(1)$.

We conclude this subsection with a discussion of Condition 2.1.3. We expect $\boldsymbol{m}$ to approximate the barycenter of $S_N(\boldsymbol{G})$, and therefore that $\mathscr{S}_\star(\lambda_1, \lambda_2)$ is maximized by $(\lambda_1, \lambda_2) = (1, 0)$, corresponding to $\boldsymbol{\Lambda}_{\lambda_1,\lambda_2}(\dot{\boldsymbol{H}}) = \operatorname{th}(\dot{\boldsymbol{H}}) = \boldsymbol{M}$. Let

$$\overline{\mathscr{S}}_\star(\lambda_1, \lambda_2) = \mathscr{S}_\star(\boldsymbol{\Lambda}_{\lambda_1,\lambda_2}, \sqrt{1-q_0}),$$

which is an upper bound for $\mathscr{S}_\star$.

Lemma 2.2.5 (Proved in Section 2.7)**.** *The following holds.*

(a) The function $\mathscr{S}_\star(\lambda_1, \lambda_2)$ attains its supremum on $\mathbb{R}^2$.

(b) $\mathscr{S}_\star(1,0) = \overline{\mathscr{S}}_\star(1,0) = 0$.

(c) $\nabla \mathscr{S}_\star(1,0) = \nabla \overline{\mathscr{S}}_\star(1,0) = 0$.

(d) $\nabla^2 \mathscr{S}_\star(1,0) \preceq \nabla^2 \overline{\mathscr{S}}_\star(1,0)$

Claim 2.2.6 (Proved in Appendix 2.B)**.** *For $\kappa = 0$, there exists $C > 0$ such that $\nabla^2 \overline{\mathscr{S}}_\star(1,0) \preceq -CI$.*

Lemma 2.2.5 is proved for all κ, while Claim 2.2.6 is verified numerically for $\kappa = 0$ using rigorous interval arithmetic. Together, they imply that for $\kappa = 0$, $(1,0)$ is a local maximum of $\mathscr{S}_\star$ and $\overline{\mathscr{S}}_\star$. In Figure 2.1, we provide a plot of $\overline{\mathscr{S}}_\star$ for the case $\kappa = 0$. This gives numerical evidence that $\overline{\mathscr{S}}_\star$, and therefore $\mathscr{S}_\star$, has global maximum $(1,0)$.

2.3 Formal statement of results

In this section we state our main result for general κ, Theorem 2.3.6. We also reduce Theorem 2.3.6 to two primary inputs: approximate contiguity with a planted model (Lemma 2.3.8) and the upper bound for the first moment in the planted model (Proposition 2.3.9), which are proved in Section 2.4–2.6 and Section 2.7.

2.3.1 Krauth–Mézard threshold

We first define the threshold $\alpha_\star$ conjectured by [KM89], following the presentation of [DS18]. Define the standard gaussian density and complementary CDF by

$$\varphi(x) = \frac{\exp(-x^2/2)}{(2\pi)^{1/2}}, \qquad \Psi(x) = \int_x^\infty \phi(u)\, \mathsf{d}u. \tag{2.12}$$

Fix once and for all $\kappa \in \mathbb{R}$. For $q \in [0,1)$, define[5]

$$\mathcal{E}(x) = \frac{\varphi(x)}{\Psi(x)}, \qquad F_{1-q}(x) = \frac{\mathcal{E}}{(1-q)^{1/2}}\left(\frac{\kappa - x}{(1-q)^{1/2}}\right). \tag{2.13}$$

For $\psi \ge 0$ and $Z \sim \mathcal{N}(0,1)$, further define

$$P(\psi) = \mathbb{E}[\mathrm{th}(\psi^{1/2}Z)^2], \qquad R_\alpha(q) = \alpha\,\mathbb{E}[F_{1-q}(q^{1/2}Z)^2],$$

and define the Gardner free energy (or Gardner volume formula) by

$$\mathscr{G}(\alpha,q,\psi) = -\frac{(1-q)\psi}{2} + \mathbb{E}\log(2\mathrm{ch}(\psi^{1/2}Z)) - \alpha\,\mathbb{E}\log\Psi\left(\frac{\kappa - q^{1/2}Z}{(1-q)^{1/2}}\right). \tag{2.14}$$

The physical meanings of these formulas are best understood in terms of a heuristic derivation of the TAP free energy $\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n})$ and TAP equations, which we explain next. (These quantities will be formally defined in (2.15), (2.16).) If we regard $\boldsymbol{G}$ as a complete bipartite factor graph on N variables and M constraints, we can study the perceptron model by the standard **belief propagation** (BP) equations [MM09, Chapter 14]. In the mean-field (dense graph) limit, these equations simplify considerably. First, because the influence of any particular message is small, all the messages emanating from a particular variable $i \in [M]$ (resp. constraint $a \in [M]$) can be consolidated into a single message m_i (resp. n_a). The TAP variables $(\boldsymbol{m},\boldsymbol{n})$ thus represent these consolidated messages. The BP equations then become the TAP equations, and the **Bethe free energy** of this BP system becomes the TAP free energy. See [Méz17] for an example of this derivation in a related model.

Moreover, by central limit theorem considerations, we expect that the coordinates of $\dot{\boldsymbol{h}} = \mathrm{th}^{-1}(\boldsymbol{m})$ and $\widehat{\boldsymbol{h}} = F^{-1}_{1-\|\boldsymbol{m}\|^2/N}(\boldsymbol{n})$ have gaussian empirical measure. Let these gaussians have variance ψ and q, respectively; this is the physical meaning of these parameters. Then the BP consistency relations require that ψ, q satisfy the fixed-point equation $q = P(\psi)$, $\psi = R_\alpha(q)$, and the corresponding Bethe free energy is precisely $\mathscr{G}(\alpha,q,\psi)$. Finally, we expect $\alpha_\star$ to be the constraint density where this Bethe free energy crosses zero. Under the following condition, which was verified in [DS18] for $\kappa = 0$, this heuristic picture can be formalized into a definition of $\alpha_\star$.

Condition 2.3.1. There exist $0 < \alpha_{\mathrm{lb}} < \alpha_{\mathrm{ub}}$ and $0 < q_{\mathrm{lb}} < q_{\mathrm{ub}} < 1$ (depending on κ) such that the following holds. For any $\alpha \in (\alpha_{\mathrm{lb}}, \alpha_{\mathrm{ub}})$,

$$\sup_{q \in (q_{\mathrm{lb}}, q_{\mathrm{ub}})} (P \circ R_\alpha)'(q) < 1,$$

and there is a unique $q_\star = q_\star(\alpha,\kappa) \in (q_{\mathrm{lb}}, q_{\mathrm{ub}})$ such that $q_\star = P(R_\alpha(q_\star))$. Let $\psi_\star = \psi_\star(\alpha,\kappa) = R_\alpha(q_\star)$. For $\alpha \in (\alpha_{\mathrm{lb}}, \alpha_{\mathrm{ub}})$, the function $\mathscr{G}_\star(\alpha) = \mathscr{G}(\alpha, q_\star(\alpha,\kappa), \psi_\star(\alpha,\kappa))$ is strictly decreasing, with a unique root $\alpha_\star = \alpha_\star(\kappa)$.

Proposition 2.3.2 ([DS18, Proposition 1.3])**.** *For $\kappa = 0$, Condition 2.3.1 holds for $\alpha_{\mathrm{lb}} = 0.833078599$, $\alpha_{\mathrm{ub}} = 0.833078600$, $q_{\mathrm{lb}} = 0.56394907949$, $q_{\mathrm{ub}} = 0.56394908030$.*

[5]The function F_{1-q} is denoted F_q in [DS18]. We change this notation to be consistent with the meaning of $F_{\varepsilon,\varrho}$ (2.18) appearing in our proofs.

2.3.2 Main result

Throughout, let $\alpha_\star = \alpha_\star(\kappa)$ and $(q_0, \psi_0) = (q_\star(\alpha_\star, \kappa), \psi_\star(\alpha_\star, \kappa))$ be given by Condition 2.3.1. We now introduce two more numerical conditions needed for our main result, which will be verified for $\kappa = 0$ in Appendix 2.B using rigorous interval arithmetic. In the below formulas, let $Z \sim \mathcal{N}(0, 1)$.

Condition 2.3.3. We have $\alpha_\star \mathbb{E}[\mathrm{th}'(\psi_0^{1/2} Z)^2] \mathbb{E}[F'_{1-q_0}(q_0^{1/2} Z)^2] < 1$.

Condition 2.3.4. Define the functions $m : (-1, +\infty) \to (0, +\infty)$ and $\widehat{f}_0 : \mathbb{R} \to (0, +\infty)$ by

$$m(z) = \mathbb{E}[(z + \mathrm{ch}^2(\psi_0^{1/2} Z))^{-1}],$$
$$\widehat{f}_0(x) = -\frac{F'_{1-q_0}(x)}{1 + (1 - q_0) F'_{1-q_0}(x)} = \frac{\mathcal{E}'((\kappa - x)/(1 - q_0)^{1/2})}{(1 - q_0)(1 - \mathcal{E}'((\kappa - x)/(1 - q_0)^{1/2}))}.$$

(By Lemma 2.4.21(b) below, $\mathcal{E}'$ has image in $(0, 1)$, and thus $\widehat{f}_0(x) > 0$.) Then, for $d_0 = \alpha_\star \mathbb{E}[F'_{1-q_0}(q_0^{1/2} Z)]$ and $\lambda : (-1, +\infty) \to \mathbb{R}$ defined by

$$\lambda(z) = z - \alpha_\star \mathbb{E}\left[\frac{\widehat{f}_0(q_0^{1/2} Z)}{1 + m(z)\widehat{f}_0(q_0^{1/2} Z)}\right] - d_0,$$

we have $\lambda_0 \equiv \inf_{z > -1} \lambda(z) < 0$.

The following lemma shows that minimizer of λ exists and is the unique root of a decreasing function, and it suffices to check Condition 2.3.4 at the value $\lambda(z_0)$.

Lemma 2.3.5 (Proved in Section 2.6)**.** *The function λ is differerentiable with $\lambda'(z) = 1 - \alpha_\star \theta(z)$, where $\theta : (-1, +\infty) \to (0, +\infty)$ is defined by*

$$\theta(z) = \mathbb{E}[(z + \mathrm{ch}^2(\psi_0^{1/2} Z))^{-2}] \mathbb{E}\left[\left(\frac{\widehat{f}_0(q_0^{1/2} Z)}{1 + m(z)\widehat{f}_0(q_0^{1/2} Z)}\right)^2\right].$$

Moreover θ is continuous and strictly decreasing, with

$$\lim_{z \downarrow -1} \theta(z) = +\infty, \qquad \lim_{z \uparrow +\infty} \theta(z) = 0.$$

In particular θ has a well-defined inverse $\theta^{-1} : (0, +\infty) \to (-1, +\infty)$, and λ is strictly convex on $(-1, +\infty)$ with minimizer $z_0 = \theta^{-1}(\alpha_\star^{-1})$. Thus λ_0 defined in Condition 2.3.4 satisfies $\lambda_0 = \lambda(z_0)$.

Theorem 2.3.6 (Main result, general κ)**.** *For any $\kappa \in \mathbb{R}$, under Conditions 2.1.3, 2.3.1, 2.3.3, and 2.3.4 the following holds. For any $\alpha > \alpha_\star(\kappa)$, we have $\lim_{N\to\infty} \mathbb{P}(M_N(\kappa)/N \geq \alpha) = 0$.*

Remark 2.3.7. The conditions in Theorem 2.3.6 serve the following purposes.

- Condition 2.1.3 controls the first moment of the partition function in the planted model.
- Condition 2.3.1 makes the threshold $\alpha_\star(\kappa)$ well-defined.
- Condition 2.3.3 ensures that the AMP iterates converge in the sense of (2.2).
- Condition 2.3.4 ensures that $\mathcal{G}^{\varepsilon}_{\mathsf{TAP}}$ (see Subsection 2.2.4) is locally concave near late AMP iterates.

With the exception of Appendix 2.B, we will assume all conditions in Theorem 2.3.6 without further notice.

2.3.3 Proof of Theorem 2.3.6

We will carry out nearly the entire proof at constraint density $\alpha_\star$. Thus, we set $M = \lfloor \alpha_\star N \rfloor$ and define $\boldsymbol{G} \in \mathbb{R}^{M\times N}$ and $Z_N(\boldsymbol{G})$ as above.

The main step of the proof is a reduction to a planted model, formalized by Lemma 2.3.8 below. Let $\mathbb{P}$ denote the law of $\boldsymbol{G}$ with i.i.d. $\mathcal{N}(0,1)$ entries, and let $\mathbb{P}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{Pl}}$ be the planted law defined in Definition 2.4.3. This is the law of $\boldsymbol{G}$ conditional on $\nabla \mathcal{F}^{\varepsilon}_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n}) = \boldsymbol{0}$, for a perturbation $\mathcal{F}^{\varepsilon}_{\mathsf{TAP}}$ of $\mathcal{F}_{\mathsf{TAP}}$ defined in (2.24). (These will actually be probability measures over $(\boldsymbol{G}, \dot{\boldsymbol{g}}, \widehat{\boldsymbol{g}})$ for auxiliary disorder $\dot{\boldsymbol{g}}, \widehat{\boldsymbol{g}}$ defined below.) Let $\mathcal{S}_{\varepsilon,\upsilon}$ be a similar perturbation of $\mathcal{S}_\upsilon$ defined in (2.27).

Lemma 2.3.8 (Proved in Section 2.4–2.6)**.** *For any $(\boldsymbol{G}, \dot{\boldsymbol{g}}, \widehat{\boldsymbol{g}})$-measurable event $\mathscr{E}$ and any $\varepsilon, \upsilon > 0$, there exists $C = C(\varepsilon,\upsilon)$ such that*

$$\mathbb{P}(\mathscr{E}) \le C \sup_{(\boldsymbol{m},\boldsymbol{n}) \in \mathcal{S}_{\varepsilon,\upsilon}} \mathbb{P}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{Pl}}(\mathscr{E})^{1/2} + o_N(1).$$

The following proposition controls the first moment of $Z_N(\boldsymbol{G})$ in the planted model, formalizing the heuristic calculation in Subsection 2.2.6. Here $\mathbb{E}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{Pl}}$ denotes expectation with respect to $\mathbb{P}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{Pl}}$.

Proposition 2.3.9 (Proved in Section 2.7)**.** *For any $\delta > 0$, there exists $\varepsilon, \upsilon > 0$ such that*

$$\sup_{(\boldsymbol{m},\boldsymbol{n}) \in \mathcal{S}_{\varepsilon,\upsilon}} \mathbb{E}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{Pl}}[Z_N(\boldsymbol{G})] \le e^{\delta N}.$$

From these two results, Theorem 2.3.6 follows by a short argument.

Proposition 2.3.10. *For any $\delta > 0$,*

$$\mathbb{P}[Z_N(\boldsymbol{G}) \le e^{\delta N}] = 1 - o_N(1).$$

Proof. Let $\mathscr{E} = \{Z_N(\boldsymbol{G}) \le e^{\delta N}\}$. By Lemma 2.3.8 and Markov's inequality,

$$\mathbb{P}(\mathscr{E}^c) \le C \sup_{(\boldsymbol{m},\boldsymbol{n}) \in \mathcal{S}_{\varepsilon,\upsilon}} \mathbb{P}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{Pl}}(\mathscr{E}^c)^{1/2} + o_N(1) \le C e^{-\delta N/2} \sup_{(\boldsymbol{m},\boldsymbol{n}) \in \mathcal{S}_{\varepsilon,\upsilon}} \mathbb{E}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{Pl}}[Z_N(\boldsymbol{G})]^{1/2} + o_N(1).$$

By Proposition 2.3.9, we may choose ε, υ so this supremum is at most $e^{\delta N/4}$. □

Proof of Theorem 2.3.6. Let $M_{\rm all} = \lfloor \alpha N \rfloor$, and let $\boldsymbol{G}_{\rm all} = \binom{\boldsymbol{G}}{\widehat{\boldsymbol{G}}} \in \mathbb{R}^{M_{\rm all}\times N}$, where $\widehat{\boldsymbol{G}} \in \mathbb{R}^{(M_{\rm all}-M)\times N}$ has i.i.d. $\mathcal{N}(0,1)$ entries. Set $\delta < \frac{1}{2}(\alpha - \alpha_\star)\log\frac{1}{\Phi(\kappa)}$. Let $\mathscr{E} = \{Z_N(\boldsymbol{G}) \le e^{\delta N}\}$, which satisfies $\mathbb{P}(\mathscr{E}) = 1 - o_N(1)$ by Proposition 2.3.10. Then

$$\mathbb{P}(M_N(\kappa)/N \ge \alpha) = \mathbb{P}(Z_N(\boldsymbol{G}_{\rm all}) > 0) \le \mathbb{P}(\mathscr{E}^c) + \mathbb{E}[Z_N(\boldsymbol{G}_{\rm all})\boldsymbol{1}\{\mathscr{E}\}].$$

Since the rows of $\widehat{\boldsymbol{G}}$ are i.i.d. samples from $\mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_N)$ independent of $\boldsymbol{G}$, for any $\boldsymbol{x} \in \Sigma_N$,

$$\mathbb{E}[Z_N(\boldsymbol{G}_{\rm all})\boldsymbol{1}\{\mathscr{E}\}] \le e^{\delta N} \mathbb{P}_{\boldsymbol{g}\sim\mathcal{N}(\boldsymbol{0},\boldsymbol{I}_N)}\left(\frac{\langle \boldsymbol{g}, \boldsymbol{x}\rangle}{\sqrt{N}} \ge \kappa\right)^{M_{\rm all}-M} = e^{\delta N}\Phi(\kappa)^{M_{\rm all}-M} = o_N(1).$$ □

2.3.4 TAP and AMP formulas

In this subsection we provide the formulas for the TAP free energy, TAP equations, and AMP iteration mentioned above. The heuristic derivation of the former two were discussed below (2.14), and the latter is obtained by iterating the TAP equations in a suitable way.

The contents of this subsection play no formal role in the following proofs. We include these formulas for the reader's convenience, to allow a comparison with the ε-perturbed TAP free energy and AMP iteration defined in Subsection 2.4.2 below. (See also (2.36), (2.37) for the ε-perturbed TAP equations.) For $(\boldsymbol{m},\boldsymbol{n}) \in \mathbb{R}^N \times \mathbb{R}^M$, let $q(\boldsymbol{m}) = \|\boldsymbol{m}\|^2/N$ and $\psi(\boldsymbol{n}) = \|\boldsymbol{n}\|^2/N$. The TAP free energy for this model is

$$\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n}) = \sum_{i=1}^N \mathcal{H}\left(\frac{1+m_i}{2}\right) + \sum_{a=1}^M \log\Psi\left(\frac{\kappa - \frac{\langle \boldsymbol{g}^a, \boldsymbol{m}\rangle}{\sqrt{N}} + (1-q(\boldsymbol{m}))n_a}{(1-q(\boldsymbol{m}))^{1/2}}\right) + \frac{N}{2}(1-q(\boldsymbol{m}))\psi(\boldsymbol{n}). \quad (2.15)$$

(Recall $\mathcal{H}(x) = -x\log(x) - (1-x)\log(1-x)$ is the binary entropy function.) The TAP equations are the stationarity conditions of $\mathcal{F}_{\mathsf{TAP}}$, and are

$$\boldsymbol{n} = F_{1-q(\boldsymbol{m})}(\widehat{\boldsymbol{h}}) \equiv F_{1-q(\boldsymbol{m})}\left(\frac{\boldsymbol{G}\boldsymbol{m}}{\sqrt{N}} - b(\boldsymbol{m})\boldsymbol{n}\right), \qquad \boldsymbol{m} = \operatorname{th}(\dot{\boldsymbol{h}}) \equiv \operatorname{th}\left(\frac{\boldsymbol{G}^\top\boldsymbol{n}}{\sqrt{N}} - d(\boldsymbol{m},\boldsymbol{n})\boldsymbol{m}\right), \tag{2.16}$$

where

$$b(\boldsymbol{m}) = 1 - q(\boldsymbol{m}), \qquad d(\boldsymbol{m},\boldsymbol{n}) = \frac{1}{N}\sum_{a=1}^{M} F'_{1-q(\boldsymbol{m})}(n_a).$$

Recall that these are the mean-field limit of the BP equations for this model. The terms $b(\boldsymbol{m})\boldsymbol{n}$ and $d(\boldsymbol{m},\boldsymbol{n})\boldsymbol{m}$ compensate for backtracking and are known as the **Onsager correction** terms.

Let q_0, ψ_0 be as in Condition 2.3.1, and define

$$b_0 = \mathbb{E}[\operatorname{th}'(\psi_0^{1/2} Z)] = 1 - q_0, \qquad d_0 = \alpha_\star\,\mathbb{E}[F'_{1-q_0}(q_0^{1/2} Z)].$$

The AMP iteration associated to $\mathcal{F}_{\mathsf{TAP}}$ is given by $\boldsymbol{n}^{-1} = \boldsymbol{0} \in \mathbb{R}^M$, $\boldsymbol{m}^0 = q_0^{1/2}\boldsymbol{1} \in \mathbb{R}^N$, and

$$\boldsymbol{n}^k = F_{1-q_0}(\widehat{\boldsymbol{h}}^k) = F_{1-q_0}\left(\frac{\boldsymbol{G}\boldsymbol{m}^k}{\sqrt{N}} - b_0\boldsymbol{n}^{k-1}\right), \qquad \boldsymbol{m}^{k+1} = \operatorname{th}(\dot{\boldsymbol{h}}^{k+1}) = \operatorname{th}\left(\frac{\boldsymbol{G}^\top\boldsymbol{n}^k}{\sqrt{N}} - d_0\boldsymbol{m}^k\right). \tag{2.17}$$

2.4 Reduction to planted model

In this section we prove the central Lemma 2.3.8, using inputs from Section 2.5–2.6 as described below. Subsection 2.4.1–2.4.5 are devoted to this proof. Subsection 2.4.6 derives the law of the planted model $\mathbb{P}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{Pl}}$, which will be useful for calculations in the rest of the paper. To maintain a smooth presentation, we defer some proofs to Subsection 2.4.7, and routine but technical arguments to Appendix 2.A.

2.4.1 Parameter list and notations

For convenience, we record here the order in which several parameters used in the proof of Lemma 2.3.8 are set. Each item in this list can be set sufficiently small or large depending on any preceding item.

- ε, size of the perturbation to the AMP iteration and TAP free energy.
- C_{cvx} and C_{bd}, estimates for ρ_ε (defined below, see (2.22)) and its derivatives.
- η, bound on strong convexity of $\mathcal{F}^\varepsilon_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n})$ in $\boldsymbol{n}$, and C_{reg}, bound on regularity of $\nabla^2\mathcal{F}^\varepsilon_{\mathsf{TAP}}$.
- r_0, radius around late AMP iterates where there is a unique critical point of $\mathcal{F}^\varepsilon_{\mathsf{TAP}}$.
- υ_0, accuracy of AMP iterate under which there is a unique critical point of $\mathcal{F}^\varepsilon_{\mathsf{TAP}}$ nearby.
- k, index of AMP iterate $(\boldsymbol{m}^k, \boldsymbol{n}^k)$ with accuracy υ_0.
- υ, tolerance in $\mathcal{S}_{\varepsilon,\upsilon}$.
- υ_1, accuracy of AMP iterate under which, by convex-concavity considerations, the nearby unique critical point lies in $\mathcal{S}_{\varepsilon,\upsilon}$.
- ℓ, index of AMP iterate $(\boldsymbol{m}^\ell, \boldsymbol{n}^\ell)$ with accuracy υ_1.
- N, problem dimension.

This information will be reviewed when these parameters are introduced. Notations such as $o_k(1)$ will denote quantities that tend to zero as the subscripted parameter tends to zero or infinity, which may depend arbitrarily on preceding items in this list but do not depend on subsequent items. We will use the term "absolute constant" to mean a constant depending on none of these parameters (but possibly depending on $\kappa, \alpha_\star, q_0, \psi_0$, which are fixed at the outset). Note that the statement of Lemma 2.3.8 is monotone in υ, and thus υ can be set small depending on the parameters preceding it in this list.

We also define more notations appearing in the proofs. Throughout, Z, Z', Z'' denote i.i.d. standard gaussians. We use $\mathcal{P}_2(\mathbb{R}^k)$ to denote the space of probability measures on $\mathbb{R}^k$ with bounded second moment and $\mathbb{W}_2$ to denote 2-Wasserstein distance. p-lim denotes limit in probability.

We often consider functions $\mathcal{F} : \mathbb{R}^N \times \mathbb{R}^M \to \mathbb{R}$, with input $(\boldsymbol{m}, \boldsymbol{n}) \in \mathbb{R}^N \times \mathbb{R}^M$. We will write $\nabla_{\boldsymbol{m}} \mathcal{F} \in \mathbb{R}^N$, $\nabla_{\boldsymbol{n}} \mathcal{F} \in \mathbb{R}^M$ for the restriction of $\nabla \mathcal{F}$ to the coordinates corresponding to $\boldsymbol{m}$ and $\boldsymbol{n}$. The Hessian restrictions $\nabla^2_{\boldsymbol{m},\boldsymbol{m}} \mathcal{F} \in \mathbb{R}^{N\times N}$, $\nabla^2_{\boldsymbol{m},\boldsymbol{n}} \mathcal{F} \in \mathbb{R}^{N\times M}$, and $\nabla^2_{\boldsymbol{n},\boldsymbol{n}} \mathcal{F} \in \mathbb{R}^{M\times M}$ are defined similarly. $P_{\boldsymbol{m}} = \boldsymbol{m}\boldsymbol{m}^\top / \|\boldsymbol{m}\|^2 \in \mathbb{R}^{N\times N}$ denotes the projection operator onto the span of $\boldsymbol{m}$, and $P_{\boldsymbol{m}}^\perp = \boldsymbol{I}_N - P_{\boldsymbol{m}}$ denotes the projection operator onto its orthogonal complement.

2.4.2 Perturbed nonlinearities, AMP iteration, and TAP free energy

We next introduce perturbed versions of the AMP iteration (2.17) and TAP free energy (2.15). The purpose of the various perturbations is discussed in Remark 2.4.5 below. Let $\varepsilon > 0$ be small. For $\varrho \geq 0$, define

$$\overline{F}_{\varepsilon,\varrho}(x) = \log \mathbb{E}\, \chi_\varepsilon(x + \varrho^{1/2} Z), \qquad \chi_\varepsilon(x) = \exp\left(-\frac{1}{2}\varepsilon x^2\right) \mathbb{P}(x + \varepsilon^{1/2} Z' \geq \kappa).$$

Then, define the perturbed nonlinearities

$$\mathrm{th}_\varepsilon(x) = \mathrm{th}(x) + \varepsilon x, \qquad F_{\varepsilon,\varrho}(x) = \overline{F}'_{\varepsilon,\varrho}(x). \tag{2.18}$$

An elementary calculation shows that explicitly,

$$\begin{aligned}\overline{F}_{\varepsilon,\varrho}(x) &= -\frac{1}{2}\log(1+\varepsilon\varrho) - \frac{\varepsilon x^2}{2(1+\varepsilon\varrho)} + \log \Psi\left(\frac{\kappa(1+\varepsilon\varrho) - x}{\sqrt{(\varrho + \varepsilon(1+\varepsilon\varrho))(1+\varepsilon\varrho)}}\right)\\ F_{\varepsilon,\varrho}(x) &= -\frac{\varepsilon x}{1+\varepsilon\varrho} + \frac{1}{\sqrt{(\varrho + \varepsilon(1+\varepsilon\varrho))(1+\varepsilon\varrho)}} \mathcal{E}\left(\frac{\kappa(1+\varepsilon\varrho) - x}{\sqrt{(\varrho + \varepsilon(1+\varepsilon\varrho))(1+\varepsilon\varrho)}}\right).\end{aligned} \tag{2.19}$$

Let

$$\varrho_\varepsilon(q, \psi) = \frac{1 - q + \varepsilon - \varepsilon^2(\psi + \varepsilon)}{1 - 2\varepsilon(\psi + \varepsilon)}.$$

Define perturbed variants of the functions $P, R_{\alpha_\star}$ by

$$P^\varepsilon(\psi) = \mathbb{E}[\mathrm{th}_\varepsilon((\psi+\varepsilon)^{1/2} Z)^2], \qquad R^\varepsilon(q, \psi) = \alpha_\star\, \mathbb{E}[F_{\varepsilon,\varrho_\varepsilon(q,\psi)}((q+\varepsilon)^{1/2} Z)^2],$$

and let $\zeta_\varepsilon(\psi) = R^\varepsilon(P^\varepsilon(\psi), \psi)$.

Proposition 2.4.1 (Proved in Appendix 2.A)**.** *There exists $\iota > 0$ such that for all sufficiently small $\varepsilon > 0$,*

$$\sup_{\psi \in [\psi_0 - \iota, \psi_0 + \iota]} \zeta'_\varepsilon(\psi) < 1,$$

and there is a unique solution $\psi_\varepsilon \in [\psi_0 - \iota, \psi_0 + \iota]$ to $\psi_\varepsilon = \zeta_\varepsilon(\psi_\varepsilon)$. Let $q_\varepsilon = P^\varepsilon(\psi_\varepsilon)$ and $\varrho_\varepsilon = \varrho_\varepsilon(q_\varepsilon, \psi_\varepsilon)$. We further have $(q_\varepsilon, \psi_\varepsilon, \varrho_\varepsilon) \to (q_0, \psi_0, 1 - q_0)$ as $\varepsilon \downarrow 0$.

Lemma 2.4.2 (Proved in Subsection 2.4.7)**.** *We have $\varrho_\varepsilon = \mathbb{E}[\mathrm{th}'_\varepsilon((\psi_\varepsilon + \varepsilon)^{1/2} Z)]$.*

Let $d_\varepsilon = \alpha_\star \mathbb{E}[F'_{\varepsilon,\varrho_\varepsilon}((q_\varepsilon+\varepsilon)^{1/2}Z)]$. Further, let $\dot{\boldsymbol{g}} \sim \mathcal{N}(0,\boldsymbol{I}_N)$, $\widehat{\boldsymbol{g}} \sim \mathcal{N}(0,\boldsymbol{I}_M)$ be independent of $\boldsymbol{G}$. The perturbed AMP iteration is defined by $\boldsymbol{n}^{-1} = \boldsymbol{0} \in \mathbb{R}^M$, $\boldsymbol{m}^0 = q_\varepsilon^{1/2}\boldsymbol{1} \in \mathbb{R}^N$, and

$$\boldsymbol{n}^k = F_{\varepsilon,\varrho_\varepsilon}(\widehat{\boldsymbol{h}}^k) = F_{\varepsilon,\varrho_\varepsilon}\left(\frac{\boldsymbol{G}\boldsymbol{m}^k}{\sqrt{N}} + \varepsilon^{1/2}\widehat{\boldsymbol{g}} - \varrho_\varepsilon \boldsymbol{n}^{k-1}\right), \tag{2.20}$$

$$\boldsymbol{m}^{k+1} = \mathrm{th}_\varepsilon(\dot{\boldsymbol{h}}^{k+1}) = \mathrm{th}_\varepsilon\left(\frac{\boldsymbol{G}^\top\boldsymbol{n}^k}{\sqrt{N}} + \varepsilon^{1/2}\dot{\boldsymbol{g}} - d_\varepsilon \boldsymbol{m}^k\right). \tag{2.21}$$

Define the convex function $V_\varepsilon : \mathbb{R} \to \mathbb{R}$ and its dual

$$V_\varepsilon(\dot{h}) = \log(2\mathrm{ch}(\dot{h})) + \frac{1}{2}\varepsilon\dot{h}^2, \qquad V^*_\varepsilon(m) = \inf_{\dot{h}}\left\{-m\dot{h} + V_\varepsilon(\dot{h})\right\}.$$

Let $C_{\mathsf{cvx}}, C_{\mathsf{bd}} > 0$ be large in ε. Let $\rho_\varepsilon : \mathbb{R} \to \mathbb{R}$ be an (unspecified) thrice-differentiable function satisfying

$$\rho_\varepsilon(q_\varepsilon) = \varrho_\varepsilon, \qquad \rho'_\varepsilon(q_\varepsilon) = -1, \qquad \rho''_\varepsilon(q_\varepsilon) = C_{\mathsf{cvx}}, \tag{2.22}$$

such that the image of ρ_ε and its derivatives satisfies

$$\rho_\varepsilon \in [C_{\mathsf{bd}}^{-1}, C_{\mathsf{bd}}], \qquad |\rho_\varepsilon^{(p)}| \le C_{\mathsf{bd}} \text{ for } p \in \{1,2,3\}. \tag{2.23}$$

(For every C_{cvx}, there exists C_{bd} such that this is possible.) Recall that for $(\boldsymbol{m},\boldsymbol{n}) \in \mathbb{R}^N \times \mathbb{R}^M$, we defined $q(\boldsymbol{m}) = \|\boldsymbol{m}\|^2/N$ and $\psi(\boldsymbol{n}) = \|\boldsymbol{n}\|^2/N$. The perturbed TAP free energy is

$$\begin{aligned}\mathcal{F}^\varepsilon_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n}) = &\sum_{i=1}^N V^*_\varepsilon(m_i) + \varepsilon^{1/2}\langle \dot{\boldsymbol{g}}, \boldsymbol{m}\rangle + \sum_{a=1}^M \overline{F}_{\varepsilon,\rho_\varepsilon(q(\boldsymbol{m}))}\left(\frac{\langle \boldsymbol{g}^a, \boldsymbol{m}\rangle}{\sqrt{N}} + \varepsilon^{1/2}\widehat{g}_a - \rho_\varepsilon(q(\boldsymbol{m}))n_a\right) \\ &+ \frac{N}{2}\rho_\varepsilon(q(\boldsymbol{m}))\psi(\boldsymbol{n}).\end{aligned} \tag{2.24}$$

We are now ready to define the planted model.

Definition 2.4.3. For $(\boldsymbol{m},\boldsymbol{n}) \in \mathbb{R}^N \times \mathbb{R}^M$, let $\mathbb{P}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{Pl}}$ denote the law of $(\boldsymbol{G}, \dot{\boldsymbol{g}}, \widehat{\boldsymbol{g}})$ conditional on $\nabla\mathcal{F}^\varepsilon_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n}) = \boldsymbol{0}$, and $\mathbb{E}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{Pl}}$ denote the corresponding expectation. ($\mathbb{P}$ and $\mathbb{E}$ continue to refer to the law of $(\boldsymbol{G}, \dot{\boldsymbol{g}}, \widehat{\boldsymbol{g}})$ with i.i.d. standard gaussian entries.)

Remark 2.4.4. As shown in Lemma 2.4.16 below, for any fixed $(\boldsymbol{m},\boldsymbol{n})$, $\nabla\mathcal{F}^\varepsilon_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n}) = \boldsymbol{0}$ is equivalent to two linear equations (2.36), (2.37) in $(\boldsymbol{G}, \dot{\boldsymbol{g}}, \widehat{\boldsymbol{g}})$, and thus in the planted model $(\boldsymbol{G}, \dot{\boldsymbol{g}}, \widehat{\boldsymbol{g}})$ remains gaussian.

Remark 2.4.5. The above perturbations serve the following purposes.

- $V^*_\varepsilon(m_i)$ regularizes the term $\mathcal{H}(\frac{1+m_i}{2})$ in the original $\mathcal{F}_{\mathsf{TAP}}$, avoiding the singular behavior of $\mathcal{F}_{\mathsf{TAP}}$ near the boundary of $[-1,1]^N$.
- $\overline{F}_{\varepsilon,\varrho_\varepsilon}$ is chosen so that $\mathcal{F}^\varepsilon_{\mathsf{TAP}}$ is strongly convex in $\boldsymbol{n}$. As a consequence, if we define

  $$\mathcal{G}_{\mathsf{TAP}}(\boldsymbol{m}) = \inf_{\boldsymbol{n}} \mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n}), \qquad \mathcal{G}^\varepsilon_{\mathsf{TAP}}(\boldsymbol{m}) = \inf_{\boldsymbol{n}} \mathcal{F}^\varepsilon_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n}),$$

  then $\mathcal{G}^\varepsilon_{\mathsf{TAP}}(\boldsymbol{m})$ also regularizes $\mathcal{G}_{\mathsf{TAP}}(\boldsymbol{m})$, avoiding a singular behavior near the boundary of $\frac{1}{\sqrt{N}}\boldsymbol{G}\boldsymbol{m} \ge \kappa$. Indeed, $\mathcal{G}_{\mathsf{TAP}}(\boldsymbol{m}) = -\infty$ if this inequality fails in any coordinate.
- The nonlinearities th_ε and $F_{\varepsilon,\varrho_\varepsilon}$ have Lipschitz inverses, so that Euclidean distances in $(\boldsymbol{m},\boldsymbol{n})$ and $(\dot{\boldsymbol{h}}, \widehat{\boldsymbol{h}})$ are comparable.
- The perturbations $\varepsilon^{1/2}\widehat{\boldsymbol{g}}$ and $\varepsilon^{1/2}\dot{\boldsymbol{g}}$ are for technical convenience, as solutions to the original TAP equation (2.16) must lie on the codimension-one manifold

  $$\langle \dot{\boldsymbol{h}} + d(\boldsymbol{m},\boldsymbol{n})\boldsymbol{m}, \boldsymbol{m}\rangle = \frac{1}{\sqrt{N}}\langle \boldsymbol{n}, \boldsymbol{G}\boldsymbol{m}\rangle = \langle \boldsymbol{n}, \widehat{\boldsymbol{h}} + b(\boldsymbol{m})\boldsymbol{n}\rangle.$$

  With this perturbation, Kac–Rice arguments can take place on full space.
- We will see in Section 2.6 that the Hessian of $\mathcal{G}^\varepsilon_{\mathsf{TAP}}(\boldsymbol{m})$ is the sum of an anisotropic sample covariance matrix, a full-rank diagonal matrix, and a low-rank spike (recall (2.4)). The condition $\rho''_\varepsilon(q_\varepsilon) = C_{\mathsf{cvx}}$ ensures this spike cannot contribute to the top eigenvalue by adding a large negative spike to the Hessian. This simplifies the proof of strong concavity of $\mathcal{G}^\varepsilon_{\mathsf{TAP}}$ near late AMP iterates.

2.4.3 Inputs to reduction

We next state several inputs needed to prove Lemma 2.3.8. As anticipated in Subsection 2.2.2, the main input is Proposition 2.4.8, which formalizes criteria (R4) and (R5). First, we record that $\mathcal{F}^{\varepsilon}_{\mathsf{TAP}}$ is (deterministically) strongly convex in $\boldsymbol{n}$.

Proposition 2.4.6 (Proved in Subsection 2.4.7)**.** *There exists $\eta = \eta(\varepsilon, C_{\mathsf{cvx}}, C_{\mathsf{bd}}) > 0$ such that $\nabla^2_{\boldsymbol{n},\boldsymbol{n}} \mathcal{F}^{\varepsilon}_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n}) \succeq \eta \boldsymbol{I}_M$ for any $(\boldsymbol{m},\boldsymbol{n}) \in \mathbb{R}^N \times \mathbb{R}^M$.*

We next record a basic regularity estimate. Define

$$\nabla^2_{\diamond} \mathcal{F}^{\varepsilon}_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n}) = \nabla^2_{\boldsymbol{m},\boldsymbol{m}} \mathcal{F}^{\varepsilon}_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n}) - (\nabla^2_{\boldsymbol{m},\boldsymbol{n}} \mathcal{F}^{\varepsilon}_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n}))(\nabla^2_{\boldsymbol{n},\boldsymbol{n}} \mathcal{F}^{\varepsilon}_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n}))^{-1}(\nabla^2_{\boldsymbol{m},\boldsymbol{n}} \mathcal{F}^{\varepsilon}_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n}))^{\top}. \tag{2.25}$$

This arises as the Hessian of $\mathcal{G}^{\varepsilon}_{\mathsf{TAP}}$, as shown in Lemma 2.4.10 below.

Proposition 2.4.7 (Proved in Appendix 2.A)**.** *For any $D > 0$, there exists $C_{\mathsf{reg}} = C_{\mathsf{reg}}(\varepsilon, C_{\mathsf{cvx}}, C_{\mathsf{bd}}, D)$ such that over both $\mathbb{P}$ and $\mathbb{P}^{\boldsymbol{m}',\boldsymbol{n}'}_{\varepsilon,\mathsf{Pl}}$ for any $\|\boldsymbol{m}'\|^2, \|\boldsymbol{n}'\|^2 \le DN$, with high probability the following holds. For all $\|\boldsymbol{m}\|^2, \|\boldsymbol{n}\|^2 \le DN$, we have $\|\nabla^2 \mathcal{F}^{\varepsilon}_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n})\|_{\mathsf{op}} \le C_{\mathsf{reg}}$.*

For $\dot{\boldsymbol{h}} \in \mathbb{R}^N$, $\widehat{\boldsymbol{h}} \in \mathbb{R}^M$, define the coordinate empirical measures

$$\mu_{\dot{\boldsymbol{h}}} = \frac{1}{N}\sum_{i=1}^{N} \delta(\dot{h}_i), \qquad \mu_{\widehat{\boldsymbol{h}}} = \frac{1}{M}\sum_{a=1}^{M} \delta(\widehat{h}_i). \tag{2.26}$$

In words, these are probability measures on $\mathbb{R}$ with mass $1/N$ on each $\dot{h}_i$ (resp. $1/M$, $\widehat{h}_i$). For $\upsilon > 0$, let

$$\begin{aligned}
\mathcal{T}_{\varepsilon,\upsilon} &= \left\{ (\dot{\boldsymbol{h}}, \widehat{\boldsymbol{h}}) \in \mathbb{R}^N \times \mathbb{R}^M : \mathbb{W}_2(\mu_{\dot{\boldsymbol{h}}}, \mathcal{N}(0, \psi_\varepsilon + \varepsilon)), \mathbb{W}_2(\mu_{\widehat{\boldsymbol{h}}}, \mathcal{N}(0, q_\varepsilon + \varepsilon)) \le \upsilon \right\}, \\
\mathcal{S}_{\varepsilon,\upsilon} &= \left\{ (\mathrm{th}_\varepsilon(\dot{\boldsymbol{h}}), F_{\varepsilon,\varrho_\varepsilon}(\widehat{\boldsymbol{h}})) : (\dot{\boldsymbol{h}}, \widehat{\boldsymbol{h}}) \in \mathcal{T}_{\varepsilon,\upsilon} \right\}.
\end{aligned} \tag{2.27}$$

Let $(\boldsymbol{m}^k, \boldsymbol{n}^k)$ be as in (2.20), (2.21).

Proposition 2.4.8 (Proved in Section 2.5 and Section 2.6)**.** *There exist $r_0 > 0$, $k_0 : \mathbb{R}_+ \to \mathbb{N}$, $\upsilon : \mathbb{R}_+ \times \mathbb{N} \to \mathbb{R}_+$, depending on $\varepsilon, C_{\mathsf{cvx}}, C_{\mathsf{bd}}, \eta, C_{\mathsf{reg}}$, and an absolute constant $C_{\mathsf{spec}} > 0$ such that the following holds. For any $\upsilon_0 > 0$ and $k \ge k_0(\upsilon_0)$, with high probability under $\mathbb{P}$:*

(a) $(\boldsymbol{m}^k, \boldsymbol{n}^k) \in \mathcal{S}_{\varepsilon,\upsilon_0}$,

(b) $\|\nabla \mathcal{F}^{\varepsilon}_{\mathsf{TAP}}(\boldsymbol{m}^k, \boldsymbol{n}^k)\| \le \upsilon_0 \sqrt{N}$,

(c) $\nabla^2_{\diamond} \mathcal{F}^{\varepsilon}_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n}) \preceq -C_{\mathsf{spec}} \boldsymbol{I}_N$ *for all* $(\boldsymbol{m},\boldsymbol{n})$ *such that* $\|(\boldsymbol{m},\boldsymbol{n}) - (\boldsymbol{m}^k, \boldsymbol{n}^k)\| \le r_0 \sqrt{N}$.

Moreover, let $\upsilon = \upsilon(\upsilon_0, k)$. For any $(\boldsymbol{m}', \boldsymbol{n}') \in \mathcal{S}_{\varepsilon,\upsilon}$, with high probability under $\mathbb{P}^{\boldsymbol{m}',\boldsymbol{n}'}_{\varepsilon,\mathsf{Pl}}$, the above three conclusions hold and:

(d) $\|(\boldsymbol{m}^k, \boldsymbol{n}^k) - (\boldsymbol{m}', \boldsymbol{n}')\| \le \upsilon_0 \sqrt{N}$.

The following concentration estimate follows by adapting an argument of [GZ00] and provides input (R3).

Lemma 2.4.9 (Proved in Section 2.6)**.** *There exists C depending on $\varepsilon, C_{\mathsf{cvx}}$ such that for sufficiently small υ, uniformly over $(\boldsymbol{m},\boldsymbol{n}) \in \mathcal{S}_{\varepsilon,\upsilon}$,*

$$\mathbb{E}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{Pl}} \left[|\det \nabla^2 \mathcal{F}^{\varepsilon}_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n})|^2 \right]^{1/2} \le C \mathbb{E}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{Pl}} \left[|\det \nabla^2 \mathcal{F}^{\varepsilon}_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n})| \right].$$

2.4.4 Unique nearby critical point and conditioning lemma

Lemma 2.4.11 below provides a criterion under which a function has a unique critical point near a given approximate critical point. Lemma 2.4.12 is a lemma about conditioning a random function on a random vector with a unique critical point nearby, which is an adaptation of the Kac–Rice formula. This important technical tool also appears as [HMP24, Lemma 3.6], where it is used in conjunction with known results on topological trivialization to condition on the TAP fixed point selected by AMP. Here, we use it with properties of the planted model provided by Proposition 2.4.8 to prove topological trivialization itself.

Lemma 2.4.10. *Let $U_1 \subseteq \mathbb{R}^N$, $U_2 \subseteq \mathbb{R}^M$ be open and convex. Suppose $\mathcal{F} : U_1 \times U_2 \to \mathbb{R}$ is twice differentiable and satisfies $\nabla^2_{\boldsymbol{n},\boldsymbol{n}}\mathcal{F}(\boldsymbol{m},\boldsymbol{n}) \succeq \eta \boldsymbol{I}_M$ for all $(\boldsymbol{m},\boldsymbol{n}) \in U_1 \times U_2$ for some $\eta > 0$, and $\mathcal{G}(\boldsymbol{m}) \equiv \min_{\boldsymbol{n}\in U_2} \mathcal{F}(\boldsymbol{m},\boldsymbol{n})$ exists for all $\boldsymbol{m} \in U_1$. Then $\boldsymbol{n}(\boldsymbol{m}) = \arg\min_{\boldsymbol{n}\in U_2} \mathcal{F}(\boldsymbol{m},\boldsymbol{n})$ is unique and differentiable, with*

$$\nabla \boldsymbol{n}(\boldsymbol{m}) = (\nabla^2_{\boldsymbol{n},\boldsymbol{n}}\mathcal{F}(\boldsymbol{m},\boldsymbol{n}(\boldsymbol{m})))^{-1}(\nabla^2_{\boldsymbol{m},\boldsymbol{n}}\mathcal{F}(\boldsymbol{m},\boldsymbol{n}(\boldsymbol{m})))^\top. \tag{2.28}$$

Moreover $\mathcal{G}$ is twice differentiable, with

$$\nabla \mathcal{G}(\boldsymbol{m}) = \nabla_{\boldsymbol{m}}\mathcal{F}(\boldsymbol{m},\boldsymbol{n}), \qquad \nabla^2 \mathcal{G}(\boldsymbol{m}) = \nabla^2_{\diamond}\mathcal{F}(\boldsymbol{m},\boldsymbol{n}). \tag{2.29}$$

Proof. Strong convexity of $\mathcal{F}$ in $\boldsymbol{n}$ implies that $\boldsymbol{n}(\boldsymbol{m})$ is unique, and can be defined as the solution to $\nabla_{\boldsymbol{m}}\mathcal{F}(\boldsymbol{m},\boldsymbol{n}) = \boldsymbol{0}$. Then (2.28) follows from the implicit function theorem, while (2.29) follows from (2.28) and the chain rule. □

Lemma 2.4.11. *Let $\mathcal{F} : \mathbb{R}^N \times \mathbb{R}^M \to \mathbb{R}$ be twice differentiable and $(\boldsymbol{m}_0,\boldsymbol{n}_0) \in \mathbb{R}^N \times \mathbb{R}^M$. Let $\eta, C_{\mathsf{reg}}, \upsilon_0 > 0$, $r_0 = 2\eta^{-1}(1 + C_{\mathsf{reg}}\eta^{-1})^2\upsilon_0$, and $U = \mathsf{B}((\boldsymbol{m}_0,\boldsymbol{n}_0), r_0\sqrt{N})$. Suppose that:*

(C1) $\|\nabla \mathcal{F}(\boldsymbol{m}_0,\boldsymbol{n}_0)\| \le \upsilon_0\sqrt{N}$,

(C2) $\|\nabla^2 \mathcal{F}(\boldsymbol{m},\boldsymbol{n})\|_{\mathsf{op}} \le C_{\mathsf{reg}}$ *for all* $(\boldsymbol{m},\boldsymbol{n}) \in U$,

(C3) $\nabla^2_{\boldsymbol{n},\boldsymbol{n}}\mathcal{F}(\boldsymbol{m},\boldsymbol{n}) \succeq \eta \boldsymbol{I}_M$ *for all* $(\boldsymbol{m},\boldsymbol{n}) \in \mathbb{R}^N \times \mathbb{R}^M$,

(C4) $\nabla^2_{\diamond}\mathcal{F}(\boldsymbol{m},\boldsymbol{n}) \preceq -\eta \boldsymbol{I}_N$ *for all* $(\boldsymbol{m},\boldsymbol{n}) \in U$.

Then, there is a unique $(\boldsymbol{m}_*,\boldsymbol{n}_*) \in U$ such that $\nabla \mathcal{F}(\boldsymbol{m}_*,\boldsymbol{n}_*) = \boldsymbol{0}$. Moreover, for sufficiently small (possibly in N) $\iota > 0$, the image of U under the map $\nabla \mathcal{F}$ contains $\mathsf{B}(\boldsymbol{0},\iota) \subseteq \mathbb{R}^N \times \mathbb{R}^M$ and is one-to-one on this set.

Proof. Let $U_1 = \mathsf{B}(\boldsymbol{m}_0, r_0\sqrt{N}) \subseteq \mathbb{R}^N$ and $U_2 = \mathbb{R}^M$. Item (C3) implies that the hypotheses of Lemma 2.4.10 hold for $\mathcal{F}^{\varepsilon}_{\mathsf{TAP}}$ with this (U_1, U_2). Thus, for $\boldsymbol{m} \in U_1$, $\boldsymbol{n}(\boldsymbol{m})$ and $\mathcal{G}(\boldsymbol{m})$ from Lemma 2.4.10 are well-defined, with derivatives given therein. If $(\boldsymbol{m}_*,\boldsymbol{n}_*)$ is a critical point of $\mathcal{F}$, then $\boldsymbol{m}_*$ must be a critical point of $\mathcal{G}$. Item (C4) and equation (2.29) imply that $\nabla^2\mathcal{G}(\boldsymbol{m}) \preceq -\eta \boldsymbol{I}_N$ for all $\boldsymbol{m} \in U_1$. Thus $\mathcal{G}$ has at most one critical point in U_1, and $\mathcal{F}$ has at most one critical point in $U_1 \times U_2 \supseteq U$.

We now show that such a point exists. By strong concavity of $\mathcal{F}(\boldsymbol{m}_0, \cdot)$ and (C1),

$$\|\boldsymbol{n}_0 - \boldsymbol{n}(\boldsymbol{m}_0)\| \le \eta^{-1}\|\nabla_{\boldsymbol{n}}\mathcal{F}(\boldsymbol{m}_0,\boldsymbol{m}_0)\| \le \eta^{-1}\upsilon_0\sqrt{N}.$$

Because $\|\nabla^2\mathcal{F}(\boldsymbol{m},\boldsymbol{n})\|_{\mathsf{op}} \le C_{\mathsf{reg}}$, the map $(\boldsymbol{m},\boldsymbol{n}) \mapsto \nabla\mathcal{F}(\boldsymbol{m},\boldsymbol{n})$ is C_{reg}-Lipschitz. Thus

$$\|\nabla\mathcal{G}(\boldsymbol{m}_0)\| = \|\nabla\mathcal{F}(\boldsymbol{m}_0,\boldsymbol{n}(\boldsymbol{m}_0))\| \le \|\nabla\mathcal{F}(\boldsymbol{m}_0,\boldsymbol{n}_0)\| + C_{\mathsf{reg}}\|\boldsymbol{n}_0 - \boldsymbol{n}(\boldsymbol{m}_0)\| \le (1 + C_{\mathsf{reg}}\eta^{-1})\upsilon_0\sqrt{N}.$$

By strong concavity of $\mathcal{G}$, there exists a critical point $\boldsymbol{m}_*$ of $\mathcal{G}$ with

$$\|\boldsymbol{m}_0 - \boldsymbol{m}_*\| \le \eta^{-1}\|\nabla\mathcal{G}(\boldsymbol{m}_0)\| \le \eta^{-1}(1 + C_{\mathsf{reg}}\eta^{-1})\upsilon_0\sqrt{N}.$$

Then, with $\boldsymbol{n}_* = \boldsymbol{n}(\boldsymbol{m}_*)$, $(\boldsymbol{m}_*,\boldsymbol{n}_*)$ is a critical point of $\mathcal{F}$. By conditions (C2), (C3) and equation (2.28), $\boldsymbol{n}(\cdot)$ is $C_{\mathsf{reg}}\eta^{-1}$-Lipschitz. So,

$$\|\boldsymbol{n}_0 - \boldsymbol{n}_*\| \le \|\boldsymbol{n}_0 - \boldsymbol{n}(\boldsymbol{m}_0)\| + C_{\mathsf{reg}}\eta^{-1}\|\boldsymbol{m}_0 - \boldsymbol{m}_*\| \le \eta^{-1}(1 + C_{\mathsf{reg}}\eta^{-1})^2\upsilon_0\sqrt{N}.$$

This shows that $(\boldsymbol{m}_*, \boldsymbol{n}_*) \in U$, proving the first claim, and furthermore $(\boldsymbol{m}_*, \boldsymbol{n}_*)$ lies in the interior of U. To show the second claim, we first prove that any $(\boldsymbol{m}, \boldsymbol{n}) \in U$ such that $\|\nabla \mathcal{F}(\boldsymbol{m}, \boldsymbol{n})\| \le \iota$ lies in a neighborhood of $(\boldsymbol{m}^*, \boldsymbol{n}^*)$. First,

$$\|\boldsymbol{n} - \boldsymbol{n}(\boldsymbol{m})\| \le \eta^{-1} \|\nabla_{\boldsymbol{n}} \mathcal{F}(\boldsymbol{m}, \boldsymbol{n})\| \le \eta^{-1}\iota.$$

Similarly to above, $\|\nabla \mathcal{G}(\boldsymbol{m})\| \le (1 + C_{\mathsf{reg}}\eta^{-1})\iota$, so we conclude

$$\|\boldsymbol{m} - \boldsymbol{m}_*\| \le \eta^{-1}(1 + C_{\mathsf{reg}}\eta^{-1})\iota, \qquad \|\boldsymbol{n} - \boldsymbol{n}_*\| \le \eta^{-1}(1 + C_{\mathsf{reg}}\eta^{-1})^2\iota.$$

Thus $(\boldsymbol{m}, \boldsymbol{n})$ lies in a neighborhood of $(\boldsymbol{m}_*, \boldsymbol{n}_*)$, which is contained in U because $(\boldsymbol{m}_*, \boldsymbol{n}_*)$ lies in the interior of U. However, by Schur's lemma,

$$\det \nabla^2 \mathcal{F}(\boldsymbol{m}_*, \boldsymbol{n}_*) = \det \nabla^2_{\boldsymbol{n},\boldsymbol{n}} \mathcal{F}(\boldsymbol{m}_*, \boldsymbol{n}_*) \det \nabla^2_{\diamond} \mathcal{F}(\boldsymbol{m}_*, \boldsymbol{n}_*) \neq 0.$$

By the inverse function theorem, $\nabla \mathcal{F}$ is invertible in a neighborhood of $(\boldsymbol{m}_*, \boldsymbol{n}_*)$, mapping it bijectively to a neighborhood of $\mathbf{0}$. This concludes the proof. $\square$

Lemma 2.4.12. *Let $\mathcal{F} : \mathbb{R}^N \times \mathbb{R}^M \to \mathbb{R}$ be a twice differentiable random function and $(\boldsymbol{m}_0, \boldsymbol{n}_0) \in \mathbb{R}^N \times \mathbb{R}^M$ be a random vector in the same probability space. Let $\eta, C_{\mathsf{reg}}, \upsilon_0, r_0$ be as in Lemma 2.4.11, and $U = \mathsf{B}((\boldsymbol{m}_0, \boldsymbol{n}_0), r_0\sqrt{N})$ (which is now a random set). Let $D > 0$ be arbitrary and $\mathscr{E}_0$ be the event that (C1) through (C4) hold and $\|\boldsymbol{m}_0\|^2, \|\boldsymbol{n}_0\|^2 \le DN$.*

Let $\varphi_{\nabla \mathcal{F}(\boldsymbol{m},\boldsymbol{n})}$ denote the probability density of $\nabla \mathcal{F}(\boldsymbol{m}, \boldsymbol{n})$ w.r.t. Lebesgue measure on $\mathbb{R}^N \times \mathbb{R}^M$. Suppose $\varphi_{\nabla \mathcal{F}(\boldsymbol{m},\boldsymbol{n})}(\boldsymbol{z})$ is bounded for $(\boldsymbol{m}, \boldsymbol{n}) \in \mathbb{R}^N \times \mathbb{R}^M$ and $\boldsymbol{z}$ in a neighborhood of $\mathbf{0}$, and continuous in $\boldsymbol{z}$ in this neighborhood uniformly over $(\boldsymbol{m}, \boldsymbol{n})$. Then, for any event $\mathscr{E} \subseteq \mathscr{E}_0$ in the same probability space,

$$\mathbb{P}(\mathscr{E}) = \int_{\mathbb{R}^N \times \mathbb{R}^M} \mathbb{E}\left[|\det \nabla^2 \mathcal{F}(\boldsymbol{m}, \boldsymbol{n})| \mathbf{1}\{\mathscr{E} \cap \{(\boldsymbol{m}, \boldsymbol{n}) \in U\}\} \middle| \nabla \mathcal{F}(\boldsymbol{m}, \boldsymbol{n}) = \mathbf{0}\right] \varphi_{\nabla \mathcal{F}(\boldsymbol{m},\boldsymbol{n})}(\mathbf{0}) \, \mathsf{d}(\boldsymbol{m}, \boldsymbol{n}).$$

Proof. On $\mathscr{E}_0$, Lemma 2.4.11 implies there is a unique critical point $(\boldsymbol{m}_*, \boldsymbol{n}_*)$ of $\mathcal{F}$ in U. Moreover the image of U under $\nabla \mathcal{F}$ contains $\mathsf{B}(\mathbf{0}, \iota)$ for small ι and is one-to-one on this set. By the area formula, on $\mathscr{E}_0$,

$$1 = \frac{1}{|\mathsf{B}(\mathbf{0}, \iota)|} \int_U |\det \nabla^2 \mathcal{F}(\boldsymbol{m}, \boldsymbol{n})| \mathbf{1}\{\|\nabla \mathcal{F}(\boldsymbol{m}, \boldsymbol{n})\| \le \iota\} \, \mathsf{d}(\boldsymbol{m}, \boldsymbol{n}).$$

Since $\mathscr{E} \subseteq \mathscr{E}_0$, multiplying both sides by $\mathbf{1}\{\mathscr{E}\}$ and taking expectations (via Fubini's theorem) yields

$$\begin{aligned}
\mathbb{P}(\mathscr{E}) &= \frac{1}{|\mathsf{B}(\mathbf{0}, \iota)|} \mathbb{E} \int_{\mathbb{R}^N \times \mathbb{R}^M} |\det \nabla^2 \mathcal{F}(\boldsymbol{m}, \boldsymbol{n})| \mathbf{1}\{\|\nabla \mathcal{F}(\boldsymbol{m}, \boldsymbol{n})\| \le \iota\} \mathbf{1}\{\boldsymbol{m} \in U\} \, \mathsf{d}(\boldsymbol{m}, \boldsymbol{n}) \\
&= \int_{\mathbb{R}^N \times \mathbb{R}^M} \mathbb{E}\left[|\det \nabla^2 \mathcal{F}(\boldsymbol{m}, \boldsymbol{n})| \mathbf{1}\{\mathscr{E} \cap \{\boldsymbol{m} \in U\}\} \middle| \|\nabla \mathcal{F}(\boldsymbol{m}, \boldsymbol{n})\| \le \iota\right] \frac{\mathbb{P}\{\|\nabla \mathcal{F}(\boldsymbol{m}, \boldsymbol{n})\| \le \iota\}}{|\mathsf{B}(\mathbf{0}, \iota)|} \, \mathsf{d}(\boldsymbol{m}, \boldsymbol{n}).
\end{aligned}$$

We now take the limit as $\iota \to 0$. On $\mathscr{E}_0$, $|\det \nabla^2 \mathcal{F}(\boldsymbol{m}, \boldsymbol{n})| \le C_{\mathsf{reg}}^{M+N}$. Since $\boldsymbol{m}_0, \boldsymbol{n}_0$ are bounded on $\mathscr{E}_0$, $\mathbf{1}\{\boldsymbol{m} \in U\} = 0$ almost surely for $\boldsymbol{m}$ outside a compact set. Since $\varphi_{\nabla \mathcal{F}(\boldsymbol{m},\boldsymbol{n})}(\boldsymbol{z})$ is bounded and continuous in $\boldsymbol{z}$, $\mathbb{P}\{\|\nabla \mathcal{F}(\boldsymbol{m}, \boldsymbol{n})\| \le \iota\}/|\mathsf{B}(\mathbf{0}, \iota)|$ is bounded, and limits to $\varphi_{\nabla \mathcal{F}(\boldsymbol{m},\boldsymbol{n})}(\boldsymbol{z})$ as $\iota \to 0$. Taking $\iota \to 0$ gives the result by dominated convergence. $\square$

2.4.5 Proof of planted reduction

We are now ready to prove Lemma 2.3.8. As anticipated in Subsection 2.2.2, Lemma 2.4.13 deduces (R2) from (R4), and Lemma 2.4.15 deduces (R1) from (R5). Then, Lemma 2.3.8 follows readily from the Kac–Rice formula.

Lemma 2.4.13. *For any $\upsilon > 0$, $\mathcal{S}_{\varepsilon,\upsilon}$ contains a critical point of $\mathcal{F}^{\varepsilon}_{\mathsf{TAP}}$ with high probability under $\mathbb{P}$.*

Proof. Let $\eta = \min(\eta(\varepsilon, C_{\mathsf{cvx}}, C_{\mathsf{bd}}), C_{\mathsf{spec}})$, where these terms are given by Propositions 2.4.6 and 2.4.8. Then, let $D = 2\max(q_\varepsilon, \psi_\varepsilon)$ and $C_{\mathsf{reg}} = C_{\mathsf{reg}}(\varepsilon, C_{\mathsf{cvx}}, C_{\mathsf{bd}}, D)$ be given by Proposition 2.4.7. Let r_0 be given by Proposition 2.4.8. Let υ_1 be small enough in υ that, with $r_1 = 2\eta^{-1}(1+C_{\mathsf{reg}}\eta^{-1})^2\upsilon_1$, we have $r_1 \le r_0$ and

$$\bigcup_{(\widetilde{\boldsymbol{m}},\widetilde{\boldsymbol{n}})\in\mathcal{S}_{\varepsilon,\upsilon_1}} \mathsf{B}((\widetilde{\boldsymbol{m}},\widetilde{\boldsymbol{n}}), r_1\sqrt{N}) \subseteq \mathcal{S}_{\varepsilon,\upsilon}. \tag{2.30}$$

(Since $\mathcal{S}_{\varepsilon,\upsilon}$ is the image of a product of two Wasserstein-balls under $(\mathrm{th}_\varepsilon, F_{\varepsilon,\varrho_\varepsilon})$, and $\mathrm{th}_\varepsilon^{-1}, F_{\varepsilon,\varrho_\varepsilon}^{-1}$ have Lipschitz constant depending only on ε, there exists υ_1 such that this holds.) Let $\ell = k_0(\upsilon_1)$ be given by Proposition 2.4.8. By Propositions 2.4.7 and 2.4.8, with high probability under $\mathbb{P}$,

- $\|\nabla^2\mathcal{F}_{\mathsf{TAP}}^\varepsilon(\boldsymbol{m},\boldsymbol{n})\|_{\mathsf{op}} \le C_{\mathsf{reg}}$ for all $\|\boldsymbol{m}\|^2, \|\boldsymbol{n}\|^2 \le DN$,
- $(\boldsymbol{m}^\ell, \boldsymbol{n}^\ell) \in \mathcal{S}_{\varepsilon,\upsilon_1}$,
- $\|\nabla\mathcal{F}_{\mathsf{TAP}}^\varepsilon(\boldsymbol{m}^\ell,\boldsymbol{n}^\ell)\| \le \upsilon_1\sqrt{N}$,
- $\nabla_\diamond^2\mathcal{F}_{\mathsf{TAP}}^\varepsilon(\boldsymbol{m},\boldsymbol{n}) \preceq -C_{\mathsf{spec}}\boldsymbol{I}_N$ for all $\|(\boldsymbol{m},\boldsymbol{n}) - (\boldsymbol{m}^\ell,\boldsymbol{n}^\ell)\| \le r_0\sqrt{N}$.

We now apply Lemma 2.4.11 with $(\mathcal{F}_{\mathsf{TAP}}^\varepsilon, \boldsymbol{m}^\ell, \boldsymbol{n}^\ell, \upsilon_1, r_1)$ in place of $(\mathcal{F}, \boldsymbol{m}_0, \boldsymbol{n}_0, \upsilon_0, r_0)$. The above points imply that conditions (C1), (C2), (C4) hold, and condition (C3) holds by Proposition 2.4.6. By Lemma 2.4.11, $\mathcal{F}_{\mathsf{TAP}}^\varepsilon$ has a critical point in $\mathsf{B}((\boldsymbol{m}^\ell,\boldsymbol{n}^\ell), r_1\sqrt{N})$. This lies in $\mathcal{S}_{\varepsilon,\upsilon}$ by (2.30). □

The following lemma shows that the condition in Lemma 2.4.12 regarding $\varphi_{\nabla\mathcal{F}}$ holds for $\mathcal{F} = \mathcal{F}_{\mathsf{TAP}}^\varepsilon$.

Lemma 2.4.14 (Proved in Subsection 2.4.7)**.** *The density $\varphi_{\nabla\mathcal{F}_{\mathsf{TAP}}^\varepsilon(\boldsymbol{m},\boldsymbol{n})}(\boldsymbol{z})$ under $\mathbb{P}$ is bounded for $(\boldsymbol{m},\boldsymbol{n}) \in \mathbb{R}^N \times \mathbb{R}^M$ and $\boldsymbol{z}$ in a neighborhood of $\boldsymbol{0}$, and continuous in $\boldsymbol{z}$ in this neighborhood uniformly over $(\boldsymbol{m},\boldsymbol{n})$.*

Lemma 2.4.15. *Let Crt_υ denote the set of critical points of $\mathcal{F}_{\mathsf{TAP}}^\varepsilon$ in $\mathcal{S}_{\varepsilon,\upsilon}$. For small $\upsilon > 0$, $\mathbb{E}\,|\mathsf{Crt}_\upsilon| \le 1 + o_N(1)$.*

Proof. By the Kac–Rice formula,

$$\mathbb{E}\,|\mathsf{Crt}_\upsilon| = \int_{\mathcal{S}_{\varepsilon,\upsilon}} \mathbb{E}_{\varepsilon,\mathsf{Pl}}^{\boldsymbol{m},\boldsymbol{n}}\left[|\det\nabla^2\mathcal{F}_{\mathsf{TAP}}^\varepsilon(\boldsymbol{m},\boldsymbol{n})|\right]\varphi_{\nabla\mathcal{F}_{\mathsf{TAP}}^\varepsilon(\boldsymbol{m},\boldsymbol{n})}(\boldsymbol{0})\;\mathsf{d}(\boldsymbol{m},\boldsymbol{n}). \tag{2.31}$$

As above, let $\eta = \min(\eta(\varepsilon, C_{\mathsf{cvx}}, C_{\mathsf{bd}}), C_{\mathsf{spec}})$, $D = 2\max(q_\varepsilon, \psi_\varepsilon)$, and $C_{\mathsf{reg}} = C_{\mathsf{reg}}(\varepsilon, C_{\mathsf{cvx}}, C_{\mathsf{bd}}, D)$. Let r_0 be given by Proposition 2.4.8, and

$$\upsilon_0 = \frac{\eta r_0}{2(1+C_{\mathsf{reg}}\eta^{-1})^2}.$$

Then set $k = k_0(\upsilon_0)$, where $k_0(\cdot)$ is as in Proposition 2.4.8. Let $\mathscr{E}$ be the event that:

- $\|\boldsymbol{m}^k\|^2, \|\boldsymbol{n}^k\|^2 \le DN$,
- $\|\nabla^2\mathcal{F}_{\mathsf{TAP}}^\varepsilon(\boldsymbol{m},\boldsymbol{n})\|_{\mathsf{op}} \le C_{\mathsf{reg}}$ for all $\|\boldsymbol{m}\|^2, \|\boldsymbol{n}\|^2 \le DN$,
- $\|\nabla\mathcal{F}_{\mathsf{TAP}}^\varepsilon(\boldsymbol{m}^k,\boldsymbol{n}^k)\| \le \upsilon_0\sqrt{N}$,
- $\nabla_\diamond^2\mathcal{F}_{\mathsf{TAP}}^\varepsilon(\boldsymbol{m},\boldsymbol{n}) \preceq -C_{\mathsf{spec}}\boldsymbol{I}_N$ for all $\|(\boldsymbol{m},\boldsymbol{n}) - (\boldsymbol{m}^k,\boldsymbol{n}^k)\| \le r_0\sqrt{N}$.

We claim that $\mathscr{E} \subseteq \mathscr{E}_0$, where $\mathscr{E}_0$ is the event defined in Lemma 2.4.12 with $(\mathcal{F}_{\mathsf{TAP}}^\varepsilon, \boldsymbol{m}^k, \boldsymbol{n}^k)$ for $(\mathcal{F}, \boldsymbol{m}_0, \boldsymbol{n}_0)$ (and thus $U = \mathsf{B}((\boldsymbol{m}^k,\boldsymbol{n}^k), r_0\sqrt{N})$). The above points imply conditions (C1), (C2), (C4), and condition (C3) follows from Proposition 2.4.6. By Lemma 2.4.14, Lemma 2.4.12 applies. Thus,

$$1 \ge \mathbb{P}(\mathscr{E}) = \int_{\mathbb{R}^N\times\mathbb{R}^M} \mathbb{E}_{\varepsilon,\mathsf{Pl}}^{\boldsymbol{m},\boldsymbol{n}}\left[|\det\nabla^2\mathcal{F}_{\mathsf{TAP}}^\varepsilon(\boldsymbol{m},\boldsymbol{n})|\mathbf{1}\{\mathscr{E}\cap\{(\boldsymbol{m},\boldsymbol{n})\in U\}\}\right]\varphi_{\nabla\mathcal{F}_{\mathsf{TAP}}^\varepsilon(\boldsymbol{m},\boldsymbol{n})}(\boldsymbol{0})\;\mathsf{d}(\boldsymbol{m},\boldsymbol{n}). \tag{2.32}$$

Let $\upsilon \le \min(\upsilon(\upsilon_0, k), \upsilon(r_0, k))$, for $\upsilon(\cdot,\cdot)$ as in Proposition 2.4.8. Define (compare with (2.31))

$$I_1 = \int_{\mathcal{S}_{\varepsilon,\upsilon}} \mathbb{E}_{\varepsilon,\mathsf{Pl}}^{\boldsymbol{m},\boldsymbol{n}}\left[|\det\nabla^2\mathcal{F}_{\mathsf{TAP}}^\varepsilon(\boldsymbol{m},\boldsymbol{n})|\mathbf{1}\{\mathscr{E}\cap\{(\boldsymbol{m},\boldsymbol{n})\in U\}\}\right]\varphi_{\nabla\mathcal{F}_{\mathsf{TAP}}^\varepsilon(\boldsymbol{m},\boldsymbol{n})}(\boldsymbol{0})\;\mathsf{d}(\boldsymbol{m},\boldsymbol{n})$$

and $I_2 = \mathbb{E}\,|\mathsf{Crt}_\upsilon| - I_1$. By Propositions 2.4.7 and 2.4.8, for any $(\boldsymbol{m},\boldsymbol{n}) \in \mathcal{S}_{\varepsilon,\upsilon}$, we have $\mathbb{P}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{Pl}}(\mathscr{E} \cap \{(\boldsymbol{m},\boldsymbol{n}) \in U\}) \ge 1-\iota$ for some $\iota = o_N(1)$. By Cauchy–Schwarz and Lemma 2.4.9,

$$\begin{aligned}
I_2 &= \int_{\mathcal{S}_{\varepsilon,\upsilon}} \mathbb{E}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{Pl}}\left[|\det \nabla^2 \mathcal{F}^\varepsilon_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n})| \mathbf{1}\{(\mathscr{E} \cap \{(\boldsymbol{m},\boldsymbol{n}) \in U\})^c\}\right] \varphi_{\nabla \mathcal{F}^\varepsilon_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n})}(\mathbf{0})\ \mathsf{d}(\boldsymbol{m},\boldsymbol{n}) \\
&\le \int_{\mathcal{S}_{\varepsilon,\upsilon}} \mathbb{E}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{Pl}}\left[|\det \nabla^2 \mathcal{F}^\varepsilon_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n})|^2\right]^{1/2} \mathbb{P}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{Pl}}\left[(\mathscr{E} \cap \{(\boldsymbol{m},\boldsymbol{n}) \in U\})^c\right]^{1/2} \varphi_{\nabla \mathcal{F}^\varepsilon_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n})}(\mathbf{0})\ \mathsf{d}(\boldsymbol{m},\boldsymbol{n}) \\
&\le C\iota^{1/2} \int_{\mathcal{S}_{\varepsilon,\upsilon}} \mathbb{E}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{Pl}}\left[|\det \nabla^2 \mathcal{F}^\varepsilon_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n})|\right] \varphi_{\nabla \mathcal{F}^\varepsilon_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n})}(\mathbf{0})\ \mathsf{d}(\boldsymbol{m},\boldsymbol{n}) \overset{(2.31)}{=} C\iota^{1/2}\,\mathbb{E}\,|\mathsf{Crt}_\upsilon|.
\end{aligned}$$

So, $I_1 \ge (1 - C\iota^{1/2})\,\mathbb{E}\,|\mathsf{Crt}_\upsilon|$. Since (2.32) implies $I_1 \le 1$, and $\iota = o_N(1)$, the conclusion follows. □

Proof of Lemma 2.3.8. Set $\upsilon > 0$ small enough that Lemma 2.4.15 holds. Let $\mathscr{E}_1$ be the event that $\mathcal{F}^\varepsilon_{\mathsf{TAP}}$ has a critical point in $\mathcal{S}_{\varepsilon,\upsilon}$. By the Kac–Rice formula,

$$\begin{aligned}
\mathbb{P}(\mathscr{E} \cap \mathscr{E}_1) &\le \mathbb{E}[\mathbf{1}\{\mathscr{E} \cap \mathscr{E}_1\}|\mathsf{Crt}_\upsilon|] \\
&= \int_{\mathcal{S}_{\varepsilon,\upsilon}} \mathbb{E}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{Pl}}\left[|\det \nabla^2 \mathcal{F}^\varepsilon_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n})| \mathbf{1}\{\mathscr{E} \cap \mathscr{E}_1\}\right] \varphi_{\nabla \mathcal{F}^\varepsilon_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n})}(\mathbf{0})\ \mathsf{d}(\boldsymbol{m},\boldsymbol{n}).
\end{aligned}$$

This is bounded by

$$\begin{aligned}
&\int_{\mathcal{S}_{\varepsilon,\upsilon}} \mathbb{E}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{Pl}}\left[|\det \nabla^2 \mathcal{F}^\varepsilon_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n})|^2\right]^{1/2} \mathbb{P}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{Pl}}(\mathscr{E})^{1/2} \varphi_{\nabla \mathcal{F}^\varepsilon_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n})}(\mathbf{0})\ \mathsf{d}(\boldsymbol{m},\boldsymbol{n}) \\
&\le C \sup_{(\boldsymbol{m},\boldsymbol{n}) \in \mathcal{S}_{\varepsilon,\upsilon}} \mathbb{P}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{Pl}}(\mathscr{E})^{1/2} \times \int_{\mathcal{S}_{\varepsilon,\upsilon}} \mathbb{E}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{Pl}}\left[|\det \nabla^2 \mathcal{F}^\varepsilon_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n})|\right] \varphi_{\nabla \mathcal{F}^\varepsilon_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n})}(\mathbf{0})\ \mathsf{d}(\boldsymbol{m},\boldsymbol{n}) \\
&\le C \sup_{(\boldsymbol{m},\boldsymbol{n}) \in \mathcal{S}_{\varepsilon,\upsilon}} \mathbb{P}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{Pl}}(\mathscr{E})^{1/2} \cdot \mathbb{E}\,|\mathsf{Crt}_\upsilon| \overset{\textit{Lem. } 2.4.15}{\le} (1+o_N(1))C \sup_{(\boldsymbol{m},\boldsymbol{n}) \in \mathcal{S}_{\varepsilon,\upsilon}} \mathbb{P}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{Pl}}(\mathscr{E})^{1/2}. \qquad (2.33)
\end{aligned}$$

The result follows because $\mathbb{P}(\mathscr{E}) \le \mathbb{P}(\mathscr{E} \cap \mathscr{E}_1) + \mathbb{P}(\mathscr{E}_1^c)$, and $\mathbb{P}(\mathscr{E}_1^c) = o_N(1)$ by Lemma 2.4.13. □

2.4.6 Conditional law in planted model

Having proved the reduction to the planted model $\mathbb{P}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{Pl}}$, we now calculate the law of the disorder in it. This is stated in Lemma 2.4.17 for general $(\boldsymbol{m},\boldsymbol{n})$, and Corollary 2.4.18 for $(\boldsymbol{m},\boldsymbol{n}) \in \mathcal{S}_{\varepsilon,\upsilon}$. The following lemma is proved by direct differentiation of $\mathcal{F}^\varepsilon_{\mathsf{TAP}}$.

Lemma 2.4.16 (Proved in Appendix 2.A)**.** *Let $\boldsymbol{m} \in \mathbb{R}^N$, $\boldsymbol{n} \in \mathbb{R}^M$, and*

$$\acute{\boldsymbol{h}} = \frac{\boldsymbol{G}\boldsymbol{m}}{\sqrt{N}} + \varepsilon^{1/2}\widehat{\boldsymbol{g}} - \rho_\varepsilon(q(\boldsymbol{m}))\boldsymbol{n}, \qquad d_\varepsilon(\boldsymbol{m},\boldsymbol{n}) = \frac{1}{N}\sum_{a=1}^M (n_a - F_{\varepsilon,\rho_\varepsilon(q(\boldsymbol{m}))}(\acute{h}_a))^2 + F'_{\rho_\varepsilon(q(\boldsymbol{m}))}(\acute{h}_a).$$

Then,

$$\nabla_{\boldsymbol{m}} \mathcal{F}^\varepsilon_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n}) = -\mathrm{th}_\varepsilon^{-1}(\boldsymbol{m}) + \frac{\boldsymbol{G}^\top F_{\varepsilon,\rho_\varepsilon(q(\boldsymbol{m}))}(\acute{\boldsymbol{h}})}{\sqrt{N}} + \varepsilon^{1/2}\dot{\boldsymbol{g}} + \rho_\varepsilon'(q(\boldsymbol{m}))d_\varepsilon(\boldsymbol{m},\boldsymbol{n})\boldsymbol{m}, \qquad (2.34)$$

$$\nabla_{\boldsymbol{n}} \mathcal{F}^\varepsilon_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n}) = \rho_\varepsilon(q(\boldsymbol{m}))\left(\boldsymbol{n} - F_{\varepsilon,\rho_\varepsilon(q(\boldsymbol{m}))}(\acute{\boldsymbol{h}})\right). \qquad (2.35)$$

In particular $\nabla \mathcal{F}^\varepsilon_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n}) = \mathbf{0}$ if and only if, with $\dot{\boldsymbol{h}} = \mathrm{th}_\varepsilon^{-1}(\boldsymbol{m})$ and $\widehat{\boldsymbol{h}} = F^{-1}_{\varepsilon,\rho_\varepsilon(q(\boldsymbol{m}))}(\boldsymbol{n})$,

$$\frac{\boldsymbol{G}\boldsymbol{m}}{\sqrt{N}} + \varepsilon^{1/2}\widehat{\boldsymbol{g}} = \widehat{\boldsymbol{h}} + \rho_\varepsilon(q(\boldsymbol{m}))\boldsymbol{n}, \qquad (2.36)$$

$$\frac{\boldsymbol{G}^\top\boldsymbol{n}}{\sqrt{N}} + \varepsilon^{1/2}\dot{\boldsymbol{g}} = \dot{\boldsymbol{h}} - \rho_\varepsilon'(q(\boldsymbol{m}))d_\varepsilon(\boldsymbol{m},\boldsymbol{n})\boldsymbol{m}. \qquad (2.37)$$

(Note that (2.36) is equivalent to $\widehat{\boldsymbol{h}} = \acute{\boldsymbol{h}}$.)

Lemma 2.4.17. *Under $\mathbb{P}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{Pl}}$, $\boldsymbol{G}$ has law*

$$\frac{\boldsymbol{G}}{\sqrt{N}} \overset{d}{=} \frac{\widehat{\boldsymbol{h}}\boldsymbol{m}^\top}{N(q(\boldsymbol{m})+\varepsilon)} + \frac{\boldsymbol{n}\dot{\boldsymbol{h}}^\top}{N(\psi(\boldsymbol{n})+\varepsilon)} + \frac{\Delta(\boldsymbol{m},\boldsymbol{n})\boldsymbol{n}\boldsymbol{m}^\top}{N(q(\boldsymbol{m})+\psi(\boldsymbol{n})+\varepsilon)} + \frac{\widetilde{\boldsymbol{G}}}{\sqrt{N}}, \qquad \text{where} \tag{2.38}$$

$$\Delta(\boldsymbol{m},\boldsymbol{n}) = \rho_\varepsilon(q(\boldsymbol{m})) - \rho'_\varepsilon(q(\boldsymbol{m}))d_\varepsilon(\boldsymbol{m},\boldsymbol{n}) - \frac{\langle \boldsymbol{n}, \widehat{\boldsymbol{h}}\rangle}{N(q(\boldsymbol{m})+\varepsilon)} - \frac{\langle \boldsymbol{m}, \dot{\boldsymbol{h}}\rangle}{N(\psi(\boldsymbol{n})+\varepsilon)} \tag{2.39}$$

and where $\widetilde{\boldsymbol{G}}$ has the following law. Let $\dot{\mathbf{e}}_1,\dots,\dot{\mathbf{e}}_N$ and $\widehat{\mathbf{e}}_1,\dots,\widehat{\mathbf{e}}_M$ be orthonormal bases of $\mathbb{R}^N$ and $\mathbb{R}^M$ with $\dot{\mathbf{e}}_1 = \boldsymbol{m}/\|\boldsymbol{m}\|$ and $\widehat{\mathbf{e}}_1 = \boldsymbol{n}/\|\boldsymbol{n}\|$, and abbreviate $\widetilde{\boldsymbol{G}}(i,j) = \langle \widehat{\mathbf{e}}_j, \widetilde{\boldsymbol{G}}\dot{\mathbf{e}}_i\rangle$. Then the $\widetilde{\boldsymbol{G}}(i,j)$ are independent centered gaussians with variance

$$\mathbb{E}\,\widetilde{\boldsymbol{G}}(i,j)^2 = \begin{cases} \varepsilon/(q(\boldsymbol{m})+\psi(\boldsymbol{n})+\varepsilon) & i=j=1, \\ \varepsilon/(q(\boldsymbol{m})+\varepsilon) & i=1, j\neq 1, \\ \varepsilon/(\psi(\boldsymbol{n})+\varepsilon) & i\neq 1, j=1, \\ 1 & i\neq 1, j\neq 1. \end{cases} \tag{2.40}$$

Proof. This is a standard gaussian conditioning calculation, which we present briefly. For fixed $\dot{\boldsymbol{v}} \in \mathbb{R}^N$, $\widehat{\boldsymbol{v}} \in \mathbb{R}^M$ and

$$\begin{aligned} \widehat{\boldsymbol{w}} &= \frac{\langle \boldsymbol{m}, \dot{\boldsymbol{v}}\rangle}{N(q(\boldsymbol{m})+\varepsilon)}\widehat{\boldsymbol{v}} - \frac{\langle \boldsymbol{m}, \dot{\boldsymbol{v}}\rangle\langle \boldsymbol{n}, \widehat{\boldsymbol{v}}\rangle}{N^2(q(\boldsymbol{m})+\varepsilon)(q(\boldsymbol{m})+\psi(\boldsymbol{n})+\varepsilon)}\boldsymbol{n}, \\ \dot{\boldsymbol{w}} &= \frac{\langle \boldsymbol{n}, \widehat{\boldsymbol{v}}\rangle}{N(\psi(\boldsymbol{n})+\varepsilon)}\dot{\boldsymbol{v}} - \frac{\langle \boldsymbol{m}, \dot{\boldsymbol{v}}\rangle\langle \boldsymbol{n}, \widehat{\boldsymbol{v}}\rangle}{N^2(\psi(\boldsymbol{n})+\varepsilon)(q(\boldsymbol{m})+\psi(\boldsymbol{n})+\varepsilon)}\boldsymbol{m}, \end{aligned}$$

we may verify the independence

$$\frac{\langle \widehat{\boldsymbol{v}}, \boldsymbol{G}\dot{\boldsymbol{v}}\rangle}{\sqrt{N}} - \left\langle \widehat{\boldsymbol{w}}, \frac{\boldsymbol{G}\boldsymbol{m}}{\sqrt{N}} + \varepsilon^{1/2}\widehat{\boldsymbol{g}}\right\rangle - \left\langle \dot{\boldsymbol{w}}, \frac{\boldsymbol{G}^\top\boldsymbol{n}}{\sqrt{N}} + \varepsilon^{1/2}\dot{\boldsymbol{g}}\right\rangle \perp\!\!\!\perp \left\{\frac{\boldsymbol{G}\boldsymbol{m}}{\sqrt{N}} + \varepsilon^{1/2}\widehat{\boldsymbol{g}}, \frac{\boldsymbol{G}^\top\boldsymbol{n}}{\sqrt{N}} + \varepsilon^{1/2}\dot{\boldsymbol{g}}\right\}.$$

By Lemma 2.4.16, $\nabla\mathcal{F}^\varepsilon_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n}) = \boldsymbol{0}$ if and only if (2.36) and (2.37) hold. Let $\widehat{\boldsymbol{u}}, \dot{\boldsymbol{u}}$ denote the right-hand sides of (2.36), (2.37), respectively. Then, for all $\dot{\boldsymbol{v}}, \widehat{\boldsymbol{v}}$,

$$\mathbb{E}\left[\frac{\langle \widehat{\boldsymbol{v}}, \boldsymbol{G}\dot{\boldsymbol{v}}\rangle}{\sqrt{N}}\middle|(2.36),(2.37)\right] = \langle \widehat{\boldsymbol{w}}, \widehat{\boldsymbol{u}}\rangle + \langle \dot{\boldsymbol{w}}, \dot{\boldsymbol{u}}\rangle.$$

Expanding shows $\boldsymbol{G}$ has the conditional mean given in (2.38). The law (2.40) of $\widetilde{\boldsymbol{G}}$ follows from computing the covariance of the gaussian process

$$(\dot{\boldsymbol{v}}, \widehat{\boldsymbol{v}}) \mapsto \frac{\langle \widehat{\boldsymbol{v}}, \widetilde{\boldsymbol{G}}\dot{\boldsymbol{v}}\rangle}{\sqrt{N}} \equiv \frac{\langle \widehat{\boldsymbol{v}}, \boldsymbol{G}\dot{\boldsymbol{v}}\rangle}{\sqrt{N}} - \left\langle \widehat{\boldsymbol{w}}, \frac{\boldsymbol{G}\boldsymbol{m}}{\sqrt{N}} + \varepsilon^{1/2}\widehat{\boldsymbol{g}}\right\rangle - \left\langle \dot{\boldsymbol{w}}, \frac{\boldsymbol{G}^\top\boldsymbol{n}}{\sqrt{N}} + \varepsilon^{1/2}\dot{\boldsymbol{g}}\right\rangle.$$

(Note that if $\widehat{\boldsymbol{v}} \in \{\widehat{\boldsymbol{e}}_2,\dots,\widehat{\boldsymbol{e}}_M\}$, then $\langle \boldsymbol{n}, \widehat{\boldsymbol{v}}\rangle = 0$ and thus $\dot{\boldsymbol{w}} = \boldsymbol{0}$. Similarly if $\dot{\boldsymbol{v}} \in \{\dot{\boldsymbol{e}}_2,\dots,\dot{\boldsymbol{e}}_N\}$, then $\widehat{\boldsymbol{w}} = \boldsymbol{0}$. So in most cases the above formulas simplify considerably.) □

Corollary 2.4.18. *If $(\boldsymbol{m},\boldsymbol{n}) \in \mathcal{S}_{\varepsilon,\upsilon}$, then under $\mathbb{P}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{Pl}}$, $\boldsymbol{G}$ has law*

$$\frac{\boldsymbol{G}}{\sqrt{N}} \overset{d}{=} \frac{(1+o_\upsilon(1))\widehat{\boldsymbol{h}}\boldsymbol{m}^\top}{N(q_\varepsilon+\varepsilon)} + \frac{(1+o_\upsilon(1))\boldsymbol{n}\dot{\boldsymbol{h}}^\top}{N(\psi_\varepsilon+\varepsilon)} + \frac{o_\upsilon(1)\boldsymbol{n}\boldsymbol{m}^\top}{N} + \frac{\widetilde{\boldsymbol{G}}}{\sqrt{N}}, \tag{2.41}$$

where $o_\upsilon(1)$ denotes a term vanishing as $\upsilon \to 0$ and $\widetilde{\boldsymbol{G}}$ is as in Lemma 2.4.17.

This corollary is proved by a standard approximation argument, which we record as Fact 2.4.20 below.

Definition 2.4.19. A function $f : \mathbb{R} \to \mathbb{R}$ is $(2, L)$-pseudo-Lipschitz if $|f(x) - f(y)| \le L|x-y|(|x|+|y|+1)$.

Fact 2.4.20 (Proved in Appendix 2.A)**.** *Suppose $\mu, \mu' \in \mathcal{P}_2(\mathbb{R})$ and let $\mu_2 = \mathbb{E}_{x\sim\mu}[x^2]$. If f is $(2, L)$-pseudo-Lipschitz, then*

$$|\mathbb{E}_\mu[f] - \mathbb{E}_{\mu'}[f]| \le 3L\mathbb{W}_2(\mu, \mu')(\mu_2 + \mathbb{W}_2(\mu, \mu') + 1).$$

Proof of Corollary 2.4.18. Let $\dot{\boldsymbol{h}} = \mathrm{th}_\varepsilon^{-1}(\boldsymbol{m})$, $\widehat{\boldsymbol{h}} = F_{\varepsilon,\varrho_\varepsilon}^{-1}(\boldsymbol{n})$, so $(\dot{\boldsymbol{h}}, \widehat{\boldsymbol{h}}) \in \mathcal{T}_{\varepsilon,\upsilon}$. Recall $\mu_{\dot{\boldsymbol{h}}}$ defined in (2.26). Since $\dot{h} \mapsto \mathrm{th}_\varepsilon(\dot{h})^2$ is $(2, O(1))$-pseudo-Lipschitz, by Fact 2.4.20,

$$|q(\boldsymbol{m}) - q_\varepsilon| = \left|\mathbb{E}_{\dot{h}\sim\mu_{\dot{\boldsymbol{h}}}}[\mathrm{th}_\varepsilon(\dot{h})^2] - \mathbb{E}_{\dot{h}\sim\mathcal{N}(0,\psi_\varepsilon+\varepsilon)}[\mathrm{th}_\varepsilon(\dot{h})^2]\right| = o_\upsilon(1).$$

Similarly $\psi(\boldsymbol{n}) = \psi_\varepsilon + o_\upsilon(1)$ and $d_\varepsilon(\boldsymbol{m}, \boldsymbol{n}) = d_\varepsilon + o_\upsilon(1)$. Also, by gaussian integration by parts and Lemma 2.4.2,

$$\mathbb{E}_{\dot{h}\sim\mathcal{N}(0,\psi_\varepsilon+\varepsilon)}[\dot{h}\mathrm{th}_\varepsilon(\dot{h})] = (\psi_\varepsilon + \varepsilon)\varrho_\varepsilon.$$

Thus

$$\left|\frac{\langle \boldsymbol{m}, \dot{\boldsymbol{h}}\rangle}{N(\psi_\varepsilon + \varepsilon)} - \varrho_\varepsilon\right| = \left|\mathbb{E}_{\dot{h}\sim\mu_{\dot{\boldsymbol{h}}}}[\dot{h}\mathrm{th}_\varepsilon(\dot{h})] - \mathbb{E}_{\dot{h}\sim\mathcal{N}(0,\psi_\varepsilon+\varepsilon)}[\dot{h}\mathrm{th}_\varepsilon(\dot{h})]\right| = o_\upsilon(1).$$

Similarly $\frac{\langle \boldsymbol{n}, \widehat{\boldsymbol{h}}\rangle}{N(q_\varepsilon+\varepsilon)} = d_\varepsilon + o_\upsilon(1)$. Finally, equation (2.22) and regularity of $\rho_\varepsilon, \rho'_\varepsilon$ (recall (2.23)) imply

$$\rho_\varepsilon(q(\boldsymbol{m})) = \varrho_\varepsilon + o_\upsilon(1), \qquad\qquad \rho'_\varepsilon(q(\boldsymbol{m})) = -1 + o_\upsilon(1).$$

Combining these estimates shows the conditional mean of $\boldsymbol{G}$ in (2.38) simplifies to the form (2.41). In particular note that $\Delta(\boldsymbol{m}, \boldsymbol{n}) = o_\upsilon(1)$. $\square$

2.4.7 Deferred proofs

We now prove various results deferred from the above proof.

Lemma 2.4.21 ([DS18, Lemma 10.1])**.** *The function $\mathcal{E}$ satisfies the following for all $x \in \mathbb{R}$.*

(a) $0 \le \mathcal{E}(x) \le |x| + 1$.

(b) $\mathcal{E}'(x) = \mathcal{E}(x)(\mathcal{E}(x) - x) \in (0, 1)$.

(c) $\mathcal{E}''(x) \in (0, 1)$.

(d) $\mathcal{E}^{(3)} \in (-1/2, 13)$.

Proof of Lemma 2.4.2. We calculate

$$\begin{aligned}
q_\varepsilon &= \mathbb{E}[\mathrm{th}_\varepsilon((\psi_\varepsilon + \varepsilon)^{1/2}Z)^2] \\
&= \varepsilon^2(\psi_\varepsilon + \varepsilon) + 2\varepsilon\,\mathbb{E}[(\psi_\varepsilon + \varepsilon)^{1/2}Z\mathrm{th}((\psi_\varepsilon + \varepsilon)^{1/2}Z)] + \mathbb{E}[\mathrm{th}^2((\psi_\varepsilon + \varepsilon)^{1/2}Z)^2] \\
&= \varepsilon^2(\psi_\varepsilon + \varepsilon) + 2\varepsilon(\psi_\varepsilon + \varepsilon)\,\mathbb{E}[1 - \mathrm{th}^2((\psi_\varepsilon + \varepsilon)^{1/2}Z)] + \mathbb{E}[\mathrm{th}^2((\psi_\varepsilon + \varepsilon)^{1/2}Z)^2].
\end{aligned}$$

Thus

$$\mathbb{E}[\mathrm{th}^2((\psi_\varepsilon + \varepsilon)^{1/2}Z)^2] = \frac{q_\varepsilon - 2\varepsilon(\psi_\varepsilon + \varepsilon) - \varepsilon^2(\psi_\varepsilon + \varepsilon)}{1 - 2\varepsilon(\psi_\varepsilon + \varepsilon)},$$

and

$$\mathbb{E}[\mathrm{th}'_\varepsilon((\psi_\varepsilon + \varepsilon)^{1/2}Z)] = 1 + \varepsilon - \mathbb{E}[\mathrm{th}^2((\psi_\varepsilon + \varepsilon)^{1/2}Z)] = \frac{1 - q_\varepsilon + \varepsilon - \varepsilon^2(\psi_\varepsilon + \varepsilon)}{1 - 2\varepsilon(\psi_\varepsilon + \varepsilon)} = \varrho_\varepsilon.$$

$\square$

Differentiating (2.19) and applying Lemma 2.4.21(b) shows the following fact.

Fact 2.4.22. *For $\varepsilon, \varrho > 0$ and any $x \in \mathbb{R}$,*

$$-\frac{1+\varepsilon^2}{\varrho + \varepsilon(1+\varepsilon\varrho)} \le F'_{\varepsilon,\varrho}(x) \le -\frac{\varepsilon}{1+\varepsilon\varrho}.$$

Thus

$$1 + \varrho F'_{\varepsilon,\varrho}(x) \ge \frac{\varepsilon}{\varrho + \varepsilon(1+\varepsilon\varrho)}. \tag{2.42}$$

For ϱ in any compact set away from 0, $|F'_{\varepsilon,\varrho}|$, $|F''_{\varepsilon,\varrho}|$ and $|F^{(3)}_{\varepsilon,\varrho}|$ are uniformly bounded independently of ε.

Proof of Proposition 2.4.6. It is clear that $\nabla^2_{\boldsymbol{n},\boldsymbol{n}} \mathcal{F}^{\varepsilon}_{\mathsf{TAP}}(\boldsymbol{m}, \boldsymbol{n})$ is diagonal, so it suffices to check $\partial^2_{n_a} \mathcal{F}^{\varepsilon}_{\mathsf{TAP}}(\boldsymbol{m}, \boldsymbol{n}) \ge \eta$ for all $a \in [M]$. We calculate

$$\begin{aligned}\partial^2_{n_a} \mathcal{F}^{\varepsilon}_{\mathsf{TAP}}(\boldsymbol{m}, \boldsymbol{n}) &= \rho_\varepsilon(q(\boldsymbol{m})) \left(1 + \rho_\varepsilon(q(\boldsymbol{m})) F'_{\varepsilon,\varrho}\left(\frac{\langle \boldsymbol{g}^a, \boldsymbol{m}\rangle}{\sqrt{N}} + \varepsilon^{1/2}\widehat{g}_a - \rho_\varepsilon(q(\boldsymbol{m})) n_a\right)\right) \\ &\overset{(2.42)}{\ge} \frac{\varepsilon \rho_\varepsilon(q(\boldsymbol{m}))}{\rho_\varepsilon(q(\boldsymbol{m})) + \varepsilon(1 + \varepsilon\rho_\varepsilon(q(\boldsymbol{m})))}.\end{aligned}$$

Since $\rho_\varepsilon \in [C_{\mathsf{bd}}^{-1}, C_{\mathsf{bd}}]$ the result follows. □

Proof of Lemma 2.4.14. The function $x \mapsto \rho_\varepsilon(q(\boldsymbol{m})) F_{\varepsilon,\rho_\varepsilon(q(\boldsymbol{m}))}(x)$ is uniformly Lipschitz over $\boldsymbol{m} \in \mathbb{R}^N$, because $\rho_\varepsilon(q(\boldsymbol{m})) \in [C_{\mathsf{bd}}^{-1}, C_{\mathsf{bd}}]$. Note that $\widehat{\boldsymbol{g}}$ appears in (2.35) through the term $\varepsilon^{1/2}\widehat{\boldsymbol{g}}$ in $\acute{\boldsymbol{h}}$ and is independent of all other terms apeparing in (2.35). Thus $\varphi_{\nabla_{\boldsymbol{n}} \mathcal{F}^{\varepsilon}_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n})}(\boldsymbol{z})$ is bounded, and continuous for $\boldsymbol{z}$ in an neighborhood of $\boldsymbol{0}$, uniformly in $\boldsymbol{m}, \boldsymbol{n}$. Similarly, $\dot{\boldsymbol{g}}$ appears in (2.34), (2.35) only as the term $\varepsilon^{1/2}\dot{\boldsymbol{g}}$ in (2.34). This implies the conclusion. □

2.5 Analysis of AMP

In this section, we prove items (a), (b), and (d) of Proposition 2.4.8. Item (c) will be proved in Section 2.6.

2.5.1 Scalar recursions

For $q \in [0, q_\varepsilon]$, $\psi \in [0, \psi_\varepsilon]$, define

$$\begin{aligned} P_{\mathrm{AMP}}(\psi) &= \mathbb{E}[\mathrm{th}_\varepsilon((\psi+\varepsilon)^{1/2} Z + (\psi_\varepsilon - \psi)^{1/2} Z') \mathrm{th}_\varepsilon((\psi+\varepsilon)^{1/2} Z + (\psi_\varepsilon - \psi)^{1/2} Z'')], \\ R_{\mathrm{AMP}}(q) &= \alpha_\star \, \mathbb{E}[F_{\varepsilon,\varrho_\varepsilon}((q+\varepsilon)^{1/2} Z + (q_\varepsilon - q)^{1/2} Z') F_{\varepsilon,\varrho_\varepsilon}((q+\varepsilon)^{1/2} Z + (q_\varepsilon - q)^{1/2} Z'')], \end{aligned}$$

Define the sequences $(\overline{q}_k)_{k\ge 0}$ and $(\overline{\psi}_k)_{k\ge 1}$ by $\overline{q}_0 = 0$ and the recursion

$$\overline{\psi}_{k+1} = R_{\mathrm{AMP}}(\overline{q}_k), \qquad \overline{q}_k = P_{\mathrm{AMP}}(\overline{\psi}_k).$$

Lemma 2.5.1. *The sequences $(\overline{q}_k)_{k\ge 0}$, $(\overline{\psi}_k)_{k\ge 1}$ are increasing, and for small ε, we have $\overline{q}_k \uparrow q_\varepsilon$ and $\overline{\psi}_k \uparrow \psi_\varepsilon$.*

Proof. Let the functions

$$\widetilde{\mathrm{th}}_\varepsilon(x) = \mathrm{th}_\varepsilon((\psi_\varepsilon + \varepsilon)^{1/2} x), \qquad \widetilde{F}_\varepsilon(x) = F_{\varepsilon,\varrho_\varepsilon}((q_\varepsilon + \varepsilon)^{1/2} x)$$

have Hermite expansions

$$\widetilde{\mathrm{th}}_\varepsilon(x) = \sum_{p\ge 0} a_p H_p(x), \qquad \widetilde{F}_\varepsilon(x) = \sum_{p\ge 0} b_p H_p(x),$$

where $H_p(x)$ is the p-th Hermite polynomial, normalized to $\mathbb{E}\, H_p(Z)^2 = 1$. Then

$$P_{\mathrm{AMP}}(\psi) = \sum_{p\ge 0} a_p^2 \left(\frac{\psi+\varepsilon}{\psi_\varepsilon + \varepsilon}\right)^p, \qquad R_{\mathrm{AMP}}(q) = \alpha_\star \sum_{p\ge 0} b_p^2 \left(\frac{q+\varepsilon}{q_\varepsilon+\varepsilon}\right)^p.$$

So, P_{AMP} and R_{AMP} are increasing and convex. Thus $(\overline{q}_k)_{k\geq 0}$, $(\overline{\psi}_k)_{k\geq 1}$ are increasing, and their limit is the smallest fixed point of $P_{\mathrm{AMP}} \circ R_{\mathrm{AMP}}$. It remains to show this fixed point is $(q_\varepsilon, \psi_\varepsilon)$. By definition of $q_\varepsilon, \psi_\varepsilon$, $(q_\varepsilon, \psi_\varepsilon)$ is a fixed point. Since $P_{\mathrm{AMP}} \circ R_{\mathrm{AMP}}$ is convex, it suffices to show $(P_{\mathrm{AMP}} \circ R_{\mathrm{AMP}})'(q_\varepsilon) < 1$. Note that

$$(P_{\mathrm{AMP}} \circ R_{\mathrm{AMP}})'(q_\varepsilon) = P'_{\mathrm{AMP}}(\psi_\varepsilon) R'_{\mathrm{AMP}}(q_\varepsilon).$$

By gaussian integration by parts,

$$\begin{aligned}
P'_{\mathrm{AMP}}(\psi) &= \mathbb{E}[\mathrm{th}'_\varepsilon((\psi+\varepsilon)^{1/2} Z + (\psi_\varepsilon - \psi)^{1/2} Z') \mathrm{th}'_\varepsilon((\psi+\varepsilon)^{1/2} Z + (\psi_\varepsilon - \psi)^{1/2} Z'')], \\
R'_{\mathrm{AMP}}(q) &= \alpha_\star \, \mathbb{E}[F'_{\varepsilon,\varrho_\varepsilon}((q+\varepsilon)^{1/2} Z + (q_\varepsilon - q)^{1/2} Z') F'_{\varepsilon,\varrho_\varepsilon}((q+\varepsilon)^{1/2} Z + (q_\varepsilon - q)^{1/2} Z'')],
\end{aligned}$$

and in particular

$$P'_{\mathrm{AMP}}(\psi_\varepsilon) = \mathbb{E}[\mathrm{th}'_\varepsilon((\psi_\varepsilon+\varepsilon)^{1/2} Z)^2], \qquad R'_{\mathrm{AMP}}(q_\varepsilon) = \alpha_\star \, \mathbb{E}[F'_{\varepsilon,\varrho_\varepsilon}((q_\varepsilon+\varepsilon)^{1/2} Z)^2].$$

In light of Proposition 2.4.1, a simple continuity argument shows

$$\mathbb{E}[\mathrm{th}'_\varepsilon((\psi_\varepsilon+\varepsilon)^{1/2} Z)^2] \overset{\varepsilon\downarrow 0}{\to} \mathbb{E}[\mathrm{th}'(\psi_0^{1/2} Z)^2], \qquad \mathbb{E}[F'_{\varepsilon,\varrho_\varepsilon}((q_\varepsilon+\varepsilon)^{1/2} Z)^2] \overset{\varepsilon\downarrow 0}{\to} \mathbb{E}[F'_{1-q_0}(q_0^{1/2} Z)^2].$$

Thus,

$$\begin{aligned}
(P_{\mathrm{AMP}} \circ R_{\mathrm{AMP}})'(q_\varepsilon) &= \alpha_\star \, \mathbb{E}[\mathrm{th}'_\varepsilon((\psi_\varepsilon+\varepsilon)^{1/2} Z)^2] \, \mathbb{E}[F'_{\varepsilon,\varrho_\varepsilon}((q_\varepsilon+\varepsilon)^{1/2} Z)^2] \\
&\overset{\varepsilon\downarrow 0}{\to} \alpha_\star \, \mathbb{E}[\mathrm{th}'(\psi_0^{1/2} Z)^2] \, \mathbb{E}[F'_{1-q_0}(q_0^{1/2} Z)^2] \overset{\text{Cond. } 2.3.3}{<} 1.
\end{aligned}$$

Thus, $(R_{\mathrm{AMP}} \circ P_{\mathrm{AMP}})'(q_\varepsilon) < 1$ for sufficiently small ε. Hence $\overline{q}_k \uparrow q_\varepsilon$ and $\overline{\psi}_k \uparrow \psi_\varepsilon$. □

2.5.2 State evolution

The limiting overlap structure of the AMP iterates in the null model follows directly from the state evolution of [Bol14, BM11, JM13, BMN20]. Define the infinite arrays $(\dot{\Sigma}_{i,j} : i,j \geq 1)$ and $(\widehat{\Sigma}_{i,j} : i,j \geq 0)$ by

$$\dot{\Sigma}_{i,j} = \begin{cases} \psi_\varepsilon & i = j, \\ \overline{\psi}_{i\wedge j} & i \neq j, \end{cases} \qquad \widehat{\Sigma}_{i,j} = \begin{cases} q_\varepsilon & i = j, \\ \overline{q}_{i\wedge j} & i \neq j. \end{cases}$$

For any $k \geq 0$, let $\dot{\Sigma}_{\leq k} \in \mathbb{R}^{k\times k}$ and $\widehat{\Sigma}^+_{\leq k} \in \mathbb{R}^{(k+1)\times(k+1)}$ denote the sub-arrays indexed by $i,j \leq k$.

Proposition 2.5.2. *For any $k \geq 0$, as $N \to \infty$ the empirical coordinate distribution of the AMP iterates converges in $\mathbb{W}_2$ in probability under $\mathbb{P}$, to*

$$\frac{1}{N}\sum_{i=1}^N \delta(\dot{h}_i^1, \ldots, \dot{h}_i^k) \overset{\mathbb{W}_2}{\to} \mathcal{N}(0, \dot{\Sigma}_{\leq k} + \varepsilon \mathbf{1}\mathbf{1}^\top), \qquad \frac{1}{M}\sum_{a=1}^M \delta(\widehat{h}_a^0, \ldots, \widehat{h}_a^k) \overset{\mathbb{W}_2}{\to} \mathcal{N}(0, \widehat{\Sigma}_{\leq k} + \varepsilon \mathbf{1}\mathbf{1}^\top). \tag{2.43}$$

Proof. The state evolution [BMN20, Theorem 1] implies that (in probability)

$$\frac{1}{N}\sum_{i=1}^N \delta(\dot{h}_i^1, \ldots, \dot{h}_i^k) \overset{\mathbb{W}_2}{\to} \mathcal{N}(0, \dot{\Sigma}^{(0)}_{\leq k} + \varepsilon \mathbf{1}\mathbf{1}^\top), \qquad \frac{1}{M}\sum_{a=1}^M \delta(\widehat{h}_a^0, \ldots, \widehat{h}_a^k) \overset{\mathbb{W}_2}{\to} \mathcal{N}(0, \widehat{\Sigma}^{(0)}_{\leq k} + \varepsilon \mathbf{1}\mathbf{1}^\top).$$

holds for arrays $\dot{\Sigma}^{(0)}$, $\widehat{\Sigma}^{(0)}$ defined as follows. As initialization, $\widehat{\Sigma}^{(0)}_{0,i} = \widehat{\Sigma}^{(0)}_{i,0} = \widehat{\Sigma}_{0,i}$ for all $i \geq 0$. Then, for $(\widehat{H}_0, \ldots, \widehat{H}_k) \sim \mathcal{N}(0, \widehat{\Sigma}^{(0)}_{\leq k} + \varepsilon \mathbf{1}\mathbf{1}^\top)$ and $0 \leq i \leq k$, define recursively

$$\dot{\Sigma}^{(0)}_{k+1,i+1} = \dot{\Sigma}^{(0)}_{i+1,k+1} = \alpha_\star \, \mathbb{E}[F_{\varepsilon,\varrho_\varepsilon}(\widehat{H}_i) F_{\varepsilon,\varrho_\varepsilon}(\widehat{H}_k)].$$

For $(\dot{H}_0, \ldots, \dot{H}_{k+1}) \sim \mathcal{N}(0, \dot{\Sigma}^{(0)}_{\leq k+1} + \varepsilon \mathbf{1}\mathbf{1}^\top)$ and $1 \leq i \leq k+1$, let

$$\widehat{\Sigma}^{(0)}_{k+1,i} = \widehat{\Sigma}^{(0)}_{i,k+1} = \mathbb{E}[\mathrm{th}_\varepsilon(\dot{H}_i) \mathrm{th}_\varepsilon(\dot{H}_{k+1})].$$

It remains to show $\dot{\Sigma}^{(0)}, \widehat{\Sigma}^{(0)}$ coincide with $\dot{\Sigma}, \widehat{\Sigma}$. Since $\widehat{\Sigma}_{0,0} = q_\varepsilon$, induction shows the diagonal entries are

$$\dot{\Sigma}^{(0)}_{k,k} = \psi_\varepsilon = \dot{\Sigma}_{k,k}, \qquad\qquad \widehat{\Sigma}^{(0)}_{k,k} = q_\varepsilon = \widehat{\Sigma}_{k,k}.$$

Then, the above recursion gives $\dot{\Sigma}^{(0)}_{i+1,j+1} = R_{\mathrm{AMP}}(\widehat{\Sigma}^{(0)}_{i,j})$, $\widehat{\Sigma}^{(0)}_{i,j} = P_{\mathrm{AMP}}(\dot{\Sigma}^{(0)}_{i,j})$. By induction, for $i \neq j$,

$$\dot{\Sigma}^{(0)}_{i,j} = \overline{\psi}_{i\wedge j} = \dot{\Sigma}_{i,j}, \qquad\qquad \widehat{\Sigma}^{(0)}_{i,j} = \overline{q}_{i\wedge j} = \widehat{\Sigma}_{i,j}.$$

Thus $\dot{\Sigma}^{(0)} = \dot{\Sigma}$ and $\widehat{\Sigma}^{(0)} = \widehat{\Sigma}$. □

The following proposition characterizes the limiting overlap structure in the planted model. To conserve notation, we will denote the planted solution by $(\boldsymbol{m}, \boldsymbol{n})$, rather than $(\boldsymbol{m}', \boldsymbol{n}')$ as in Proposition 2.4.8.

Proposition 2.5.3. *Let $(\boldsymbol{m}, \boldsymbol{n}) \in \mathcal{S}_{\varepsilon, o_N(1)}$, $\dot{\boldsymbol{h}} = \mathrm{th}_\varepsilon^{-1}(\boldsymbol{m})$, $\widehat{\boldsymbol{h}} = F^{-1}_{\varepsilon,\varrho_\varepsilon}(\boldsymbol{n})$, and $(\boldsymbol{G}, \dot{\boldsymbol{g}}, \widehat{\boldsymbol{g}}) \sim \mathbb{P}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{Pl}}$. For any $k \geq 0$, as $N \to \infty$ the empirical coordinate distribution of $(\dot{\boldsymbol{h}}, \widehat{\boldsymbol{h}})$ and the AMP iterates converges in $\mathbb{W}_2$ in probability under $\mathbb{P}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{Pl}}$, to*

$$\frac{1}{N}\sum_{i=1}^N \delta(\dot{h}^1_i, \ldots, \dot{h}^k_i, \dot{h}_i) \xrightarrow{\mathbb{W}_2} \mathcal{N}(0, \dot{\Sigma}_{\leq k+1} + \varepsilon \boldsymbol{1}\boldsymbol{1}^\top), \qquad \frac{1}{M}\sum_{a=1}^M \delta(\widehat{h}^0_a, \ldots, \widehat{h}^k_a, \widehat{h}_a) \xrightarrow{\mathbb{W}_2} \mathcal{N}(0, \widehat{\Sigma}_{\leq k+1} + \varepsilon \boldsymbol{1}\boldsymbol{1}^\top).$$

We prove this proposition by introducing an auxiliary AMP iteration. We fix $\boldsymbol{m}, \boldsymbol{n}, \dot{\boldsymbol{h}}, \widehat{\boldsymbol{h}}$ as in Proposition 2.5.3. Let $\widetilde{\boldsymbol{G}} \in \mathbb{R}^{M\times N}$ be given by (2.40) and $\widehat{\boldsymbol{G}} \in \mathbb{R}^{M\times N}$ have i.i.d. $\mathcal{N}(0,1)$ entries, and couple these matrices so that a.s.

$$P^\perp_{\boldsymbol{n}} \widetilde{\boldsymbol{G}} P^\perp_{\boldsymbol{m}} = P^\perp_{\boldsymbol{n}} \widehat{\boldsymbol{G}} P^\perp_{\boldsymbol{m}}, \tag{2.44}$$

and, with $\overline{\boldsymbol{G}}$ denoting this common value, $\widetilde{\boldsymbol{G}} - \overline{\boldsymbol{G}}$ and $\widehat{\boldsymbol{G}} - \overline{\boldsymbol{G}}$ are independent. Further, let $Z \sim \mathcal{N}(0,1)$, $\dot{\boldsymbol{\xi}} \sim \mathcal{N}(0, \boldsymbol{I}_N)$, $\widehat{\boldsymbol{\xi}} \sim \mathcal{N}(0, \boldsymbol{I}_M)$ be coupled to $\widetilde{\boldsymbol{G}}$ such that

$$\widetilde{\boldsymbol{G}} + \boldsymbol{\Delta} = \overline{\boldsymbol{G}} - \sqrt{\frac{\varepsilon}{q(\boldsymbol{m})+\varepsilon}} \cdot \frac{\widehat{\boldsymbol{\xi}}\boldsymbol{m}^\top}{\|\boldsymbol{m}\|} - \sqrt{\frac{\varepsilon}{\psi(\boldsymbol{n})+\varepsilon}} \cdot \frac{\boldsymbol{n}\dot{\boldsymbol{\xi}}^\top}{\|\boldsymbol{n}\|}, \qquad \text{where} \tag{2.45}$$

$$\boldsymbol{\Delta} = \sqrt{\frac{\varepsilon}{q(\boldsymbol{m})+\varepsilon} + \frac{\varepsilon}{\psi(\boldsymbol{n})+\varepsilon} - \frac{\varepsilon}{q(\boldsymbol{m})+\psi(\boldsymbol{n})+\varepsilon}}\, \frac{\boldsymbol{n}\boldsymbol{m}^\top}{\|\boldsymbol{n}\|\|\boldsymbol{m}\|} Z. \tag{2.46}$$

(Such a coupling exists by (2.40).) The auxiliary AMP iteration has initialization $\boldsymbol{n}^{(1),-1} = \boldsymbol{0}$, $\boldsymbol{m}^{(1),0} = q_\varepsilon^{1/2}\boldsymbol{1}$, and iteration

$$\boldsymbol{m}^{(1),k} = \mathrm{th}_\varepsilon(\dot{\boldsymbol{h}}^{(1),k}), \qquad\qquad \boldsymbol{n}^{(1),k} = F_{\varepsilon,\varrho_\varepsilon}(\widehat{\boldsymbol{h}}^{(1),k}),$$

for $\dot{\boldsymbol{h}}^{(1),k}, \widehat{\boldsymbol{h}}^{(1),k}$ as follows. Let $\overline{\psi}_0 = 0$, and

$$\widehat{\boldsymbol{h}}^{(1),k} = \frac{1}{\sqrt{N}}\widehat{\boldsymbol{G}}\left(\boldsymbol{m}^{(1),k} - \frac{\overline{q}_k}{q_\varepsilon}\boldsymbol{m}\right) + \frac{\sqrt{\varepsilon}(q_\varepsilon - \overline{q}_k)}{\sqrt{q_\varepsilon(q_\varepsilon+\varepsilon)}}\widehat{\boldsymbol{\xi}} + \frac{\overline{q}_k+\varepsilon}{q_\varepsilon+\varepsilon}\widehat{\boldsymbol{h}} - \varrho_\varepsilon\left(\boldsymbol{n}^{(1),k-1} - \frac{\overline{\psi}_k}{\psi_\varepsilon}\boldsymbol{n}\right) \tag{2.47}$$

$$\dot{\boldsymbol{h}}^{(1),k+1} = \frac{1}{\sqrt{N}}\widehat{\boldsymbol{G}}^\top\left(\boldsymbol{n}^{(1),k} - \frac{\overline{\psi}_{k+1}}{\psi_\varepsilon}\boldsymbol{n}\right) + \frac{\sqrt{\varepsilon}(\psi_\varepsilon - \psi_{k+1})}{\sqrt{\psi_\varepsilon(\psi_\varepsilon+\varepsilon)}}\dot{\boldsymbol{\xi}} + \frac{\overline{\psi}_{k+1}+\varepsilon}{\psi_\varepsilon+\varepsilon}\dot{\boldsymbol{h}} - d_\varepsilon\left(\boldsymbol{m}^{(1),k} - \frac{\overline{q}_k}{q_\varepsilon}\boldsymbol{m}\right).$$

Define augmented arrays $(\dot{\Sigma}^+_{i,j} : i,j \in \{\diamond, \bowtie\} \cup \mathbb{Z}_{\geq 1})$ and $(\widehat{\Sigma}^+_{i,j} : i,j \in \{\diamond, \bowtie\} \cup \mathbb{Z}_{\geq 0})$ by

$$\dot{\Sigma}^+_{i,j} = \begin{cases} \psi_\varepsilon + \varepsilon & i = j \geq 1 \text{ or } i = j = \diamond, \\ \overline{\psi}_j + \varepsilon & i > j \geq 1, \\ \overline{\psi}_i + \varepsilon & i \geq 1, j = \diamond, \\ \frac{\sqrt{\varepsilon}(\psi_\varepsilon - \overline{\psi}_i)}{\sqrt{\psi_\varepsilon(\psi_\varepsilon+\varepsilon)}} & i \geq 1, j = \bowtie, \\ 1 & i = j = \bowtie, \\ 0 & i = \diamond, j = \bowtie, \end{cases} \qquad \widehat{\Sigma}^+_{i,j} = \begin{cases} q_\varepsilon + \varepsilon & i = j \geq 0 \text{ or } i = j = \diamond, \\ \overline{q}_j + \varepsilon & i > j \geq 0, \\ \overline{q}_i + \varepsilon & i \geq 0, j = \diamond, \\ \frac{\sqrt{\varepsilon}(q_\varepsilon - \overline{q}_i)}{\sqrt{q_\varepsilon(q_\varepsilon+\varepsilon)}} & i \geq 0, j = \bowtie, \\ 1 & i = j = \bowtie, \\ 0 & i = \diamond, j = \bowtie, \end{cases}$$

with the remaining entries defined by symmetry over the diagonal. Note that on indices (i,j) where $\{i,j\} \cap \{\diamond, \bowtie\} = \emptyset$, these arrays coincide with $\dot{\Sigma} + \varepsilon \mathbf{1}\mathbf{1}^\top$ and $\widehat{\Sigma} + \varepsilon \mathbf{1}\mathbf{1}^\top$. Let $\dot{\Sigma}^+_{\le k} \in \mathbb{R}^{(k+2)\times(k+2)}$ and $\widehat{\Sigma}^+_{\le k} \in \mathbb{R}^{(k+3)\times(k+3)}$ denote the sub-arrays indexed by $\{\diamond, \bowtie\}$ and $\{1,\dots,k\}$ (resp. $\{0,\dots,k\}$).

Proposition 2.5.4 (Proved in Appendix 2.A)**.** *For any $k \ge 0$, as $N \to \infty$ we have the convergence in $\mathbb{W}_2$ in probability under $\mathbb{P}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{Pl}}$*

$$\frac{1}{N}\sum_{i=1}^N \delta(\dot{h}_i, \dot{\xi}_i, \dot{h}_i^{(1),1}, \dots, \dot{h}_i^{(1),k}) \xrightarrow{\mathbb{W}_2} \mathcal{N}(0, \dot{\Sigma}^+_{\le k}), \qquad \frac{1}{M}\sum_{a=1}^M \delta(\widehat{h}_a, \widehat{\xi}_a, \widehat{h}_a^{(1),0}, \dots, \widehat{h}_a^{(1),k}) \xrightarrow{\mathbb{W}_2} \mathcal{N}(0, \widehat{\Sigma}^+_{\le k}).$$

This is proved by applying state evolution, analogously to Proposition 2.5.2. We next show that this AMP iteration approximates the original one, in the following sense.

Proposition 2.5.5 (Proved in Appendix 2.A)**.** *For any $k \ge 0$, as $N \to \infty$ we have $\|\widehat{\boldsymbol{h}}^{(1),k} - \widehat{\boldsymbol{h}}^k\|/\sqrt{N} \to 0$ in probability under $\mathbb{P}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{Pl}}$ and if $k \ge 1$, $\|\dot{\boldsymbol{h}}^{(1),k} - \dot{\boldsymbol{h}}^k\|/\sqrt{N} \to 0$ in probability under $\mathbb{P}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{Pl}}$.*

Proof of Proposition 2.5.3. If we identify index $\diamond$ with $k+1$, the array $\{\dot{\Sigma}^+_{i,j} : i,j \in \{\diamond\} \cup \{1,\dots,k\}\}$ coincides with $\dot{\Sigma}_{\le k+1} + \varepsilon \mathbf{1}\mathbf{1}^\top$, and similarly $\{\widehat{\Sigma}^+_{i,j} : i,j \in \{\diamond\} \cup \{0,\dots,k\}\}$ coincides with $\widehat{\Sigma}_{\le k+1} + \varepsilon \mathbf{1}\mathbf{1}^\top$. By Proposition 2.5.4,

$$\frac{1}{N}\sum_{i=1}^N \delta(\dot{h}_i^{(1),1}, \dots, \dot{h}_i^{(1),k}, \dot{h}_i) \xrightarrow{\mathbb{W}_2} \mathcal{N}(0, \dot{\Sigma}^+_{\le k+1} + \varepsilon \mathbf{1}\mathbf{1}^\top),$$
$$\frac{1}{M}\sum_{a=1}^M \delta(\widehat{h}_a^{(1),0}, \dots, \widehat{h}_a^{(1),k}, \widehat{h}_a) \xrightarrow{\mathbb{W}_2} \mathcal{N}(0, \widehat{\Sigma}^+_{\le k+1} + \varepsilon \mathbf{1}\mathbf{1}^\top)$$

in probability under $\mathbb{P}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{Pl}}$. Proposition 2.5.5 implies the conclusion. ☐

2.5.3 Completion of the proof

We separately prove Proposition 2.4.8 under $\mathbb{P}$ and $\mathbb{P}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{Pl}}$.

Proof of Proposition 2.4.8(a)(b), under $\mathbb{P}$. By Proposition 2.5.2, for *any* k,

$$\mu_{\dot{\boldsymbol{h}}^k} \xrightarrow{\mathbb{W}_2} \mathcal{N}(0, \psi_\varepsilon + \varepsilon), \qquad \mu_{\widehat{\boldsymbol{h}}^k} \xrightarrow{\mathbb{W}_2} \mathcal{N}(0, q_\varepsilon + \varepsilon)$$

in probability. So, with high probability, $(\dot{\boldsymbol{h}}^k, \widehat{\boldsymbol{h}}^k) \in \mathcal{T}_{\varepsilon,\upsilon_0}$ and thus item (a) holds. Approximation arguments similar to the proof of Corollary 2.4.18 using Fact 2.4.20 yield

$$q(\boldsymbol{m}^k) \to \mathbb{E}[\mathrm{th}_\varepsilon((\psi_\varepsilon + \varepsilon)^{1/2} Z)^2] = q_\varepsilon$$

in probability. Regularity of ρ_ε and its derivatives then implies

$$\rho_\varepsilon(q(\boldsymbol{m}^k)) \to \varrho_\varepsilon, \qquad \rho'_\varepsilon(q(\boldsymbol{m}^k)) \to -1$$

in probability. Proposition 2.5.2 also implies

$$\lim_{k\to\infty} \operatorname*{p-lim}_{N\to\infty} \frac{1}{N}\|\dot{\boldsymbol{h}}^{k+1} - \dot{\boldsymbol{h}}^k\|^2 = \lim_{k\to\infty} \operatorname*{p-lim}_{N\to\infty} \frac{1}{N}\|\widehat{\boldsymbol{h}}^{k+1} - \widehat{\boldsymbol{h}}^k\|^2 = 0.$$

Below, let $o_{k,P}(\sqrt{N})$ denote a random vector $\boldsymbol{v}$ such that $\lim_{k\to\infty} \operatorname{p-lim}_{N\to\infty} \frac{1}{\sqrt{N}}\|\boldsymbol{v}\| = 0$, and $o_{k,P}(1)$ denote a random scalar ι such that $\lim_{k\to\infty} \operatorname{p-lim}_{N\to\infty} |\iota| = 0$. Let

$$\acute{\boldsymbol{h}}^k = \frac{\boldsymbol{G}\boldsymbol{m}^k}{\sqrt{N}} + \varepsilon^{1/2}\widehat{\boldsymbol{g}} - \rho_\varepsilon(q(\boldsymbol{m}^k))\boldsymbol{n}^k.$$

By Lemma 2.4.2,

$$\widehat{\boldsymbol{h}}^k = \frac{\boldsymbol{G}\boldsymbol{m}^k}{\sqrt{N}} + \varepsilon^{1/2}\widehat{\boldsymbol{g}} - \varrho_\varepsilon \boldsymbol{n}^{k-1}.$$

The above discussion implies $\widehat{\boldsymbol{h}}^k - \acute{\boldsymbol{h}}^k = o_{k,P}(\sqrt{N})$, and thus $\boldsymbol{n}^k - F_{\varepsilon,\rho_\varepsilon(q(\boldsymbol{m}))}(\acute{\boldsymbol{h}}^k) = o_{k,P}(\sqrt{N})$. By (2.35),

$$\nabla_{\boldsymbol{n}} \mathcal{F}^\varepsilon_{\mathsf{TAP}}(\boldsymbol{m}^k, \boldsymbol{n}^k) = o_{k,P}(\sqrt{N}).$$

Moreover,

$$d_\varepsilon(\boldsymbol{m}^k, \boldsymbol{n}^k) = \frac{1}{N}\sum_{a=1}^M F'_{\varepsilon,\varrho_\varepsilon}(\widehat{h}^k) + o_{k,P}(1) = d_\varepsilon + o_{k,P}(1),$$

for d_ε defined below Lemma 2.4.2. So

$$\nabla_{\boldsymbol{m}} \mathcal{F}^\varepsilon_{\mathsf{TAP}}(\boldsymbol{m}^k, \boldsymbol{n}^k) = -\mathrm{th}_\varepsilon^{-1}(\boldsymbol{m}^k) + \frac{\boldsymbol{G}^\top \boldsymbol{n}^k}{\sqrt{N}} + \varepsilon^{1/2}\dot{\boldsymbol{g}} - d_\varepsilon \boldsymbol{m}^k + \left(1 + \frac{\|\boldsymbol{G}\|_{\mathsf{op}}}{\sqrt{N}}\right) o_{k,P}(\sqrt{N}).$$

Since $\|\boldsymbol{G}\|_{\mathsf{op}} \le C\sqrt{N}$ w.h.p.,

$$\begin{aligned}\nabla_{\boldsymbol{m}} \mathcal{F}^\varepsilon_{\mathsf{TAP}}(\boldsymbol{m}^k, \boldsymbol{n}^k) &= -\dot{\boldsymbol{h}}^k + \frac{\boldsymbol{G}^\top \boldsymbol{n}^k}{\sqrt{N}} + \varepsilon^{1/2}\dot{\boldsymbol{g}} - d_\varepsilon \boldsymbol{m}^k + o_{k,P}(\sqrt{N}) \\ &= \dot{\boldsymbol{h}}^{k+1} - \dot{\boldsymbol{h}}^k + o_{k,P}(\sqrt{N}) = o_{k,P}(\sqrt{N}),\end{aligned}$$

proving item (b). ☐

Proof of Proposition 2.4.8(a)(b)(d), under $\mathbb{P}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{Pl}}$. Suppose first $(\boldsymbol{m},\boldsymbol{n}) \in \mathcal{S}_{\varepsilon,o_N(1)}$, and let $\dot{\boldsymbol{h}} = \mathrm{th}_\varepsilon^{-1}(\boldsymbol{m})$, $\widehat{\boldsymbol{h}} = F^{-1}_{\varepsilon,\varrho_\varepsilon}(\boldsymbol{n})$. The above argument, using Proposition 2.5.3 in place of Proposition 2.5.2, shows items (a) and (b) hold with high probability under $\mathbb{P}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{Pl}}$. Proposition 2.5.3 also yields

$$\lim_{k\to\infty} \operatorname*{p-lim}_{N\to\infty} \frac{1}{N}\|\dot{\boldsymbol{h}}^k - \dot{\boldsymbol{h}}\|^2 = \lim_{k\to\infty} \operatorname*{p-lim}_{N\to\infty} \frac{1}{N}\|\widehat{\boldsymbol{h}}^k - \widehat{\boldsymbol{h}}\|^2 = 0.$$

Thus item (d) holds with high probability under $\mathbb{P}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{Pl}}$. Finally, we show this remains true for $(\boldsymbol{m},\boldsymbol{n}) \in \mathcal{S}_{\varepsilon,\upsilon}$, for suitably small υ. Let $(\overline{\boldsymbol{m}},\overline{\boldsymbol{n}}) \in \mathcal{S}_{\varepsilon,o_N(1)}$ be such that $\frac{1}{N}\|\boldsymbol{m}-\overline{\boldsymbol{m}}\|^2, \frac{1}{N}\|\boldsymbol{n}-\overline{\boldsymbol{n}}\|^2 = o_\upsilon(1)$. We will show there is a coupling of $(\boldsymbol{G},\dot{\boldsymbol{g}},\widehat{\boldsymbol{g}}) \sim \mathbb{P}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{Pl}}$ and $(\overline{\boldsymbol{G}},\dot{\overline{\boldsymbol{g}}},\widehat{\overline{\boldsymbol{g}}}) \sim \mathbb{P}^{\overline{\boldsymbol{m}},\overline{\boldsymbol{n}}}_{\varepsilon,\mathsf{Pl}}$ such that

$$\|\boldsymbol{G}-\overline{\boldsymbol{G}}\|_{\mathsf{op}}, \|\dot{\boldsymbol{g}}-\dot{\overline{\boldsymbol{g}}}\|, \|\widehat{\boldsymbol{g}}-\widehat{\overline{\boldsymbol{g}}}\| \le o_\upsilon(1)\sqrt{N}. \tag{2.48}$$

If $(\boldsymbol{m}^k,\boldsymbol{n}^k)$ are the AMP iterates under $\mathbb{P}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{Pl}}$ and $(\overline{\boldsymbol{m}}^k,\overline{\boldsymbol{n}}^k)$ are the AMP iterates under $\mathbb{P}^{\overline{\boldsymbol{m}},\overline{\boldsymbol{n}}}_{\varepsilon,\mathsf{Pl}}$, this implies $\|\boldsymbol{m}^k-\overline{\boldsymbol{m}}^k\|, \|\boldsymbol{n}^k-\overline{\boldsymbol{n}}^k\| \le o_\upsilon(1)\sqrt{N}$ (this uses crucially that υ is set small depending on k). This implies (a) and (d) continue to hold, and similar approximation arguments to above show (b) continues to hold.

We now prove (2.48). Let $\dot{\overline{\boldsymbol{h}}} = \mathrm{th}_\varepsilon^{-1}(\overline{\boldsymbol{m}})$ and $\widehat{\overline{\boldsymbol{h}}} = F^{-1}_{\varepsilon,\rho_\varepsilon(q(\overline{\boldsymbol{m}}))}(\overline{\boldsymbol{n}})$. Another approximation argument shows $\|\dot{\boldsymbol{h}}-\dot{\overline{\boldsymbol{h}}}\|, \|\widehat{\boldsymbol{h}}-\widehat{\overline{\boldsymbol{h}}}\| \le o_\upsilon(1)\sqrt{N}$. The conditional means of $\boldsymbol{G},\overline{\boldsymbol{G}}$ are given by (2.38), and an approximation argument shows

$$\left\|\mathbb{E}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{Pl}}[\boldsymbol{G}] - \mathbb{E}^{\overline{\boldsymbol{m}},\overline{\boldsymbol{n}}}_{\varepsilon,\mathsf{Pl}}[\overline{\boldsymbol{G}}]\right\|_{\mathsf{op}} \le o_\upsilon(1)\sqrt{N}.$$

We couple the random parts $\widetilde{\boldsymbol{G}},\widetilde{\overline{\boldsymbol{G}}}$ as follows. Let $\dot{\boldsymbol{e}}_1,\widehat{\boldsymbol{e}}_1$ (resp. $\dot{\overline{\boldsymbol{h}}}_1,\widehat{\overline{\boldsymbol{h}}}_1$) be the the unit vectors parallel to $\boldsymbol{m},\boldsymbol{n}$ (resp. $\overline{\boldsymbol{m}},\overline{\boldsymbol{n}}$). Let $\dot{T},\hat{T}$ be rotation operators on $\mathbb{R}^N,\mathbb{R}^M$ with $\dot{T}\dot{\boldsymbol{e}}_1 = \dot{\overline{\boldsymbol{h}}}_1$ and $\hat{T}\widehat{\boldsymbol{e}}_1 = \widehat{\overline{\boldsymbol{h}}}_1$. These can be set so $\|\dot{T}-\boldsymbol{I}_N\|_{\mathsf{op}}, \|\hat{T}-\boldsymbol{I}_M\|_{\mathsf{op}} \le o_\upsilon(1)$. By (2.40), we can couple $\widetilde{\boldsymbol{G}},\widetilde{\overline{\boldsymbol{G}}}$ such that $\widetilde{\overline{\boldsymbol{G}}} = \hat{T}\widetilde{\boldsymbol{G}}\dot{T}^{-1}$. Since, for some absolute constant C, $\|\widetilde{\boldsymbol{G}}\|_{\mathsf{op}} \le C\sqrt{N}$ with high probability, on this event

$$\|\widetilde{\boldsymbol{G}}-\widetilde{\overline{\boldsymbol{G}}}\|_{\mathsf{op}} \le \|\widetilde{\boldsymbol{G}}\|_{\mathsf{op}}(\|\dot{T}-\boldsymbol{I}_N\|_{\mathsf{op}} + \|\hat{T}-\boldsymbol{I}_M\|_{\mathsf{op}}) = o_\upsilon(1)\sqrt{N}.$$

Thus $\|\boldsymbol{G}-\overline{\boldsymbol{G}}\|_{\mathsf{op}} \le o_\upsilon(1)\sqrt{N}$. The stationary equations (2.36), (2.37) then imply $\|\dot{\boldsymbol{g}}-\dot{\overline{\boldsymbol{g}}}\|_{\mathsf{op}}, \|\widehat{\boldsymbol{g}}-\widehat{\overline{\boldsymbol{g}}}\|_{\mathsf{op}} \le o_\upsilon(1)\sqrt{N}$. This proves (2.48). ☐

2.6 Local concavity of perturbed TAP free energy

In this section, we prove Lemmas 2.3.5 and 2.4.9 and Proposition 2.4.8(c).

2.6.1 Description of spectral gap bound

We first define a quantity λ_ε, which is a perturbed analog of the value $\lambda_0 = \inf_{z>-1} \lambda(z)$ defined in Condition 2.3.4. We will see that λ_ε upper bounds the maximum eigenvalue of $\nabla^2_\diamond \mathcal{F}^\varepsilon_{\mathsf{TAP}}$ near late AMP iterates. To define λ_ε, we introduce ε-perturbed variants of quantities appearing in Condition 2.3.4 and Lemma 2.3.5. Let

$$\dot{f}_\varepsilon(x) = \frac{\mathrm{ch}^2 x}{1+\varepsilon \mathrm{ch}^2(x)}, \qquad \widehat{f}_\varepsilon(x) = -\frac{F'_{\varepsilon,\varrho_\varepsilon}(x)}{1+\varrho_\varepsilon F'_{\varepsilon,\varrho_\varepsilon}(x)}.$$

We extend these definitions to $\varepsilon = 0$ by defining $\dot{f}_0(x) = \mathrm{ch}^2(x)$ and $\widehat{f}_0$ as in Condition 2.3.4; this extension will be used solely in Lemma 2.6.1 and the proof of Lemma 2.3.5 below.

Note that $\dot{f}_\varepsilon$ and $\widehat{f}_\varepsilon$ are positive, the latter because Fact 2.4.22 implies $F'_{\varepsilon,\varrho_\varepsilon}(x) < 0$ and $1+\varrho_\varepsilon F'_{\varepsilon,\varrho_\varepsilon}(x) > 0$, and $\dot{f}_\varepsilon(x)$ has minimum $\dot{f}_\varepsilon(0) = \frac{1}{1+\varepsilon}$. The function $\widehat{f}_0$ is also positive, as Lemma 2.4.21(b) implies $F'_{1-q_0}(x) < 0$ and $1+(1-q_0)F'_{1-q_0}(x) > 0$. In the below, it will be convenient to abbreviate $\widetilde{q}_\varepsilon = q_\varepsilon + \varepsilon$, $\widetilde{\psi}_\varepsilon = \psi_\varepsilon + \varepsilon$.

Lemma 2.6.1. *For any $\varepsilon \ge 0$ (including $\varepsilon = 0$), the functions $m_\varepsilon, \theta_\varepsilon : (-\frac{1}{1+\varepsilon}, +\infty) \to (0,+\infty)$ defined by*

$$\begin{aligned} m_\varepsilon(z) &= \mathbb{E}[(z+\dot{f}_\varepsilon(\widetilde{\psi}_\varepsilon^{1/2} Z))^{-1}], \\ \theta_\varepsilon(z) &= \mathbb{E}[(z+\dot{f}_\varepsilon(\widetilde{\psi}_\varepsilon^{1/2} Z))^{-2}]\, \mathbb{E}\left[\left(\frac{\widehat{f}_\varepsilon(\widetilde{q}_\varepsilon^{1/2} Z)}{1+m_\varepsilon(z)\widehat{f}_\varepsilon(\widetilde{q}_\varepsilon^{1/2} Z)}\right)^2\right] \end{aligned}$$

are continuous and strictly decreasing, with

$$\lim_{z\downarrow -(1+\varepsilon)^{-1}} m_\varepsilon(z) = \lim_{z\downarrow -(1+\varepsilon)^{-1}} \theta_\varepsilon(z) = +\infty, \qquad \lim_{z\uparrow +\infty} m_\varepsilon(z) = \lim_{z\uparrow +\infty} \theta_\varepsilon(z) = 0.$$

In particular θ_ε has a well-defined inverse $\theta_\varepsilon^{-1} : (0,+\infty) \to (-\frac{1}{1+\varepsilon}, +\infty)$.

Proof of Lemma 2.6.1. Note that $m_\varepsilon(z)$ is clearly decreasing on $(-\frac{1}{1+\varepsilon}, +\infty)$ with $\lim_{z\uparrow+\infty} m_\varepsilon(z) = 0$. To show the other limit, let

$$\dot{g}_\varepsilon(x) = \dot{f}_\varepsilon(x) - \frac{1}{1+\varepsilon} = \frac{\mathrm{sh}^2(x)}{(1+\varepsilon)(1+\varepsilon\mathrm{ch}^2(x))}.$$

For $z = -\frac{1}{1+\varepsilon} + \iota$, with $\iota > 0$ small,

$$m_\varepsilon(z) = \mathbb{E}[(\iota + \dot{g}_\varepsilon(\widetilde{\psi}_\varepsilon^{1/2} Z))^{-1}] \ge \mathbb{E}[\mathbf{1}\{|Z| \le \iota^{1/2}\}(\iota + \dot{g}_\varepsilon(\widetilde{\psi}_\varepsilon^{1/2} Z))^{-1}] \ge \Omega(\iota^{-1/2}).$$

Thus $\lim_{z\downarrow -(1+\varepsilon)^{-1}} m_\varepsilon(z) = +\infty$. We can write $\theta_\varepsilon(z)$ as

$$\theta_\varepsilon(z) = \frac{\mathbb{E}[(z+\dot{f}_\varepsilon(\widetilde{\psi}_\varepsilon^{1/2} Z))^{-2}]}{\mathbb{E}[(z+\dot{f}_\varepsilon(\widetilde{\psi}_\varepsilon^{1/2} Z))^{-1}]^2}\, \mathbb{E}\left[\frac{(m_\varepsilon(z)\widehat{f}_\varepsilon(\widetilde{q}_\varepsilon^{1/2} Z))^2}{(1+m_\varepsilon(z)\widehat{f}_\varepsilon \widetilde{q}_\varepsilon^{1/2} Z))^2}\right]. \tag{2.49}$$

Since $m_\varepsilon(z)$ is decreasing and $\widehat{f}_\varepsilon$ is positive, the second factor of (2.49) is manifestly decreasing. The z-derivative of the first is

$$\frac{-\mathbb{E}[(z+\dot{f}_\varepsilon(\widetilde{\psi}_\varepsilon^{1/2} Z))^{-1}]\,\mathbb{E}[(z+\dot{f}_\varepsilon(\widetilde{\psi}_\varepsilon^{1/2} Z))^{-3}] + \mathbb{E}[(z+\dot{f}_\varepsilon(\widetilde{\psi}_\varepsilon^{1/2} Z))^{-2}]^2}{\mathbb{E}[(z+\dot{f}_\varepsilon(\widetilde{\psi}_\varepsilon^{1/2} Z))^{-1}]^3} < 0$$

by Cauchy–Schwarz. Thus θ_ε is decreasing on $(-\frac{1}{1+\varepsilon}, +\infty)$. We now calculate its limits as $z \downarrow -\frac{1}{1+\varepsilon}$ and $z \uparrow +\infty$. Consider first $z = -\frac{1}{1+\varepsilon} + \iota$ for ι small. Then the first factor of (2.49) is

$$\frac{\mathbb{E}[(\iota + \dot{g}_\varepsilon(\widetilde{\psi}_\varepsilon^{1/2} Z))^{-2}]}{\mathbb{E}[(\iota + \dot{g}_\varepsilon(\widetilde{\psi}_\varepsilon^{1/2} Z))^{-1}]^2} \ge \frac{\mathbb{E}[\mathbf{1}\{|Z| \le \iota^{1/2}\}(\iota + \dot{g}_\varepsilon(\widetilde{\psi}_\varepsilon^{1/2} Z))^{-2}]}{\mathbb{E}[\mathbf{1}\{|Z| \le \iota^{1/3}\}(\iota + \dot{g}_\varepsilon(\widetilde{\psi}_\varepsilon^{1/2} Z))^{-1} + O(\iota^{-2/3})]^2} = \frac{\Omega(\iota^{-3/2})}{O(\iota^{-4/3})},$$

which diverges as $\iota \downarrow 0$. The second factor of (2.49) tends to 1 in this limit by dominated convergence. Thus $\lim_{z \downarrow -(1+\varepsilon)^{-1}} \theta_\varepsilon(z) = +\infty$. We can write the first factor of (2.49) as

$$\frac{\mathbb{E}[(1 + z^{-1} \dot{f}_\varepsilon(\widetilde{\psi}_\varepsilon^{1/2} Z))^{-2}]}{\mathbb{E}[(1 + z^{-1} \dot{f}_\varepsilon(\widetilde{\psi}_\varepsilon^{1/2} Z))^{-1}]^2},$$

which tends to 1 as $z \uparrow +\infty$ by dominated convergence. In this limit, the second factor of (2.49) tends to 0 by dominated convergence, so $\lim_{z \uparrow +\infty} \theta_\varepsilon(z) = 0$. This completes the proof. $\square$

Proof of Lemma 2.3.5. Note that

$$m'(z) = -\mathbb{E}[(z + \mathrm{ch}^2(\psi_0^{1/2} Z))^{-2}].$$

Thus, differentiating λ yields

$$\lambda'(z) = 1 + \alpha_\star m'(z)\, \mathbb{E}\left[\left(\frac{\widehat{f}_0(q_0^{1/2} Z)}{1 + m(z)\widehat{f}_0(q_0^{1/2} Z)}\right)^2\right] = 1 - \alpha_\star \theta(z).$$

The assertions about θ follow from Lemma 2.6.1, with $\varepsilon = 0$. Since θ is strictly decreasing on $(-1, +\infty)$, λ' is strictly increasing on this interval, and therefore λ is strictly convex on this interval. Since $\theta^{-1} : (0, +\infty) \to (-1, +\infty)$ is well-defined, we may define $z_0 = \theta^{-1}(\alpha_\star^{-1})$. This point satisfies the stationarity condition $\lambda'(z_0) = 0$ and is thus the unique minimizer of λ on $(-1, +\infty)$. $\square$

Recall from below Lemma 2.4.2 that $d_\varepsilon = \alpha_\star\, \mathbb{E}[F'_{\varepsilon,\varrho_\varepsilon}(\widetilde{q}_\varepsilon^{1/2} Z)]$. We now define the threshold λ_ε.

Definition 2.6.2. Let $z_\varepsilon = \theta_\varepsilon^{-1}(\alpha_\star^{-1})$ and

$$\lambda_\varepsilon \equiv z_\varepsilon - \alpha_\star\, \mathbb{E}\left[\frac{\widehat{f}_\varepsilon(\widetilde{q}_\varepsilon^{1/2} Z)}{1 + m_\varepsilon(z_\varepsilon)\widehat{f}_\varepsilon(\widetilde{q}_\varepsilon^{1/2} Z)}\right] - d_\varepsilon. \tag{2.50}$$

Lemma 2.6.3. *As $\varepsilon \downarrow 0$, $\lambda_\varepsilon \to \lambda_0$ (defined in Condition 2.3.4).*

Proof. By Proposition 2.4.1, as $\varepsilon \downarrow 0$, $(\widetilde{q}_\varepsilon, \widetilde{\psi}_\varepsilon) \to (q_0, \psi_0)$. Thus, for $\dot{f}_0(x) = \mathrm{ch}^2(x)$, the push-forwards $(\dot{f}_\varepsilon)_\# \mathcal{N}(0, \widetilde{\psi}_\varepsilon)$ and $(\widehat{f}_\varepsilon)_\# \mathcal{N}(0, \widetilde{q}_\varepsilon)$ converge weakly to $(\dot{f}_0)_\# \mathcal{N}(0, \psi_0)$ and $(\widehat{f}_0)_\# \mathcal{N}(0, q_0)$.

For any $z > -1$ and small ε, the integrand of $m_\varepsilon(z)$ is bounded independently of ε, and thus $\lim_{\varepsilon \downarrow 0} m_\varepsilon(z) = m(z)$ by dominated convergence. Similarly, all three integrands in (2.49) are bounded, so $\lim_{\varepsilon \downarrow 0} \theta_\varepsilon(z) = \theta(z)$. Moreover, one easily checks that on any compact subset of $(-1, +\infty)$, the derivatives of $m_\varepsilon, \theta_\varepsilon$ are bounded independently of ε. Thus $m_\varepsilon \to m$, $\theta_\varepsilon \to \theta$ uniformly on compact subsets of $(-1, +\infty)$.

By Lemma 2.3.5, $\lim_{z \downarrow -1} \theta(z) = +\infty$, so $z_0 = \theta^{-1}(\alpha_\star^{-1})$ is bounded away from -1. The above uniform convergence then implies $z_\varepsilon \to z_0$ and $m_\varepsilon(z_\varepsilon) \to m(z_0)$. Since the below integrands are bounded,

$$\mathbb{E}\left[\frac{\widehat{f}_\varepsilon(\widetilde{q}_\varepsilon^{1/2} Z)}{1 + m_\varepsilon(z_\varepsilon)\widehat{f}_\varepsilon(\widetilde{q}_\varepsilon^{1/2} Z)}\right] \to \mathbb{E}\left[\frac{\widehat{f}_0(q_0^{1/2} Z)}{1 + m(z_0)\widehat{f}_0(q_0^{1/2} Z)}\right].$$

Finally, as $F'_{\varepsilon,\varrho_\varepsilon}$ is bounded (by Fact 2.4.22) and limits to the bounded function F'_{1-q_0}, we have $d_\varepsilon \to d_0$. $\square$

2.6.2 Hessian estimate

We next prove the following upper bound on $\nabla^2_\diamond \mathcal{F}^\varepsilon_{\mathsf{TAP}}$.

Lemma 2.6.4. *Suppose $(\boldsymbol{m},\boldsymbol{n}) \in \mathcal{S}_{\varepsilon,r_0}$, and $\|\boldsymbol{G}\|_{\mathsf{op}}, \|\widehat{\boldsymbol{g}}\| \le C\sqrt{N}$ for some absolute constant C (i.e. independent of all parameters in Subsection 2.4.1). Let $\dot{\boldsymbol{h}} \in \mathbb{R}^N$, $\acute{\boldsymbol{h}} \in \mathbb{R}^M$ be defined (as in Lemma 2.4.16) by*

$$\dot{\boldsymbol{h}} = \mathrm{th}_\varepsilon^{-1}(\boldsymbol{m}), \qquad \acute{\boldsymbol{h}} = \frac{\boldsymbol{G}\boldsymbol{m}}{\sqrt{N}} + \varepsilon^{1/2}\widehat{\boldsymbol{g}} - \rho_\varepsilon(q(\boldsymbol{m}))\boldsymbol{n},$$

and

$$\boldsymbol{D}_1 = \mathrm{diag}(\dot{f}_\varepsilon(\dot{\boldsymbol{h}})), \qquad \boldsymbol{D}_2 = \mathrm{diag}(\widehat{f}_\varepsilon(\acute{\boldsymbol{h}})).$$

Then,

$$\nabla^2_\diamond \mathcal{F}^\varepsilon_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n}) \preceq P^\perp_{\boldsymbol{m}}\left(-\boldsymbol{D}_1 - \frac{1}{N}\boldsymbol{G}^\top \boldsymbol{D}_2 \boldsymbol{G} - d_\varepsilon \boldsymbol{I}_N\right)P^\perp_{\boldsymbol{m}} + \frac{\lambda_\varepsilon \boldsymbol{m}\boldsymbol{m}^\top}{\|\boldsymbol{m}\|^2} + (o_{C_{\mathsf{cvx}}}(1) + o_{r_0}(1))\boldsymbol{I}_N.$$

(Recall the meaning of $o_{C_{\mathsf{cvx}}}(1), o_{r_0}(1)$ discussed in Subsection 2.4.1.)

Fact 2.6.5 (Proved in Appendix 2.A)**.** *Let $\boldsymbol{m} \in \mathbb{R}^N$, $\boldsymbol{n} \in \mathbb{R}^M$, and let $\acute{\boldsymbol{h}}, \dot{\boldsymbol{h}}$ be as above. Let $F = F_{\varepsilon,\rho_\varepsilon(q(\boldsymbol{m}))}$ and*

$$\boldsymbol{D}_3 = \mathrm{diag}\left(F'(\acute{\boldsymbol{h}})\right), \qquad \boldsymbol{D}_4 = \boldsymbol{I}_M + \rho_\varepsilon(q(\boldsymbol{m}))\boldsymbol{D}_3.$$

Then,

$$\begin{aligned}
\nabla^2_{\boldsymbol{m},\boldsymbol{m}} \mathcal{F}^\varepsilon_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n}) &= -\boldsymbol{D}_1 + \frac{\boldsymbol{G}^\top \boldsymbol{D}_3 \boldsymbol{G}}{N} + \rho'_\varepsilon(q(\boldsymbol{m}))d_\varepsilon(\boldsymbol{m},\boldsymbol{n})\boldsymbol{I}_N \\
&\quad + \rho'_\varepsilon(q(\boldsymbol{m})) \cdot \frac{\boldsymbol{G}^\top (F''(\acute{\boldsymbol{h}}) + 2\boldsymbol{D}_3(F(\acute{\boldsymbol{h}}) - \boldsymbol{n}))\boldsymbol{m}^\top + \boldsymbol{m}(F''(\acute{\boldsymbol{h}}) + 2\boldsymbol{D}_3(F(\acute{\boldsymbol{h}}) - \boldsymbol{n}))^\top \boldsymbol{G}}{N^{3/2}} \\
&\quad + \left\{\rho''_\varepsilon(q(\boldsymbol{m}))d_\varepsilon(\boldsymbol{m},\boldsymbol{n}) + \frac{\rho'_\varepsilon(q(\boldsymbol{m}))^2}{N}\sum_{a=1}^M \left(2F'(\acute{h}_a)^2 + F^{(3)}(\acute{h}_a)\right)\right\}\frac{\boldsymbol{m}\boldsymbol{m}^\top}{N} \\
\nabla^2_{\boldsymbol{m},\boldsymbol{n}} \mathcal{F}^\varepsilon_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n}) &= -\frac{\rho_\varepsilon(q(\boldsymbol{m}))}{\sqrt{N}}\boldsymbol{G}^\top \boldsymbol{D}_3 - \rho'_\varepsilon(q(\boldsymbol{m}))\frac{\boldsymbol{m}(\rho_\varepsilon(q(\boldsymbol{m}))F''(\acute{\boldsymbol{h}}) + 2\boldsymbol{D}_4(F(\acute{\boldsymbol{h}}) - \boldsymbol{n}))^\top}{N} \\
\nabla^2_{\boldsymbol{n},\boldsymbol{n}} \mathcal{F}^\varepsilon_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n}) &= \rho_\varepsilon(q(\boldsymbol{m}))\boldsymbol{D}_4,
\end{aligned}$$

Furthermore, for

$$\widetilde{\boldsymbol{D}}_2 = -\boldsymbol{D}_3 + \rho_\varepsilon(q(\boldsymbol{m}))\boldsymbol{D}_3^2\boldsymbol{D}_4^{-1} = \mathrm{diag}\left(-\frac{F'(\acute{\boldsymbol{h}})}{1 + \rho_\varepsilon(q(\boldsymbol{m}))F'(\acute{\boldsymbol{h}})}\right),$$

we have

$$\begin{aligned}
\nabla^2_\diamond \mathcal{F}^\varepsilon_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n}) &= -\boldsymbol{D}_1 - \frac{\boldsymbol{G}^\top \widetilde{\boldsymbol{D}}_2 \boldsymbol{G}}{N} + \rho'_\varepsilon(q(\boldsymbol{m}))d_\varepsilon(\boldsymbol{m},\boldsymbol{n})\boldsymbol{I}_N \\
&\quad + \rho'_\varepsilon(q(\boldsymbol{m})) \cdot \frac{\boldsymbol{G}^\top \boldsymbol{D}_4^{-1} F''(\acute{\boldsymbol{h}})\boldsymbol{m}^\top + \boldsymbol{m}F''(\acute{\boldsymbol{h}})^\top \boldsymbol{D}_4^{-1}\boldsymbol{G}}{N^{3/2}} \\
&\quad + \left\{\rho''_\varepsilon(q(\boldsymbol{m}))d_\varepsilon(\boldsymbol{m},\boldsymbol{n}) + \frac{\rho'_\varepsilon(q(\boldsymbol{m}))^2}{N}\sum_{a=1}^M \left(2F'(\acute{h}_a)^2 + F^{(3)}(\acute{h}_a)\right.\right. \\
&\qquad \left.\left. - \frac{(\rho_\varepsilon(q(\boldsymbol{m}))F''(\acute{h}_a) + 2(F(\acute{h}_a) - n_a)(1 + \rho_\varepsilon(q(\boldsymbol{m}))F'(\acute{h}_a)))^2}{\rho_\varepsilon(q(\boldsymbol{m}))(1 + \rho_\varepsilon(q(\boldsymbol{m}))F'(\acute{h}_a))}\right)\right\}\frac{\boldsymbol{m}\boldsymbol{m}^\top}{N}.
\end{aligned}$$

Lemma 2.6.6 (Proved in Appendix 2.A)**.** *Suppose* $(\boldsymbol{m},\boldsymbol{n}) \in \mathcal{S}_{\varepsilon,r_0}$ *and* $\|\boldsymbol{G}\|_{\mathsf{op}}, \|\widehat{\boldsymbol{g}}\| \le C\sqrt{N}$ *for an absolute constant* C*. The following estimates hold for sufficiently small* r_0 *(depending on* $\varepsilon, C_{\mathsf{cvx}}, C_{\mathsf{bd}}, \eta$*).*

(a) Up to additive $o_{r_0}(1)$ *error,* $q(\boldsymbol{m}) \approx q_\varepsilon$*,* $\psi(\boldsymbol{n}) \approx \psi_\varepsilon$*, and* $d_\varepsilon(\boldsymbol{m},\boldsymbol{n}) \approx d_\varepsilon$*,* $\rho_\varepsilon(q(\boldsymbol{m})) \approx \varrho_\varepsilon$*,* $\rho'_\varepsilon(q(\boldsymbol{m})) \approx -1$*,* $\rho''_\varepsilon(q(\boldsymbol{m})) \approx C_{\mathsf{cvx}}$*.*

(b) $\|\widetilde{\boldsymbol{D}}_2 - \boldsymbol{D}_2\|_{\mathsf{op}} = o_{r_0}(1)$*.*

(c) $\frac{1}{N}\sum_{a=1}^M (2F'(\acute{h}_a)^2 + F^{(3)}(\acute{h}_a))$ *is bounded by an absolute constant.*

(d) $\frac{1}{\sqrt{N}}\|\boldsymbol{D}_4^{-1}F''(\acute{\boldsymbol{h}})\|$ *is bounded, with bound depending only on* ε*.*

Proof of Lemma 2.6.4. By Fact 2.6.5 and Lemma 2.6.6,

$$\nabla^2_\diamond \mathcal{F}^\varepsilon_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n}) \preceq -\boldsymbol{D}_1 - \frac{\boldsymbol{G}^\top \widetilde{\boldsymbol{D}}_2 \boldsymbol{G}}{N} - d_\varepsilon \boldsymbol{I}_N + \frac{\boldsymbol{G}^\top \boldsymbol{v}_1 \boldsymbol{m}^\top + \boldsymbol{m}\boldsymbol{v}_1^\top \boldsymbol{G}}{N} + (C_{\mathsf{cvx}} d_\varepsilon + C_1)\frac{\boldsymbol{m}\boldsymbol{m}^\top}{N} + o_{r_0}(1)\boldsymbol{I}_N,$$

for $C_1 \in \mathbb{R}$, $\boldsymbol{v}_1 \in \mathbb{R}^N$ with $|C_1|, \|\boldsymbol{v}_1\|$ bounded depending only on ε. By the assumption on $\|\boldsymbol{G}\|_{\mathsf{op}}$, $\frac{1}{\sqrt{N}}\|\boldsymbol{G}^\top \boldsymbol{v}_1\|$ is also bounded depending only on ε. Note that

$$-\boldsymbol{D}_1 \preceq -P^\perp_{\boldsymbol{m}} \boldsymbol{D}_1 P^\perp_{\boldsymbol{m}} - (P^\perp_{\boldsymbol{m}} \boldsymbol{D}_1 P_{\boldsymbol{m}} + P_{\boldsymbol{m}} \boldsymbol{D}_1 P^\perp_{\boldsymbol{m}}) = -P^\perp_{\boldsymbol{m}} \boldsymbol{D}_1 P^\perp_{\boldsymbol{m}} - \frac{(P^\perp_{\boldsymbol{m}} \boldsymbol{D}_1 \boldsymbol{m})\boldsymbol{m}^\top + \boldsymbol{m}(P^\perp_{\boldsymbol{m}} \boldsymbol{D}_1 \boldsymbol{m})}{q(\boldsymbol{m})N}$$

and similarly

$$-\frac{1}{N}\boldsymbol{G}^\top \boldsymbol{D}_2 \boldsymbol{G} \preceq -P^\perp_{\boldsymbol{m}} \boldsymbol{G}^\top \boldsymbol{D}_2 \boldsymbol{G} P^\perp_{\boldsymbol{m}} - \frac{(P^\perp_{\boldsymbol{m}} \boldsymbol{G}^\top \boldsymbol{D}_2 \boldsymbol{G}\boldsymbol{m})\boldsymbol{m}^\top + \boldsymbol{m}(P^\perp_{\boldsymbol{m}} \boldsymbol{G}^\top \boldsymbol{D}_2 \boldsymbol{G}\boldsymbol{m})^\top}{q(\boldsymbol{m})N^2}.$$

Moreover $\|\boldsymbol{D}_1\|_{\mathsf{op}}, \|\boldsymbol{D}_2\|_{\mathsf{op}} \le O(\varepsilon^{-1})$, the latter by (2.42). So, there exists $C_2 \in \mathbb{R}$, $\boldsymbol{v}_2 \in \mathbb{R}^N$ with $|C_2|, \|\boldsymbol{v}_2\|$ bounded depending only on ε, such that

$$\nabla^2_\diamond \mathcal{F}^\varepsilon_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n}) \preceq P^\perp_{\boldsymbol{m}} \left(-\boldsymbol{D}_1 - \frac{\boldsymbol{G}^\top \widetilde{\boldsymbol{D}}_2 \boldsymbol{G}}{N}\right) P^\perp_{\boldsymbol{m}} - d_\varepsilon \boldsymbol{I}_N + \frac{\boldsymbol{v}_2 \boldsymbol{m}^\top + \boldsymbol{m}\boldsymbol{v}_2^\top}{N^{1/2}} + (C_{\mathsf{cvx}} d_\varepsilon + C_2)\frac{\boldsymbol{m}\boldsymbol{m}^\top}{N} + o_{r_0}(1)\boldsymbol{I}_N.$$

Note that $d_\varepsilon < 0$, because $F'_{\varepsilon,\varrho_\varepsilon} < 0$ by Fact 2.4.22. So, for large C_{cvx},

$$(C_{\mathsf{cvx}} d_\varepsilon + C_2)\frac{\boldsymbol{m}\boldsymbol{m}^\top}{N} + \frac{\boldsymbol{v}_2 \boldsymbol{m}^\top + \boldsymbol{m}\boldsymbol{v}_2^\top}{N^{1/2}} \preceq \frac{(\lambda_\varepsilon + d_\varepsilon)\boldsymbol{m}\boldsymbol{m}^\top}{\|\boldsymbol{m}\|^2} + \frac{\boldsymbol{v}_2 \boldsymbol{v}_2^\top}{C_{\mathsf{cvx}}|d_\varepsilon| - C_2 + (\lambda_\varepsilon + d_\varepsilon)/q(\boldsymbol{m})}.$$

The final term has operator norm $o_{C_{\mathsf{cvx}}}(1)$. $\square$

2.6.3 Null model: post-AMP Gordon's inequality

We turn to the proof of Proposition 2.4.8(c), first under the measure $\mathbb{P}$. In light of Lemma 2.6.4, we define

$$\boldsymbol{R}(\boldsymbol{m},\boldsymbol{n}) = P^\perp_{\boldsymbol{m}} \left(-\boldsymbol{D}_1 - \frac{1}{N}\boldsymbol{G}^\top \boldsymbol{D}_2 \boldsymbol{G}\right) P^\perp_{\boldsymbol{m}}, \tag{2.51}$$

where, as in that lemma, $\boldsymbol{D}_1 = \operatorname{diag}(\dot{f}_\varepsilon(\dot{\boldsymbol{h}}))$, $\boldsymbol{D}_2 = \operatorname{diag}(\widehat{f}_\varepsilon(\acute{\boldsymbol{h}}(\boldsymbol{m},\boldsymbol{n},\boldsymbol{G})))$ for $\dot{\boldsymbol{h}} = \operatorname{th}_\varepsilon^{-1}(\boldsymbol{m})$ and

$$\acute{\boldsymbol{h}}(\boldsymbol{m},\boldsymbol{n},\boldsymbol{G}) = \frac{\boldsymbol{G}\boldsymbol{m}}{\sqrt{N}} + \varepsilon^{1/2}\widehat{\boldsymbol{g}} - \rho_\varepsilon(q(\boldsymbol{m}))\boldsymbol{n}.$$

Proposition 2.6.7. *With high probability under* $\mathbb{P}$*,* $\boldsymbol{R}(\boldsymbol{m},\boldsymbol{n}) \preceq (\lambda_\varepsilon + d_\varepsilon + o_{r_0}(1) + o_k(1))P^\perp_{\boldsymbol{m}}$ *for all* $\|(\boldsymbol{m},\boldsymbol{n}) - (\boldsymbol{m}^k,\boldsymbol{n}^k)\| \le r_0\sqrt{N}$*.*

For z_ε defined in Definition 2.6.2, let

$$r_\varepsilon^2 = \mathbb{E}[(z_\varepsilon + \dot{f}_\varepsilon(\widetilde{\psi}_\varepsilon^{1/2} Z))^{-2}]^{-1}.$$

Define the AMP iterates $\boldsymbol{m}^0, \boldsymbol{n}^0, \ldots, \boldsymbol{m}^k, \boldsymbol{n}^k$ and $\widehat{\boldsymbol{h}}^0, \dot{\boldsymbol{h}}^1, \widehat{\boldsymbol{h}}^1, \ldots, \dot{\boldsymbol{h}}^k, \widehat{\boldsymbol{h}}^k$ as in (2.20), (2.21), and

$$\mathsf{DATA} = (\dot{\boldsymbol{g}}, \dot{\boldsymbol{h}}^1, \ldots, \dot{\boldsymbol{h}}^k, \widehat{\boldsymbol{g}}, \widehat{\boldsymbol{h}}^0, \ldots, \widehat{\boldsymbol{h}}^k).$$

Let $U(r_0) = \{(\boldsymbol{m}, \boldsymbol{n}) : \|(\boldsymbol{m}, \boldsymbol{n}) - (\boldsymbol{m}^k, \boldsymbol{n}^k)\| \le r_0 \sqrt{N}\}$. Let $\acute{\boldsymbol{h}}^k \equiv \acute{\boldsymbol{h}}(\boldsymbol{m}^k, \boldsymbol{n}^k, \boldsymbol{G})$, and note that

$$\acute{\boldsymbol{h}}^k = \widehat{\boldsymbol{h}}^k + \varrho_\varepsilon \boldsymbol{n}^{k-1} - \rho_\varepsilon(q(\boldsymbol{m}^k))\boldsymbol{n}^k \tag{2.52}$$

is DATA-measurable. Let $U'(r_0) = \{\acute{\boldsymbol{h}} : \|\acute{\boldsymbol{h}} - \acute{\boldsymbol{h}}^k\| \le C r_0 \sqrt{N}\}$, for a suitably large absolute constant C. Since $\|\boldsymbol{G}\|_{\mathsf{op}} = O(\sqrt{N})$ with high probability, on this event $\acute{\boldsymbol{h}}(\boldsymbol{m}, \boldsymbol{n}, \boldsymbol{G}) \in U'(r_0)$ for all $(\boldsymbol{m}, \boldsymbol{n}) \in U(r_0)$.

Below, we will write $\boldsymbol{D}_2(\acute{\boldsymbol{h}}) = \operatorname{diag}(\widehat{f}_\varepsilon(\acute{\boldsymbol{h}}))$ for a varying $\acute{\boldsymbol{h}}$ which is not necessarily $\acute{\boldsymbol{h}}(\boldsymbol{m}, \boldsymbol{n}, \boldsymbol{G})$. On the other hand $\boldsymbol{D}_1$ always refers to the function of $\boldsymbol{m}$ defined above. The starting point of our proof of Proposition 2.6.7 is to recast the maximum eigenvalue as a minimax program, as follows:

$$\begin{aligned}
&\sup_{(\boldsymbol{m},\boldsymbol{n}) \in U(r_0)} \sup_{\substack{\|\dot{\boldsymbol{v}}\|=1 \\ \dot{\boldsymbol{v}} \perp \boldsymbol{m}}} \dot{\boldsymbol{v}}^\top \left(-\boldsymbol{D}_1 - \frac{1}{N} \boldsymbol{G}^\top \boldsymbol{D}_2(\acute{\boldsymbol{h}}(\boldsymbol{m}, \boldsymbol{n}, \boldsymbol{G})) \boldsymbol{G} \right) \dot{\boldsymbol{v}} \\
&= \sup_{(\boldsymbol{m},\boldsymbol{n}) \in U(r_0)} \sup_{\substack{\|\dot{\boldsymbol{v}}\|=1 \\ \dot{\boldsymbol{v}} \perp \boldsymbol{m}}} \inf_{\widehat{\boldsymbol{v}} \in \mathbb{R}^M} \left\{ -\langle \boldsymbol{D}_1 \dot{\boldsymbol{v}}, \dot{\boldsymbol{v}} \rangle + \langle \boldsymbol{D}_2(\acute{\boldsymbol{h}}(\boldsymbol{m}, \boldsymbol{n}, \boldsymbol{G}))^{-1} \widehat{\boldsymbol{v}}, \widehat{\boldsymbol{v}} \rangle + \frac{2}{\sqrt{N}} \langle \boldsymbol{G} \dot{\boldsymbol{v}}, \widehat{\boldsymbol{v}} \rangle \right\}.
\end{aligned}$$

Here we used that $\boldsymbol{D}_1, \boldsymbol{D}_2$ are positive definite, by positivity of $\dot{f}_\varepsilon$, $\widehat{f}_\varepsilon$. On the high probability event that $\|\boldsymbol{G}\|_{\mathsf{op}} = O(\sqrt{N})$, this is bounded by

$$\sup_{\substack{(\boldsymbol{m},\boldsymbol{n}) \in U(r_0) \\ \acute{\boldsymbol{h}} \in U'(r_0)}} \sup_{\substack{\|\dot{\boldsymbol{v}}\|=1 \\ \dot{\boldsymbol{v}} \perp \boldsymbol{m}}} \inf_{\substack{\|\widehat{\boldsymbol{v}}\|=r_\varepsilon \\ \widehat{\boldsymbol{v}} \perp \boldsymbol{n}}} \left\{ -\langle \boldsymbol{D}_1 \dot{\boldsymbol{v}}, \dot{\boldsymbol{v}} \rangle + \langle \boldsymbol{D}_2(\acute{\boldsymbol{h}})^{-1} \widehat{\boldsymbol{v}}, \widehat{\boldsymbol{v}} \rangle + \frac{2}{\sqrt{N}} \langle \boldsymbol{G} \dot{\boldsymbol{v}}, \widehat{\boldsymbol{v}} \rangle \right\}. \tag{2.53}$$

We will control (2.53) by applying Gordon's minimax inequality conditional on the AMP iterates; we explain this next. Let

$$\dot{\mu}_{\mathrm{AMP}} = \frac{1}{N} \sum_{i=1}^N \delta(\varepsilon^{1/2} \dot{g}, \dot{h}_i^1, \ldots, \dot{h}_i^k), \qquad \widehat{\mu}_{\mathrm{AMP}} = \frac{1}{M} \sum_{a=1}^M \delta(\varepsilon^{1/2} \widehat{g}, \widehat{h}_a^0, \ldots, \widehat{h}_a^k).$$

Further let $(\dot{\Sigma}^+_{i,j})_{i,j \ge 0}$ and $(\widehat{\Sigma}^+_{i,j})_{i,j \ge -1}$ be augmented versions of $(\dot{\Sigma}_{i,j})_{i,j \ge 1}, (\widehat{\Sigma}_{i,j})_{i,j \ge 0}$ where we add a row and column of zeros, i.e. $\dot{\Sigma}^+_{0,i} = \dot{\Sigma}^+_{i,0} = \widehat{\Sigma}^+_{-1,i} = \widehat{\Sigma}^+_{i,-1} = 0$.

Lemma 2.6.8. *For any $\upsilon > 0$, with high probability,*

$$\mathbb{W}_2(\dot{\mu}_{\mathrm{AMP}}, \mathcal{N}(0, \dot{\Sigma}^+_{\le k} + \varepsilon \mathbf{1}\mathbf{1}^\top)), \mathbb{W}_2(\widehat{\mu}_{\mathrm{AMP}}, \mathcal{N}(0, \widehat{\Sigma}^+_{\le k} + \varepsilon \mathbf{1}\mathbf{1}^\top)) \le \upsilon. \tag{2.54}$$

Proof. Follows from AMP state evolution, identically to Proposition 2.5.2. □

We now let υ be sufficiently small depending on r_0, k and condition on a realization of DATA such that (2.54) holds. (Note that (2.54) is DATA-measurable.) Define $\bar{\boldsymbol{h}}^i = \dot{\boldsymbol{h}}^i - \varepsilon^{1/2} \dot{\boldsymbol{g}}$, $\breve{\boldsymbol{h}}^i = \widehat{\boldsymbol{h}}^i - \varepsilon^{1/2} \widehat{\boldsymbol{g}}$, and

$$\begin{aligned}
\boldsymbol{M}_{(k)} &= (\boldsymbol{m}^0, \ldots, \boldsymbol{m}^k) \in \mathbb{R}^{N \times (k+1)}, & \boldsymbol{N}_{(k)} &= (\boldsymbol{n}^0, \ldots, \boldsymbol{n}^{k-1}) \in \mathbb{R}^{M \times k}, \\
\bar{\boldsymbol{H}}_{(k)} &= (\bar{\boldsymbol{h}}^1, \ldots, \bar{\boldsymbol{h}}^k) \in \mathbb{R}^{N \times k}, & \breve{\boldsymbol{H}}_{(k)} &= (\breve{\boldsymbol{h}}^0, \ldots, \breve{\boldsymbol{h}}^k) \in \mathbb{R}^{M \times (k+1)}.
\end{aligned}$$

Note that on event (2.54),

$$\frac{1}{N}\boldsymbol{M}_{(k)}^\top \boldsymbol{M}_{(k)} = \widehat{\Sigma}_{\le k} + o_\upsilon(1), \qquad \frac{1}{N}\boldsymbol{N}_{(k)}^\top \boldsymbol{N}_{(k)} = \dot{\Sigma}_{\le k} + o_\upsilon(1), \tag{2.55}$$

$$\frac{1}{N}\bar{\boldsymbol{H}}_{(k)}^\top \bar{\boldsymbol{H}}_{(k)} = \dot{\Sigma}_{\le k} + o_\upsilon(1), \qquad \frac{1}{M}\breve{\boldsymbol{H}}_{(k)}^\top \breve{\boldsymbol{H}}_{(k)} = \widehat{\Sigma}_{\le k} + o_\upsilon(1), \tag{2.56}$$

where $o_\upsilon(1)$ denotes an additive error of operator norm $o_\upsilon(1)$. That is, $\{\boldsymbol{n}^0, \ldots, \boldsymbol{n}^{k-1}\}$ and $\{\bar{\boldsymbol{h}}^1, \ldots, \bar{\boldsymbol{h}}^k\}$ span k-dimensional subspaces of $\mathbb{R}^M$ and $\mathbb{R}^N$, and the linear mapping between them that sends $\boldsymbol{n}^i$ to $\bar{\boldsymbol{h}}^{i+1}$ is an approximate isometry. The same is true, after scaling by a factor $\alpha_\star$, for $\{\boldsymbol{m}^0, \ldots, \boldsymbol{m}^k\}$ and $\{\breve{\boldsymbol{h}}^0, \ldots, \breve{\boldsymbol{h}}^k\}$. Define the linear maps

$$\dot{\boldsymbol{T}} = \bar{\boldsymbol{H}}_{(k)}(\boldsymbol{N}_{(k)}^\top \boldsymbol{N}_{(k)})^{-1}\boldsymbol{N}_{(k)}^\top, \qquad \hat{\boldsymbol{T}} = \breve{\boldsymbol{H}}_{(k)}(\boldsymbol{M}_{(k)}^\top \boldsymbol{M}_{(k)})^{-1}\boldsymbol{M}_{(k)}^\top.$$

(The inverses are well-defined because the matrices are full-rank, by (2.55).) That is, $\dot{\boldsymbol{T}}$ (resp. $\hat{\boldsymbol{T}}$) projects onto the span of $\{\boldsymbol{n}^0, \ldots, \boldsymbol{n}^{k-1}\}$ (resp. $\{\boldsymbol{m}^0, \ldots, \boldsymbol{m}^k\}$) and then applies the linear map that sends $\boldsymbol{n}^i$ to $\dot{\boldsymbol{h}}^{i+1}$ (resp. $\boldsymbol{m}^i$ to $\widehat{\boldsymbol{h}}^i$).

Lemma 2.6.9 (Post-AMP Gordon's inequality)**.** *Conditional on any realization of* DATA *satisfying event* (2.54)*, the following holds. Let* $\dot{\boldsymbol{\xi}} \sim \mathcal{N}(0, \boldsymbol{I}_N)$*,* $\widehat{\boldsymbol{\xi}} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_M)$*,* $Z \sim \mathcal{N}(0,1)$ *be independent of everything else and*

$$\dot{\boldsymbol{g}}'_{\mathrm{AMP}}(\widehat{\boldsymbol{v}}) = \sqrt{N}\dot{\boldsymbol{T}}\widehat{\boldsymbol{v}} + \|P_{\boldsymbol{N}_{(k)}}^\perp \widehat{\boldsymbol{v}}\| P_{\boldsymbol{M}_{(k)}}^\perp \dot{\boldsymbol{\xi}}, \qquad \widehat{\boldsymbol{g}}'_{\mathrm{AMP}}(\dot{\boldsymbol{v}}) = \sqrt{N}\hat{\boldsymbol{T}}\dot{\boldsymbol{v}} + \|P_{\boldsymbol{M}_{(k)}}^\perp \dot{\boldsymbol{v}}\| P_{\boldsymbol{N}_{(k)}}^\perp \widehat{\boldsymbol{\xi}}.$$

For any continuous $f : \mathbb{R}^N \times \mathbb{R}^M \times \mathbb{R}^N \times (\mathbb{R}^M)^2 \times \mathbb{R}^{N\times(k+1)} \times \mathbb{R}^{M\times(k+2)} \to \mathbb{R}$*,*

$$\sup_{\substack{(\boldsymbol{m},\boldsymbol{n})\in U(r_0)\\ \acute{\boldsymbol{h}}\in U'(r_0)}} \sup_{\substack{\|\dot{\boldsymbol{v}}\|=1\\ \dot{\boldsymbol{v}}\perp\boldsymbol{m}}} \inf_{\substack{\|\widehat{\boldsymbol{v}}\|=r_\varepsilon,\\ \widehat{\boldsymbol{v}}\perp\boldsymbol{n}}} \left\{ f(\dot{\boldsymbol{v}}, \widehat{\boldsymbol{v}}; \boldsymbol{m}, \boldsymbol{n}, \acute{\boldsymbol{h}}, \mathsf{DATA}) + \frac{2}{\sqrt{N}}\langle \boldsymbol{G}\dot{\boldsymbol{v}}, \widehat{\boldsymbol{v}}\rangle + \frac{2\|P_{\boldsymbol{N}_{(k)}}^\perp \widehat{\boldsymbol{v}}\|\|P_{\boldsymbol{M}_{(k)}}^\perp \dot{\boldsymbol{v}}\|}{\sqrt{N}} Z \right\}$$

is stochastically dominated by

$$\sup_{\substack{(\boldsymbol{m},\boldsymbol{n})\in U(r_0)\\ \acute{\boldsymbol{h}}\in U'(r_0)}} \sup_{\substack{\|\dot{\boldsymbol{v}}\|=1\\ \dot{\boldsymbol{v}}\perp\boldsymbol{m}^k}} \inf_{\substack{\|\widehat{\boldsymbol{v}}\|=r_\varepsilon,\\ \widehat{\boldsymbol{v}}\perp\boldsymbol{n}^k}} \left\{ f(\dot{\boldsymbol{v}}, \widehat{\boldsymbol{v}}; \boldsymbol{m}, \boldsymbol{n}, \acute{\boldsymbol{h}}, \mathsf{DATA}) + \frac{2}{\sqrt{N}}\langle \dot{\boldsymbol{v}}, \dot{\boldsymbol{g}}'_{\mathrm{AMP}}(\widehat{\boldsymbol{v}})\rangle + \frac{2}{\sqrt{N}}\langle \widehat{\boldsymbol{v}}, \widehat{\boldsymbol{g}}'_{\mathrm{AMP}}(\dot{\boldsymbol{v}})\rangle \right\} + o_\upsilon(1).$$

Proof. We will first show that conditional on DATA,

$$\frac{1}{\sqrt{N}}\boldsymbol{G} \overset{d}{=} \dot{\boldsymbol{T}}^\top + \hat{\boldsymbol{T}} + o_\upsilon(1) + \frac{P_{\boldsymbol{N}_{(k)}^\perp}\overline{\boldsymbol{G}}P_{\boldsymbol{M}_{(k)}}^\perp}{\sqrt{N}}, \tag{2.57}$$

where $o_\upsilon(1)$ is a deterministic error of operator norm $o_\upsilon(1)$ and $\overline{\boldsymbol{G}}$ is an i.i.d. copy of $\boldsymbol{G}$. Conditioning on DATA amounts to conditioning on the linear relations

$$\frac{1}{\sqrt{N}}\boldsymbol{G}\boldsymbol{m}^i = \breve{\boldsymbol{h}}^i + \varrho_\varepsilon \boldsymbol{n}^{i-1}, \qquad \frac{1}{\sqrt{N}}\boldsymbol{G}^\top \boldsymbol{n}^i = \bar{\boldsymbol{h}}^{i+1} + d_\varepsilon \boldsymbol{m}^i \tag{2.58}$$

for $0 \le i \le k$ and $0 \le i \le k-1$. So, $P_{\boldsymbol{N}_{(k)}}^\perp \boldsymbol{G} P_{\boldsymbol{M}_{(k)}}^\perp$ is independent of DATA and $\boldsymbol{G} - P_{\boldsymbol{N}_{(k)}}^\perp \boldsymbol{G} P_{\boldsymbol{M}_{(k)}}^\perp$ is DATA-measurable. It suffices to show the latter part is $\dot{\boldsymbol{T}}^\top + \hat{\boldsymbol{T}}$, up to $o_\upsilon(1)$ additive operator norm error. Recall from (2.55) that the condition number of $\frac{1}{N}\boldsymbol{M}_{(k)}^\top \boldsymbol{M}_{(k)}$ and $\frac{1}{N}\boldsymbol{N}_{(k)}^\top \boldsymbol{N}_{(k)}$ is bounded depending on k. So it suffices to show

$$\left\|\frac{1}{\sqrt{N}}\boldsymbol{G}\boldsymbol{M}_{(k)} - (\dot{\boldsymbol{T}}^\top + \hat{\boldsymbol{T}})\boldsymbol{M}_{(k)}\right\|_{\mathsf{op}} = o_\upsilon(1)\sqrt{N}, \quad \left\|\frac{1}{\sqrt{N}}\boldsymbol{G}^\top \boldsymbol{N}_{(k)} - (\dot{\boldsymbol{T}} + \hat{\boldsymbol{T}}^\top)\boldsymbol{N}_{(k)}\right\|_{\mathsf{op}} = o_\upsilon(1)\sqrt{N}. \tag{2.59}$$

By (2.58) and the definition of $\dot{\boldsymbol{T}}, \hat{\boldsymbol{T}}$,

$$\frac{1}{\sqrt{N}}\boldsymbol{G}\boldsymbol{M}_{(k)} = \breve{\boldsymbol{H}}_{(k)} + \varrho_\varepsilon[\boldsymbol{0}, \boldsymbol{N}_{(k)}], \qquad \frac{1}{\sqrt{N}}\boldsymbol{G}^\top\boldsymbol{N}_{(k)} = \bar{\boldsymbol{H}}_{(k)} + d_\varepsilon\boldsymbol{M}_{(k-1)},$$
$$\hat{\boldsymbol{T}}\boldsymbol{M}_{(k)} = \breve{\boldsymbol{H}}_{(k)}, \qquad \dot{\boldsymbol{T}}\boldsymbol{N}_{(k)} = \bar{\boldsymbol{H}}_{(k)}.$$

For all $i,j \geq 1$, we have by gaussian integration by parts

$$\begin{aligned}\frac{1}{N}\langle \bar{\boldsymbol{h}}^i, \boldsymbol{m}^j\rangle &= \frac{1}{N}\langle \bar{\boldsymbol{h}}^i, \mathrm{th}_\varepsilon(\bar{\boldsymbol{h}}^j + \varepsilon^{1/2}\dot{\boldsymbol{g}})\rangle \\ &= \mathbb{E}[(\overline{\psi}_{i\wedge j}^{1/2}Z + (\psi_\varepsilon + \varepsilon - \overline{\psi}_{i\wedge j})Z')\mathrm{th}_\varepsilon(\overline{\psi}_{i\wedge j}^{1/2}Z + (\psi_\varepsilon + \varepsilon - \overline{\psi}_{i\wedge j})^{1/2}Z'')] + o_\upsilon(1) \\ &= \varrho_\varepsilon\overline{\psi}_{i\wedge j} + o_\upsilon(1).\end{aligned}$$

Moreover $\frac{1}{N}\langle \bar{\boldsymbol{h}}^i, \boldsymbol{m}^0\rangle = o_\upsilon(1)$. Thus,

$$\begin{aligned}\dot{\boldsymbol{T}}^\top\boldsymbol{M}_{(k)} &= \boldsymbol{N}_{(k)}\left(\frac{1}{N}\boldsymbol{N}_{(k)}^\top\boldsymbol{N}_{(k)}\right)^{-1}\left(\frac{1}{N}\bar{\boldsymbol{H}}_{(k)}^\top\boldsymbol{M}_{(k)}\right) \\ &= \boldsymbol{N}_{(k)}\left(\dot{\Sigma}_{\leq k} + o_\upsilon(1)\right)^{-1}\left([0, \varrho_\varepsilon\dot{\Sigma}_{\leq k}] + o_\upsilon(1)\right) = \varrho_\varepsilon[\boldsymbol{0}, \boldsymbol{N}_{(k)}] + o_\upsilon(1)\sqrt{N},\end{aligned}$$

where the errors are all in operator norm. A similar calculation shows

$$\hat{\boldsymbol{T}}^\top\boldsymbol{N}_{(k)} = d_\varepsilon\boldsymbol{M}_{(k-1)} + o_\upsilon(1)\sqrt{N}.$$

Combining proves (2.59) and thus (2.57). So, conditional on DATA,

$$\frac{1}{\sqrt{N}}\langle \boldsymbol{G}\dot{\boldsymbol{v}}, \widehat{\boldsymbol{v}}\rangle \stackrel{d}{=} \langle \dot{\boldsymbol{v}}, \dot{\boldsymbol{T}}\widehat{\boldsymbol{v}}\rangle + \langle \widehat{\boldsymbol{v}}, \hat{\boldsymbol{T}}\dot{\boldsymbol{v}}\rangle + o_\upsilon(1) + \frac{1}{\sqrt{N}}\langle \overline{\boldsymbol{G}}P^\perp_{\boldsymbol{M}_{(k)}}\dot{\boldsymbol{v}}, P^\perp_{\boldsymbol{N}_{(k)}}\widehat{\boldsymbol{v}}\rangle$$

By Gordon's inequality applied to $\overline{\boldsymbol{G}}$,

$$\begin{aligned}\sup_{\substack{(\boldsymbol{m},\boldsymbol{n})\in U(r_0)\\ \acute{\boldsymbol{h}}\in U'(r_0)}}\ \sup_{\substack{\|\dot{\boldsymbol{v}}\|=1\\ \dot{\boldsymbol{v}}\perp\boldsymbol{m}}}\ \inf_{\substack{\|\widehat{\boldsymbol{v}}\|=r_\varepsilon,\\ \widehat{\boldsymbol{v}}\perp\boldsymbol{n}}}\Bigg\{& f(\dot{\boldsymbol{v}}, \widehat{\boldsymbol{v}}; \boldsymbol{m}, \boldsymbol{n}, \acute{\boldsymbol{h}}, \mathsf{DATA}) + 2\langle \dot{\boldsymbol{v}}, \dot{\boldsymbol{T}}\widehat{\boldsymbol{v}}\rangle + 2\langle \widehat{\boldsymbol{v}}, \hat{\boldsymbol{T}}\dot{\boldsymbol{v}}\rangle \\ &+ \frac{2}{\sqrt{N}}\langle \overline{\boldsymbol{G}}P^\perp_{\boldsymbol{M}_{(k)}}\dot{\boldsymbol{v}}, P^\perp_{\boldsymbol{N}_{(k)}}\widehat{\boldsymbol{v}}\rangle + \frac{2\|P^\perp_{\boldsymbol{N}_{(k)}}\widehat{\boldsymbol{v}}\|\|P^\perp_{\boldsymbol{M}_{(k)}}\dot{\boldsymbol{v}}\|}{\sqrt{N}}Z\Bigg\}\end{aligned}$$

is stochastically dominated by

$$\begin{aligned}\sup_{\substack{(\boldsymbol{m},\boldsymbol{n})\in U(r_0)\\ \acute{\boldsymbol{h}}\in U'(r_0)}}\ \sup_{\substack{\|\dot{\boldsymbol{v}}\|=1\\ \dot{\boldsymbol{v}}\perp\boldsymbol{m}}}\ \inf_{\substack{\|\widehat{\boldsymbol{v}}\|=r_\varepsilon,\\ \widehat{\boldsymbol{v}}\perp\boldsymbol{n}}}\Bigg\{& f(\dot{\boldsymbol{v}}, \widehat{\boldsymbol{v}}; \boldsymbol{m}, \boldsymbol{n}, \acute{\boldsymbol{h}}, \mathsf{DATA}) + 2\langle \dot{\boldsymbol{v}}, \dot{\boldsymbol{T}}\widehat{\boldsymbol{v}}\rangle + 2\langle \widehat{\boldsymbol{v}}, \hat{\boldsymbol{T}}\dot{\boldsymbol{v}}\rangle \\ &+ \frac{2\|P^\perp_{\boldsymbol{N}_{(k)}}\widehat{\boldsymbol{v}}\|}{\sqrt{N}}\langle \dot{\boldsymbol{v}}, P^\perp_{\boldsymbol{M}_{(k)}}\dot{\boldsymbol{\xi}}\rangle + \frac{2\|P^\perp_{\boldsymbol{M}_{(k)}}\dot{\boldsymbol{v}}\|}{\sqrt{N}}\langle \widehat{\boldsymbol{v}}, P^\perp_{\boldsymbol{N}_{(k)}}\widehat{\boldsymbol{\xi}}\rangle\Bigg\}.\end{aligned}$$

The quantity inside the sup-inf is precisely $f(\dot{\boldsymbol{v}}, \widehat{\boldsymbol{v}}, \mathsf{DATA}) + \frac{2}{\sqrt{N}}\langle \dot{\boldsymbol{v}}, \dot{\boldsymbol{g}}'_{\mathrm{AMP}}(\widehat{\boldsymbol{v}})\rangle + \frac{2}{\sqrt{N}}\langle \widehat{\boldsymbol{v}}, \widehat{\boldsymbol{g}}'_{\mathrm{AMP}}(\dot{\boldsymbol{v}})\rangle$. □

Define

$$\dot{\boldsymbol{g}}_{\mathrm{AMP}}(\widehat{\boldsymbol{v}}) = \sqrt{N}\dot{\boldsymbol{T}}\widehat{\boldsymbol{v}} + \|P^\perp_{\boldsymbol{N}_{(k)}}\widehat{\boldsymbol{v}}\|\dot{\boldsymbol{\xi}}, \qquad \widehat{\boldsymbol{g}}_{\mathrm{AMP}}(\dot{\boldsymbol{v}}) = \sqrt{N}\hat{\boldsymbol{T}}\dot{\boldsymbol{v}} + \|P^\perp_{\boldsymbol{M}_{(k)}}\dot{\boldsymbol{v}}\|\widehat{\boldsymbol{\xi}}.$$

Note that

$$\frac{1}{\sqrt{N}}\|\dot{\boldsymbol{g}}_{\mathrm{AMP}}(\widehat{\boldsymbol{v}}) - \dot{\boldsymbol{g}}'_{\mathrm{AMP}}(\widehat{\boldsymbol{v}})\| \leq \frac{r_\varepsilon}{\sqrt{N}}\|P_{\boldsymbol{M}_{(k)}}\dot{\boldsymbol{\xi}}\|, \qquad \frac{1}{\sqrt{N}}\|\widehat{\boldsymbol{g}}_{\mathrm{AMP}}(\dot{\boldsymbol{v}}) - \dot{\boldsymbol{g}}'_{\mathrm{AMP}}(\dot{\boldsymbol{v}})\| \leq \frac{1}{\sqrt{N}}\|P_{\boldsymbol{N}_{(k)}}\widehat{\boldsymbol{\xi}}\|,$$

are both bounded by υ with high probability, and similarly $|Z|/\sqrt{N} \le \upsilon$ with high probability. Below, let err denote an error term of order $o_{r_0}(1) + o_k(1) + o_\upsilon(1)$. By (2.53), Lemma 2.6.9, and these observations, it suffices to show that with high probability,

$$\sup_{\substack{(\boldsymbol{m},\boldsymbol{n})\in U(r_0)\\ \acute{\boldsymbol{h}}\in U'(r_0)}} \sup_{\substack{\|\dot{\boldsymbol{v}}\|=1\\ \dot{\boldsymbol{v}}\perp\boldsymbol{m}}} \inf_{\substack{\|\widehat{\boldsymbol{v}}\|=r_\varepsilon,\\ \widehat{\boldsymbol{v}}\perp\boldsymbol{n}}} \Bigg\{ -\langle \boldsymbol{D}_1\dot{\boldsymbol{v}},\dot{\boldsymbol{v}}\rangle + \langle \boldsymbol{D}_2(\acute{\boldsymbol{h}})^{-1}\widehat{\boldsymbol{v}},\widehat{\boldsymbol{v}}\rangle + \frac{2}{\sqrt{N}}\langle \dot{\boldsymbol{v}}, \dot{\boldsymbol{g}}_{\rm AMP}(\widehat{\boldsymbol{v}})\rangle + \frac{2}{\sqrt{N}}\langle \widehat{\boldsymbol{v}}, \widehat{\boldsymbol{g}}_{\rm AMP}(\dot{\boldsymbol{v}})\rangle \Bigg\} \le \lambda_\varepsilon + d_\varepsilon + \mathsf{err}. \tag{2.60}$$

Lemma 2.6.10. *Let*

$$\dot{\mu}'_{\rm AMP} = \frac{1}{N}\sum_{i=1}^N \delta(\dot{\xi}_i, \bar{h}_i^1,\ldots,\bar{h}_i^k), \qquad \widehat{\mu}'_{\rm AMP} = \frac{1}{M}\sum_{a=1}^M \delta(\widehat{\xi}_a, \breve{h}_a^0,\ldots,\breve{h}_a^k).$$

Conditional on a realization of DATA *such that* (2.54) *holds, with high probability,*

$$\mathbb{W}_2(\dot{\mu}'_{\rm AMP}, \mathcal{N}(0,1)\times\mathcal{N}(0,\dot{\Sigma}_{\le k})), \mathbb{W}_2(\widehat{\mu}'_{\rm AMP}, \mathcal{N}(0,1)\times\mathcal{N}(0,\widehat{\Sigma}_{\le k})) \le 2\upsilon. \tag{2.61}$$

Proof. Under event (2.54), the $\mathbb{W}_2$-distance of the marginal of $\dot{\mu}'_{\rm AMP}$ on all but the first coordinate to $\mathcal{N}(0,\dot{\Sigma}_{\le k}))$ is deterministically at most υ. Since $\dot{\boldsymbol{\xi}}$ is independent of DATA, it follows that $\mathbb{W}_2(\dot{\mu}'_{\rm AMP}, \mathcal{N}(0,1)\times \mathcal{N}(0,\dot{\Sigma}_{\le k})) \le 2\upsilon$ with high probability. The estimate for $\widehat{\mu}'_{\rm AMP}$ is analogous. □

Fact 2.6.11 (Proved in Appendix 2.A)**.** *Let $\mu,\mu' \in \mathcal{P}_2(\mathbb{R}^3)$, and suppose the marginals of μ have fourth moments. Suppose f_1,f_2,f_3 are L-Lipschitz functions, and f_3 is bounded by L. Then there exists $C = C(\mu, L)$ such that*

$$|\mathbb{E}_{(x,y,z)\sim\mu} f_1(x)f_2(y)f_3(z) - \mathbb{E}_{(x',y',z')\sim\mu'} f_1(x')f_2(y')f_3(z')| \le C\max(\mathbb{W}_2(\mu,\mu'), \mathbb{W}_2(\mu,\mu')^2). \tag{2.62}$$

Lemma 2.6.12. *Suppose* (2.61) *holds. Uniformly over $(\boldsymbol{m},\boldsymbol{n})\in U(r_0)$, $\acute{\boldsymbol{h}}\in U'(r_0)$, $\dot{\boldsymbol{v}}\in\{\|\dot{\boldsymbol{v}}\|=1, \dot{\boldsymbol{v}}\perp\boldsymbol{m}\}$,*

$$\mathbb{W}_2\left(\frac{1}{M}\sum_{a=1}^M \delta(\widehat{h}_a^k, \acute{h}_a, n_a, \widehat{g}_{\rm AMP}(\dot{\boldsymbol{v}})_a), (\widetilde{q}_\varepsilon^{1/2}Z, \widetilde{q}_\varepsilon^{1/2}Z, F_{\varepsilon,\varrho_\varepsilon}(\widetilde{q}_\varepsilon^{1/2}Z), Z')\right) \le \mathsf{err}. \tag{2.63}$$

Similarly, uniformly over $(\boldsymbol{m},\boldsymbol{n})\in U(r_0)$, $\widehat{\boldsymbol{v}}\in\{\|\widehat{\boldsymbol{v}}\| = r_\varepsilon, \widehat{\boldsymbol{v}}\perp\boldsymbol{n}\}$,

$$\mathbb{W}_2\left(\frac{1}{N}\sum_{i=1}^N \delta(\dot{h}_i^k, m_i, \dot{g}_{\rm AMP}(\widehat{\boldsymbol{v}})_i), (\widetilde{\psi}_\varepsilon^{1/2}Z, \mathrm{th}_\varepsilon(\widetilde{\psi}_\varepsilon^{1/2}Z), r_\varepsilon Z')\right) \le \mathsf{err}. \tag{2.64}$$

Proof. We first show that for any $\dot{\boldsymbol{v}}' \in \{\|\dot{\boldsymbol{v}}'\|=1, \dot{\boldsymbol{v}}'\perp\boldsymbol{m}\}$,

$$\mathbb{W}_2\left(\frac{1}{M}\sum_{a=1}^M \delta(\widehat{h}_a^k, \widehat{g}_{\rm AMP}(\dot{\boldsymbol{v}}')_a), (\widetilde{q}_\varepsilon^{1/2}Z, Z')\right) = o_\upsilon(1). \tag{2.65}$$

Indeed, let $\dot{\boldsymbol{v}}' = \frac{1}{\sqrt{N}}\boldsymbol{M}_{(k)}\dot{\vec{v}} + P^\perp_{\boldsymbol{M}_{(k)}}\dot{\boldsymbol{v}}'$ for some $\dot{\vec{v}}\in\mathbb{R}^{k+1}$, so that $\widehat{\boldsymbol{g}}_{\rm AMP}(\dot{\boldsymbol{v}}') = \breve{\boldsymbol{H}}_{(k)}\dot{\vec{v}} + \|P^\perp_{\boldsymbol{M}_{(k)}}\dot{\boldsymbol{v}}'\|\widehat{\boldsymbol{\xi}}$. By the approximate isometry (2.55), (2.56), since $\frac{1}{\sqrt{N}}\boldsymbol{M}_{(k)}\dot{\vec{v}}\perp\boldsymbol{m}^k$, we have $\frac{1}{N}\langle\breve{\boldsymbol{h}}^k, \breve{\boldsymbol{H}}_{(k)}\dot{\vec{v}}\rangle = o_\upsilon(1)$. (Since υ is small depending on k, we may take it much smaller than the condition number of $\widehat{\Sigma}_{\le k}$.) By this isometry,

$$\mathbb{W}_2\left(\frac{1}{M}\sum_{a=1}^M \delta(\widehat{h}_a^k, (\breve{\boldsymbol{H}}_{(k)}\dot{\vec{v}})_a), (\widetilde{q}_\varepsilon^{1/2}Z, \|P_{\boldsymbol{M}_{(k)}}\dot{\boldsymbol{v}}'\|Z')\right) = o_\upsilon(1).$$

Then (2.61) implies (2.65). Now consider $(\boldsymbol{m},\boldsymbol{n})\in U(r_0)$ and let T be a rotation operator mapping $\boldsymbol{m}/\|\boldsymbol{m}\|$ to $\boldsymbol{m}^k/\|\boldsymbol{m}^k\|$. Note that $\|T-I\|_{\mathsf{op}} = o_{r_0}(1)$. Consider any $\dot{\boldsymbol{v}}\in\{\|\dot{\boldsymbol{v}}\|=1, \dot{\boldsymbol{v}}\perp\boldsymbol{m}\}$, and let $\dot{\boldsymbol{v}}' = T\dot{\boldsymbol{v}}$. Then,

$$\|\widehat{\boldsymbol{g}}_{\rm AMP}(\dot{\boldsymbol{v}}') - \widehat{\boldsymbol{g}}_{\rm AMP}(\dot{\boldsymbol{v}})\| \le (\sqrt{N}\|\hat{\boldsymbol{T}}\|_{\mathsf{op}} + \|\widehat{\boldsymbol{\xi}}\|)\|\dot{\boldsymbol{v}}' - \dot{\boldsymbol{v}}\| \le \sqrt{N}(\|\hat{\boldsymbol{T}}\|_{\mathsf{op}} + O(1))o_{r_0}(1).$$

Note that

$$\|\hat{\boldsymbol{T}}\|_{\mathsf{op}} = \sup_{\dot{\vec{v}} \in \mathbb{R}^{k+1}} \frac{\|\hat{\boldsymbol{T}}\boldsymbol{M}_{(k)}\dot{\vec{v}}\|}{\|\boldsymbol{M}_{(k)}\dot{\vec{v}}\|} = \sup_{\dot{\vec{v}} \in \mathbb{R}^{k+1}} \frac{\|\breve{\boldsymbol{H}}\dot{\vec{v}}\|}{\|\boldsymbol{M}_{(k)}\dot{\vec{v}}\|} = \sup_{\dot{\vec{v}} \in \mathbb{R}^{k+1}} \sqrt{\frac{\langle \frac{1}{N}\breve{\boldsymbol{H}}^\top \breve{\boldsymbol{H}}, \dot{\vec{v}}^{\otimes 2}\rangle}{\langle \frac{1}{N}\boldsymbol{M}^\top \boldsymbol{M}, \dot{\vec{v}}^{\otimes 2}\rangle}}$$

is bounded by an absolute constant by (2.55), (2.56). Thus $\|\widehat{\boldsymbol{g}}_{\mathrm{AMP}}(\dot{\boldsymbol{v}}') - \widehat{\boldsymbol{g}}_{\mathrm{AMP}}(\dot{\boldsymbol{v}})\| \le o_{r_0}(1)\sqrt{N}$. By (2.52) and definition of $U'(r_0)$,

$$\|\widehat{\boldsymbol{h}}^k - \acute{\boldsymbol{h}}\| \le \|\widehat{\boldsymbol{h}}^k - \acute{\boldsymbol{h}}^k\| + \|\acute{\boldsymbol{h}}^k - \acute{\boldsymbol{h}}\| \le (o_k(1) + o_{r_0}(1))\sqrt{N}. \tag{2.66}$$

Similarly,

$$\|F_{\varepsilon,\varrho_\varepsilon}(\widehat{\boldsymbol{h}}^k) - \boldsymbol{n}\| = \|\boldsymbol{n}^k - \boldsymbol{n}\| \le o_{r_0}(1)\sqrt{N}. \tag{2.67}$$

Combining these bounds with (2.65) proves (2.63). (2.64) is proved similarly. $\square$

Proposition 2.6.13. *If* (2.61) *holds, uniformly over* $(\boldsymbol{m},\boldsymbol{n}) \in U(r_0)$, $\acute{\boldsymbol{h}} \in U'(r_0)$, $\dot{\boldsymbol{v}} \in \{\|\dot{\boldsymbol{v}}\| = 1, \dot{\boldsymbol{v}} \perp \boldsymbol{m}\}$,

$$\inf_{\substack{\|\widehat{\boldsymbol{v}}\|=r_\varepsilon,\\ \widehat{\boldsymbol{v}}\perp\boldsymbol{n}}} \langle \boldsymbol{D}_2(\acute{\boldsymbol{h}})^{-1}\widehat{\boldsymbol{v}}, \widehat{\boldsymbol{v}}\rangle + \frac{2}{\sqrt{N}}\langle \widehat{\boldsymbol{v}}, \widehat{\boldsymbol{g}}_{\mathrm{AMP}}(\dot{\boldsymbol{v}})\rangle \le -\alpha_\star\,\mathbb{E}\left[\frac{\widehat{f}_\varepsilon(\widetilde{q}_\varepsilon^{1/2}Z)}{1 + m_\varepsilon(z_\varepsilon)\widehat{f}_\varepsilon(\widetilde{q}_\varepsilon^{1/2}Z)}\right] - m_\varepsilon(z_\varepsilon)r_\varepsilon^2 + \mathsf{err}.$$

Proof. Let

$$\widehat{\boldsymbol{v}}' = -\frac{1}{\sqrt{N}}\left(\boldsymbol{D}_2(\acute{\boldsymbol{h}})^{-1} + m_\varepsilon(z_\varepsilon)I\right)^{-1}\widehat{\boldsymbol{g}}_{\mathrm{AMP}}(\dot{\boldsymbol{v}}).$$

Note the identity

$$\alpha_\star\,\mathbb{E}\left[\left(\frac{\widehat{f}_\varepsilon(\widetilde{q}_\varepsilon^{1/2}Z)}{1 + m_\varepsilon(z_\varepsilon)\widehat{f}_\varepsilon(\widetilde{q}_\varepsilon^{1/2}Z)}\right)^2\right] = \frac{\alpha_\star\theta_\varepsilon(z_\varepsilon)}{\mathbb{E}[(z_\varepsilon + \dot{f}_\varepsilon(\widetilde{\psi}_\varepsilon^{1/2}Z))^{-2}]} = r_\varepsilon^2. \tag{2.68}$$

Then,

$$\begin{aligned}
\|\widehat{\boldsymbol{v}}'\|^2 &= \frac{1}{N}\widehat{\boldsymbol{g}}_{\mathrm{AMP}}(\dot{\boldsymbol{v}})^\top\left(\widetilde{\boldsymbol{D}}_2(\acute{\boldsymbol{h}})^{-1} + m_\varepsilon(z_\varepsilon)I\right)^{-2}\widehat{\boldsymbol{g}}_{\mathrm{AMP}}(\dot{\boldsymbol{v}}) \\
&= \frac{\alpha_\star}{M}\sum_{a=1}^M \left(\frac{\widehat{f}_\varepsilon(\acute{h}_a)}{1 + m_\varepsilon(z_\varepsilon)\widehat{f}_\varepsilon(\acute{h}_a)}\right)^2 \widehat{g}_{\mathrm{AMP}}(\dot{\boldsymbol{v}})_a^2 \\
&= \alpha_\star\,\mathbb{E}\left[\left(\frac{\widehat{f}_\varepsilon(\widetilde{q}_\varepsilon^{1/2}Z)}{1 + m_\varepsilon(z_\varepsilon)\widehat{f}_\varepsilon(\widetilde{q}_\varepsilon^{1/2}Z)}\right)^2 (Z')^2\right] + \mathsf{err} = r_\varepsilon^2 + \mathsf{err}.
\end{aligned}$$

In the last line we used Lemma 2.6.12 and Fact 2.6.11, with $f_1(x) = f_2(x) = x$, $f_3(x) = (\frac{\widehat{f}_\varepsilon(x)}{1+m_\varepsilon(z_\varepsilon)\widehat{f}_\varepsilon(x)})^2$. (Note that we have not shown the coordinate empirical measure in (2.63) has bounded fourth moments, but it suffices for Fact 2.6.11 that the gaussian approximating it does.) Similarly,

$$\begin{aligned}
\frac{1}{\sqrt{N}}\langle \widehat{\boldsymbol{v}}', \boldsymbol{n}\rangle &= -\frac{\alpha_\star}{M}\sum_{a=1}^M \left(\frac{\widehat{f}_\varepsilon(\acute{h}_a)}{1 + m_\varepsilon(z_\varepsilon)\widehat{f}_\varepsilon(\acute{h}_a)}\right) n_a \widehat{\boldsymbol{g}}_{\mathrm{AMP}}(\dot{\boldsymbol{v}})_a \\
&= -\alpha_\star\,\mathbb{E}\left[\left(\frac{\widehat{f}_\varepsilon(\widetilde{q}_\varepsilon^{1/2}Z)}{1 + m_\varepsilon(z_\varepsilon)\widehat{f}_\varepsilon(\widetilde{q}_\varepsilon^{1/2}Z)}\right) F_{\varepsilon,\varrho_\varepsilon}(\widetilde{q}_\varepsilon^{1/2}Z)Z'\right] + \mathsf{err} = \mathsf{err}.
\end{aligned}$$

Likewise,

$$\begin{aligned}
\langle (\boldsymbol{D}_2(\acute{\boldsymbol{h}})^{-1} + m_\varepsilon(z_\varepsilon)\boldsymbol{I}_M)\widehat{\boldsymbol{v}}', \widehat{\boldsymbol{v}}'\rangle = -\frac{1}{\sqrt{N}}\langle \widehat{\boldsymbol{v}}', \widehat{\boldsymbol{g}}_{\mathrm{AMP}}(\dot{\boldsymbol{v}})\rangle &= \frac{\alpha_\star}{M}\sum_{a=1}^M \left(\frac{\widehat{f}_\varepsilon(\acute{h}_a)}{1 + m_\varepsilon(z_\varepsilon)\widehat{f}_\varepsilon(\acute{h}_a)}\right)\widehat{\boldsymbol{g}}_{\mathrm{AMP}}(\dot{\boldsymbol{v}})_a^2 \\
&= \alpha_\star\,\mathbb{E}\left[\frac{\widehat{f}_\varepsilon(\widetilde{q}_\varepsilon^{1/2}Z)}{1 + m_\varepsilon(z_\varepsilon)\widehat{f}_\varepsilon(\widetilde{q}_\varepsilon^{1/2}Z)}\right] + \mathsf{err}.
\end{aligned}$$

From this, it follows that

$$\langle \boldsymbol{D}_2(\acute{\boldsymbol{h}})^{-1}\widehat{\boldsymbol{v}}', \widehat{\boldsymbol{v}}'\rangle + \frac{2}{\sqrt{N}}\langle \widehat{\boldsymbol{v}}', \widehat{\boldsymbol{g}}_{\mathrm{AMP}}(\dot{\boldsymbol{v}})\rangle = -\alpha_\star\,\mathbb{E}\left[\frac{\widehat{f}_\varepsilon(\widetilde{q}_\varepsilon^{1/2}Z)}{1+m_\varepsilon(z_\varepsilon)\widehat{f}_\varepsilon(\widetilde{q}_\varepsilon^{1/2}Z)}\right] - m_\varepsilon(z_\varepsilon)r_\varepsilon^2 + \mathsf{err}.$$

By the above estimates on $\|\widehat{\boldsymbol{v}}'\|^2$ and $\frac{1}{\sqrt{N}}\langle \widehat{\boldsymbol{v}}', \boldsymbol{n}\rangle$, we can find $\widehat{\boldsymbol{v}}$ such that $\|\widehat{\boldsymbol{v}}\| = r_\varepsilon$, $\widehat{\boldsymbol{v}} \perp \boldsymbol{n}$, and $\|\widehat{\boldsymbol{v}} - \widehat{\boldsymbol{v}}'\| \le \mathsf{err}$. Since $\boldsymbol{D}_2(\acute{\boldsymbol{h}})^{-1}$ has operator norm bounded independently of r_0, k, υ,

$$|\langle \boldsymbol{D}_2(\acute{\boldsymbol{h}})^{-1}\widehat{\boldsymbol{v}}, \widehat{\boldsymbol{v}}\rangle - \langle \boldsymbol{D}_2^{-1}\widehat{\boldsymbol{v}}', \widehat{\boldsymbol{v}}'\rangle| \le 2\|\boldsymbol{D}_2^{-1}(\acute{\boldsymbol{h}})\|_{\mathsf{op}}\|\widehat{\boldsymbol{v}} - \widehat{\boldsymbol{v}}'\| \le \mathsf{err}.$$

By Cauchy–Schwarz,

$$\frac{2}{\sqrt{N}}|\langle \widehat{\boldsymbol{v}}, \widehat{\boldsymbol{g}}_{\mathrm{AMP}}(\dot{\boldsymbol{v}})\rangle - \langle \widehat{\boldsymbol{v}}', \widehat{\boldsymbol{g}}_{\mathrm{AMP}}(\dot{\boldsymbol{v}})\rangle| \le \frac{2}{\sqrt{N}}\|\widehat{\boldsymbol{g}}_{\mathrm{AMP}}(\dot{\boldsymbol{v}})\|\|\widehat{\boldsymbol{v}} - \widehat{\boldsymbol{v}}'\| \le \mathsf{err}.$$

This completes the proof. $\square$

Proposition 2.6.14. *If* (2.61) *holds, uniformly over* $(\boldsymbol{m}, \boldsymbol{n}) \in U(r_0)$, $\widehat{\boldsymbol{v}} \in \{\|\widehat{\boldsymbol{v}}\| = r_\varepsilon, \widehat{\boldsymbol{v}} \perp \boldsymbol{n}\}$, *we have*

$$\sup_{\substack{\|\dot{\boldsymbol{v}}\|=1\\ \dot{\boldsymbol{v}}\perp\boldsymbol{m}}} -\langle \boldsymbol{D}_1\dot{\boldsymbol{v}}, \dot{\boldsymbol{v}}\rangle + \frac{2}{\sqrt{N}}\langle \dot{\boldsymbol{v}}, \dot{\boldsymbol{g}}_{\mathrm{AMP}}(\widehat{\boldsymbol{v}})\rangle \le z_\varepsilon + m_\varepsilon(z_\varepsilon)r_\varepsilon^2 + \mathsf{err}.$$

Proof. Fix any $(\boldsymbol{m}, \boldsymbol{n})$ and $\widehat{\boldsymbol{v}}$ satisfying the stated conditions. We estimate

$$\sup_{\substack{\|\dot{\boldsymbol{v}}\|=1\\ \dot{\boldsymbol{v}}\perp\boldsymbol{m}}} -\langle \boldsymbol{D}_1\dot{\boldsymbol{v}}, \dot{\boldsymbol{v}}\rangle + \frac{2}{\sqrt{N}}\langle \dot{\boldsymbol{v}}, \dot{\boldsymbol{g}}_{\mathrm{AMP}}(\widehat{\boldsymbol{v}})\rangle \le \sup_{\dot{\boldsymbol{v}}\perp\boldsymbol{m}} -\langle \boldsymbol{D}_1\dot{\boldsymbol{v}}, \dot{\boldsymbol{v}}\rangle + \frac{2}{\sqrt{N}}\langle \dot{\boldsymbol{v}}, \dot{\boldsymbol{g}}_{\mathrm{AMP}}(\widehat{\boldsymbol{v}})\rangle - z_\varepsilon\left(\|\dot{\boldsymbol{v}}\|^2 - 1\right). \tag{2.69}$$

Note that $-\boldsymbol{D}_1 - z_\varepsilon \boldsymbol{I}_N$ is negative definite, as $z_\varepsilon > -\frac{1}{1+\varepsilon} = \max_{x\in\mathbb{R}}\{-\dot{f}(x)\}$. So, the supremum on the right-hand side of (2.69) is maximized by $\dot{\boldsymbol{v}}$ solving the stationarity condition (in $\mathsf{span}(\boldsymbol{m})^\perp$):

$$\dot{\boldsymbol{v}} = \frac{1}{\sqrt{N}}P_{\boldsymbol{m}}^\perp(\boldsymbol{D}_1 + z_\varepsilon \boldsymbol{I}_N)^{-1}P_{\boldsymbol{m}}^\perp\dot{\boldsymbol{g}}_{\mathrm{AMP}}(\widehat{\boldsymbol{v}}).$$

Let

$$\dot{\boldsymbol{v}}' = \frac{1}{\sqrt{N}}(\boldsymbol{D}_1 + z_\varepsilon \boldsymbol{I}_N)^{-1}\dot{\boldsymbol{g}}_{\mathrm{AMP}}(\widehat{\boldsymbol{v}}).$$

Note that, by Fact 2.6.11 and Lemma 2.6.12,

$$\begin{aligned}\langle (\boldsymbol{D}_1 + z_\varepsilon \boldsymbol{I}_N)\dot{\boldsymbol{v}}', \dot{\boldsymbol{v}}'\rangle = \frac{1}{\sqrt{N}}\langle \dot{\boldsymbol{v}}', \dot{\boldsymbol{g}}_{\mathrm{AMP}}(\widehat{\boldsymbol{v}})\rangle &= \frac{1}{N}\sum_{i=1}^N \dot{g}_{\mathrm{AMP}}(\widehat{\boldsymbol{v}})_i^2(\dot{f}_\varepsilon(\dot{h}_i) + z_\varepsilon)^{-1}\\ &= r_\varepsilon^2\,\mathbb{E}\left[(\dot{f}_\varepsilon(\widetilde{\psi}_\varepsilon Z) + z_\varepsilon)^{-1}\right] + \mathsf{err}\\ &= m_\varepsilon(z_\varepsilon)r_\varepsilon^2 + \mathsf{err}.\end{aligned}$$

Thus

$$-\langle \boldsymbol{D}_1\dot{\boldsymbol{v}}', \dot{\boldsymbol{v}}'\rangle + \frac{2}{\sqrt{N}}\langle \dot{\boldsymbol{v}}', \dot{\boldsymbol{g}}_{\mathrm{AMP}}(\widehat{\boldsymbol{v}})\rangle - z_\varepsilon\left(\|\dot{\boldsymbol{v}}'\|^2 - 1\right) = z_\varepsilon + m_\varepsilon(z_\varepsilon)r_\varepsilon^2 + \mathsf{err}.$$

We now estimate $\|\dot{\boldsymbol{v}} - \dot{\boldsymbol{v}}'\|$. Note that

$$\|\dot{\boldsymbol{v}} - \dot{\boldsymbol{v}}'\| \le \|(\boldsymbol{D}_1 + z_\varepsilon \boldsymbol{I}_N)^{-1}\|_{\mathsf{op}}\|P_{\boldsymbol{m}}\dot{\boldsymbol{g}}_{\mathrm{AMP}}(\widehat{\boldsymbol{v}})\| + \|P_{\boldsymbol{m}}(\boldsymbol{D}_1 + z_\varepsilon \boldsymbol{I}_N)^{-1}\dot{\boldsymbol{g}}_{\mathrm{AMP}}(\widehat{\boldsymbol{v}})\|,$$

and by Fact 2.6.11 and Lemma 2.6.12, both terms on the right-hand side are bounded by err. Since $\boldsymbol{D}_1 + z_\varepsilon \boldsymbol{I}_N$ has bounded operator norm,

$$|\langle (\boldsymbol{D}_1 + z_\varepsilon \boldsymbol{I}_N)\dot{\boldsymbol{v}}, \dot{\boldsymbol{v}}\rangle - \langle (\boldsymbol{D}_1 + z_\varepsilon \boldsymbol{I}_N)\dot{\boldsymbol{v}}', \dot{\boldsymbol{v}}'\rangle| \le 2\|\boldsymbol{D}_1 + z_\varepsilon \boldsymbol{I}_N\|_{\mathsf{op}}\|\dot{\boldsymbol{v}} - \dot{\boldsymbol{v}}'\| \le \mathsf{err}.$$

By Cauchy–Schwarz,

$$\frac{2}{\sqrt{N}}|\langle \dot{\boldsymbol{v}}', \dot{\boldsymbol{g}}_{\mathrm{AMP}}(\widehat{\boldsymbol{v}})\rangle - \langle \dot{\boldsymbol{v}}, \dot{\boldsymbol{g}}_{\mathrm{AMP}}(\widehat{\boldsymbol{v}})\rangle| \le \frac{2}{\sqrt{N}}\|\dot{\boldsymbol{g}}_{\mathrm{AMP}}(\widehat{\boldsymbol{v}})\|\|\dot{\boldsymbol{v}} - \dot{\boldsymbol{v}}'\| \le \mathsf{err}.$$

Combining completes the proof. $\square$

Proof of Proposition 2.6.7. By Propositions 2.6.13 and 2.6.14, on the high probability event (2.61), the left-hand side of (2.60) is bounded by

$$z_\varepsilon - \alpha_\star \mathbb{E}\left[\frac{\widehat{f}_\varepsilon(\widetilde{q}_\varepsilon^{1/2} Z)}{1 + m_\varepsilon(z_\varepsilon)\widehat{f}_\varepsilon(\widetilde{q}_\varepsilon^{1/2} Z)}\right] + \mathsf{err} = \lambda_\varepsilon + d_\varepsilon + \mathsf{err}.$$

This proves (2.60), and by the discussion leading to (2.60) the proposition follows. □

Proof of Proposition 2.4.8(c), under $\mathbb{P}$. By Proposition 2.4.8(a), with high probability, $(\boldsymbol{m}^k, \boldsymbol{n}^k) \in \mathcal{S}_{\varepsilon,\upsilon_0}$. Recall that $\mathrm{th}_\varepsilon, F_{\varepsilon,\varrho_\varepsilon}$ are $O(1)$-Lipschitz, with $O_\varepsilon(1)$-Lipschitz inverses (i.e. Lipschitz constant depending only on ε). On this event, for υ_0 small depending on r_0 and some $C_\varepsilon = O_\varepsilon(1)$,

$$U(r_0) \subseteq \mathcal{S}_{\varepsilon,\upsilon_0 + C_\varepsilon r_0} \subseteq \mathcal{S}_{\varepsilon,2C_\varepsilon r_0}. \tag{2.70}$$

Since $\|\boldsymbol{G}\|_{\mathsf{op}}, \|\widehat{\boldsymbol{g}}\| \le C\sqrt{N}$ holds with high probability under $\mathbb{P}$, Lemma 2.6.4 applies. Applying this lemma with $2C_\varepsilon r_0$ in place of r_0 shows that for all $(\boldsymbol{m}, \boldsymbol{n}) \in U(r_0)$,

$$\nabla^2_\diamond \mathcal{F}^\varepsilon_{\mathsf{TAP}}(\boldsymbol{m}, \boldsymbol{n}) \preceq \boldsymbol{R}(\boldsymbol{m}, \boldsymbol{n}) + \lambda_\varepsilon P_{\boldsymbol{m}} + (o_{C_{\mathsf{cvx}}}(1) + o_{r_0}(1))\boldsymbol{I}_N.$$

Combined with Proposition 2.6.7, this gives that with high probability,

$$\nabla^2_\diamond \mathcal{F}(\boldsymbol{m}, \boldsymbol{n}) \preceq (\lambda_\varepsilon + o_{C_{\mathsf{cvx}}}(1) + o_{r_0}(1) + o_k(1))\boldsymbol{I}_N.$$

By Lemma 2.6.3,

$$\nabla^2_\diamond \mathcal{F}(\boldsymbol{m}, \boldsymbol{n}) \preceq (\lambda_0 + o_\varepsilon(1) + o_{C_{\mathsf{cvx}}}(1) + o_{r_0}(1) + o_k(1))\boldsymbol{I}_N.$$

Under Condition 2.3.4, $\lambda_0 < 0$. The conclusion follows by setting the parameters so the error term in the last display is bounded by $|\lambda_0|/2$. □

Remark 2.6.15. The bound $\lambda_\varepsilon + d_\varepsilon$ in Proposition 2.6.7 is tight. One way to see this is to calculate the upper edge of the limiting spectral measure of

$$\boldsymbol{A} = P^\perp_{\boldsymbol{M}_{(k)}} (-\boldsymbol{D}_1 - \boldsymbol{W}) P^\perp_{\boldsymbol{M}_{(k)}}, \qquad \text{where} \qquad \boldsymbol{W} = \frac{1}{N}\boldsymbol{G}^\top P^\perp_{\boldsymbol{N}^{(k)}} \boldsymbol{D}_2 P^\perp_{\boldsymbol{N}^{(k)}} \boldsymbol{G},$$

using free probability [Voi91]. We now outline this calculation. Note that conditional on DATA, $-\boldsymbol{D}_1$ and $-\boldsymbol{W}$ are orthogonally invariant as quadratic forms on $\mathsf{span}(\boldsymbol{m}^0, \ldots, \boldsymbol{m}^k)^\perp$. The inverse Cauchy transform of $-\boldsymbol{D}_1$ is approximated within err by $m_\varepsilon^{-1}(t)$. By e.g. [BS98, Equation 1.2], the inverse Cauchy transform of $-\boldsymbol{W}$ is approximated within err by

$$\frac{1}{t} - \alpha_\star \mathbb{E}\left[\frac{\widehat{f}_\varepsilon(\widetilde{q}_\varepsilon^{1/2} Z)}{1 + t\widehat{f}_\varepsilon(\widetilde{q}_\varepsilon^{1/2} Z)}\right],$$

Since R-transforms add under free additive convolution, $\boldsymbol{A}$ has limiting inverse Cauchy transform

$$\vartheta_\varepsilon(t) = m_\varepsilon^{-1}(t) - \alpha_\star \mathbb{E}\left[\frac{\widehat{f}_\varepsilon(\widetilde{q}_\varepsilon^{1/2} Z)}{1 + t\widehat{f}_\varepsilon(\widetilde{q}_\varepsilon^{1/2} Z)}\right].$$

One calculates that

$$\vartheta_\varepsilon'(t) = -\mathbb{E}[(m_\varepsilon^{-1}(t) + \dot{f}_\varepsilon(\widetilde{\psi}_\varepsilon^{1/2} Z))^{-2}]^{-1} + \mathbb{E}\left[\left(\frac{\widehat{f}_\varepsilon(\widetilde{q}_\varepsilon^{1/2} Z)}{1 + t\widehat{f}_\varepsilon(\widetilde{q}_\varepsilon^{1/2} Z)}\right)^2\right]$$

has the same sign as $\theta_\varepsilon(m_\varepsilon^{-1}(t)) - \alpha_\star^{-1}$. Thus $\vartheta_\varepsilon(t)$ is decreasing on $(0, m_\varepsilon(z_\varepsilon)]$ and increasing $[m_\varepsilon(z_\varepsilon), +\infty)$. It follows that the limiting spectral measure of $\boldsymbol{A}$ has upper edge $\vartheta_\varepsilon(m_\varepsilon(z_\varepsilon)) = \lambda_\varepsilon + d_\varepsilon$. By the Weyl inequalities the same is true for $\boldsymbol{R}(\boldsymbol{m}, \boldsymbol{n})$, so Proposition 2.6.7 is tight.

2.6.4 Planted model

The proof of Proposition 2.4.8(c) in the planted model is only simpler, as we will be able to apply Gordon's inequality directly rather than conditional on AMP iterates. The main step is the following proposition. Let υ be sufficiently small depending on r_0, k.

Proposition 2.6.16. *Suppose $(\boldsymbol{m}', \boldsymbol{n}') \in \mathcal{S}_{\varepsilon,\upsilon}$. With high probability under $\mathbb{P}^{\boldsymbol{m}',\boldsymbol{n}'}_{\varepsilon,\mathsf{Pl}}$, $\boldsymbol{R}(\boldsymbol{m},\boldsymbol{n}) \preceq (\lambda_\varepsilon + d_\varepsilon + \mathsf{err})P^\perp_{\boldsymbol{m}}$ for all $\|(\boldsymbol{m},\boldsymbol{n}) - (\boldsymbol{m}',\boldsymbol{n}')\| \le 2r_0\sqrt{N}$.*

Let $\dot{\boldsymbol{h}}' = \mathrm{th}_\varepsilon^{-1}(\boldsymbol{m}')$, $\widehat{\boldsymbol{h}}' = F^{-1}_{\varepsilon,\rho_\varepsilon(q(\boldsymbol{m}))}(\boldsymbol{n}')$. By Lemma 2.4.16, under $\mathbb{P}^{\boldsymbol{m},\boldsymbol{n}'}_{\varepsilon,\mathsf{Pl}}$ we have $\acute{\boldsymbol{h}}(\boldsymbol{m}',\boldsymbol{n}',\boldsymbol{G}) = \widehat{\boldsymbol{h}}'$.

For this subsection, let $U(r_0) = \{(\boldsymbol{m},\boldsymbol{n}) : \|(\boldsymbol{m},\boldsymbol{n}) - (\boldsymbol{m}',\boldsymbol{n}')\| \le 2r_0\sqrt{N}\}$ and $U'(r_0) = \{\acute{\boldsymbol{h}} : \|\acute{\boldsymbol{h}} - \widehat{\boldsymbol{h}}'\| \le Cr_0\sqrt{N}\}$, for suitably large constant C. Identically to the discussion above (2.53), to prove Proposition 2.6.16 it suffices to show, with high probability,

$$\sup_{(\boldsymbol{m},\boldsymbol{n})\in U(r_0)} \sup_{\substack{\|\dot{\boldsymbol{v}}\|=1\\ \dot{\boldsymbol{v}}\perp\boldsymbol{m}}} \inf_{\substack{\|\widehat{\boldsymbol{v}}\|=r_\varepsilon,\\ \widehat{\boldsymbol{v}}\perp\boldsymbol{n}}} \left\{-\langle \boldsymbol{D}_1\dot{\boldsymbol{v}},\dot{\boldsymbol{v}}\rangle + \langle \boldsymbol{D}_2(\acute{\boldsymbol{h}})^{-1}\widehat{\boldsymbol{v}},\widehat{\boldsymbol{v}}\rangle + \frac{2}{\sqrt{N}}\langle \boldsymbol{G}\dot{\boldsymbol{v}},\widehat{\boldsymbol{v}}\rangle\right\} \le \lambda_\varepsilon + d_\varepsilon + \mathsf{err}.$$

Lemma 2.6.17. *Let $\dot{\boldsymbol{\xi}}, \dot{\boldsymbol{\xi}}' \sim \mathcal{N}(0,\boldsymbol{I}_N)$, $\widehat{\boldsymbol{\xi}}, \widehat{\boldsymbol{\xi}}' \sim \mathcal{N}(\boldsymbol{0},\boldsymbol{I}_M)$, $Z, Z' \sim \mathcal{N}(0,1)$ be independent of everything else and*

$$\dot{\boldsymbol{g}}'_{\mathsf{Pl}}(\widehat{\boldsymbol{v}}) = \frac{\|P_{\boldsymbol{n}'}\widehat{\boldsymbol{v}}\|(\dot{\boldsymbol{h}}' + \varepsilon^{1/2}P^\perp_{\boldsymbol{m}'}\dot{\boldsymbol{\xi}}')}{\widetilde{\psi}_\varepsilon^{1/2}} + \|P^\perp_{\boldsymbol{n}'}\widehat{\boldsymbol{v}}\|P^\perp_{\boldsymbol{m}'}\dot{\boldsymbol{\xi}}, \quad \widehat{\boldsymbol{g}}'_{\mathsf{Pl}}(\dot{\boldsymbol{v}}) = \frac{\|P_{\boldsymbol{m}'}\dot{\boldsymbol{v}}\|(\widehat{\boldsymbol{h}}' + \varepsilon^{1/2}P^\perp_{\boldsymbol{n}'}\widehat{\boldsymbol{\xi}}')}{\widetilde{q}_\varepsilon^{1/2}} + \|P^\perp_{\boldsymbol{m}'}\dot{\boldsymbol{v}}\|P^\perp_{\boldsymbol{n}'}\widehat{\boldsymbol{\xi}}.$$

For any continuous $f : \mathbb{R}^N \times \mathbb{R}^M \times (\mathbb{R}^N)^2 \times (\mathbb{R}^M)^3 \to \mathbb{R}$,

$$\sup_{\substack{(\boldsymbol{m},\boldsymbol{n})\in U(r_0)\\ \acute{\boldsymbol{h}}\in U'(r_0)}} \sup_{\substack{\|\dot{\boldsymbol{v}}\|=1\\ \dot{\boldsymbol{v}}\perp\boldsymbol{m}}} \inf_{\substack{\|\widehat{\boldsymbol{v}}\|=r_\varepsilon,\\ \widehat{\boldsymbol{v}}\perp\boldsymbol{n}}} \left\{f(\dot{\boldsymbol{v}},\widehat{\boldsymbol{v}};\boldsymbol{m}',\boldsymbol{m},\boldsymbol{n}',\boldsymbol{n},\acute{\boldsymbol{h}}) + \frac{2}{\sqrt{N}}\langle \boldsymbol{G}\dot{\boldsymbol{v}},\widehat{\boldsymbol{v}}\rangle + \frac{2\|P^\perp_{\boldsymbol{n}'}\widehat{\boldsymbol{v}}\|\|P^\perp_{\boldsymbol{m}'}\dot{\boldsymbol{v}}\|}{\sqrt{N}}Z\right\}$$

is stochastically dominated by

$$\begin{aligned}\sup_{\substack{(\boldsymbol{m},\boldsymbol{n})\in U(r_0)\\ \acute{\boldsymbol{h}}\in U'(r_0)}} \sup_{\substack{\|\dot{\boldsymbol{v}}\|=1\\ \dot{\boldsymbol{v}}\perp\boldsymbol{m}}} \inf_{\substack{\|\widehat{\boldsymbol{v}}\|=r_\varepsilon,\\ \widehat{\boldsymbol{v}}\perp\boldsymbol{n}}} &\Bigg\{f(\dot{\boldsymbol{v}},\widehat{\boldsymbol{v}};\boldsymbol{m}',\boldsymbol{m},\boldsymbol{n}',\boldsymbol{n},\acute{\boldsymbol{h}}) + \frac{2}{\sqrt{N}}\langle \dot{\boldsymbol{v}},\dot{\boldsymbol{g}}'_{\mathsf{Pl}}(\widehat{\boldsymbol{v}})\rangle + \frac{2}{\sqrt{N}}\langle \widehat{\boldsymbol{v}},\widehat{\boldsymbol{g}}'_{\mathsf{Pl}}(\dot{\boldsymbol{v}})\rangle \\ &+ \frac{2\varepsilon^{1/2}\|P_{\boldsymbol{n}'}\widehat{\boldsymbol{v}}\|\|P_{\boldsymbol{m}'}\dot{\boldsymbol{v}}\|}{(q_\varepsilon + \psi_\varepsilon + \varepsilon)^{1/2}\sqrt{N}}Z'\Bigg\} + o_\upsilon(1).\end{aligned}$$

Proof. By Corollary 2.4.18, the gaussian process $(\dot{\boldsymbol{v}},\widehat{\boldsymbol{v}}) \mapsto \frac{1}{\sqrt{N}}\langle \boldsymbol{G}\dot{\boldsymbol{v}},\widehat{\boldsymbol{v}}\rangle$ has the form

$$\begin{aligned}\frac{1}{\sqrt{N}}\langle \boldsymbol{G}\dot{\boldsymbol{v}},\widehat{\boldsymbol{v}}\rangle &\stackrel{d}{=} \frac{\langle \dot{\boldsymbol{h}}',\dot{\boldsymbol{v}}\rangle\langle \boldsymbol{n}',\widehat{\boldsymbol{v}}\rangle}{N\widetilde{\psi}_\varepsilon} + \frac{\langle \boldsymbol{m}',\dot{\boldsymbol{v}}\rangle\langle \widehat{\boldsymbol{h}}',\widehat{\boldsymbol{v}}\rangle}{N\widetilde{q}_\varepsilon} + o_\upsilon(1) + \frac{1}{\sqrt{N}}\langle \widetilde{\boldsymbol{G}}\dot{\boldsymbol{v}},\widehat{\boldsymbol{v}}\rangle \\ &= \frac{\|P_{\boldsymbol{n}'}\widehat{\boldsymbol{v}}\|\langle \dot{\boldsymbol{h}}',\dot{\boldsymbol{v}}\rangle}{\widetilde{\psi}_\varepsilon^{1/2}\sqrt{N}} + \frac{\|P_{\boldsymbol{m}'}\dot{\boldsymbol{v}}\|\langle \widehat{\boldsymbol{h}}',\widehat{\boldsymbol{v}}\rangle}{\widetilde{q}_\varepsilon^{1/2}\sqrt{N}} + o_\upsilon(1) + \frac{1}{\sqrt{N}}\langle \widetilde{\boldsymbol{G}}\dot{\boldsymbol{v}},\widehat{\boldsymbol{v}}\rangle.\end{aligned}$$

Here the $o_\upsilon(1)$ is uniform over bounded $\|\dot{\boldsymbol{v}}\|, \|\widehat{\boldsymbol{v}}\|$. Moreover, by (2.40), the random part $\langle \widetilde{\boldsymbol{G}}\dot{\boldsymbol{v}},\widehat{\boldsymbol{v}}\rangle$ expands as

$$\begin{aligned}\langle \widetilde{\boldsymbol{G}}\dot{\boldsymbol{v}},\widehat{\boldsymbol{v}}\rangle &= \langle \widetilde{\boldsymbol{G}}P^\perp_{\boldsymbol{m}'}\dot{\boldsymbol{v}},P^\perp_{\boldsymbol{n}'}\widehat{\boldsymbol{v}}\rangle + \langle \widetilde{\boldsymbol{G}}P^\perp_{\boldsymbol{m}'}\dot{\boldsymbol{v}},P_{\boldsymbol{n}'}\widehat{\boldsymbol{v}}\rangle + \langle \widetilde{\boldsymbol{G}}P_{\boldsymbol{m}'}\dot{\boldsymbol{v}},P^\perp_{\boldsymbol{n}'}\widehat{\boldsymbol{v}}\rangle + \langle \widetilde{\boldsymbol{G}}P_{\boldsymbol{m}'}\dot{\boldsymbol{v}},P_{\boldsymbol{n}'}\widehat{\boldsymbol{v}}\rangle \\ &\stackrel{d}{=} \langle \widetilde{\boldsymbol{G}}P^\perp_{\boldsymbol{m}'}\dot{\boldsymbol{v}},P^\perp_{\boldsymbol{n}'}\widehat{\boldsymbol{v}}\rangle + \frac{\varepsilon^{1/2}}{\widetilde{\psi}_\varepsilon^{1/2}}\|P_{\boldsymbol{n}'}\widehat{\boldsymbol{v}}\|\langle P^\perp_{\boldsymbol{m}'}\dot{\boldsymbol{\xi}}',\dot{\boldsymbol{v}}\rangle + \frac{\varepsilon^{1/2}}{\widetilde{q}_\varepsilon^{1/2}}\|P_{\boldsymbol{m}'}\dot{\boldsymbol{v}}\|\langle P^\perp_{\boldsymbol{n}'}\widehat{\boldsymbol{\xi}}',\widehat{\boldsymbol{v}}\rangle + \frac{\varepsilon^{1/2}\|P_{\boldsymbol{n}'}\widehat{\boldsymbol{v}}\|\|P_{\boldsymbol{m}'}\dot{\boldsymbol{v}}\|}{(q_\varepsilon + \psi_\varepsilon + \varepsilon)^{1/2}}Z'.\end{aligned}$$

Thus, (as processes)

$$\frac{1}{\sqrt{N}}\langle \boldsymbol{G}\dot{\boldsymbol{v}},\widehat{\boldsymbol{v}}\rangle+\frac{\|P_{\boldsymbol{n}'}^{\perp}\widehat{\boldsymbol{v}}\|\|P_{\boldsymbol{m}'}^{\perp}\dot{\boldsymbol{v}}\|}{\sqrt{N}}Z\overset{d}{=}\frac{1}{\sqrt{N}}\langle \widetilde{\boldsymbol{G}}P_{\boldsymbol{m}'}^{\perp}\dot{\boldsymbol{v}},P_{\boldsymbol{n}'}^{\perp}\widehat{\boldsymbol{v}}\rangle+\frac{\|P_{\boldsymbol{n}'}^{\perp}\widehat{\boldsymbol{v}}\|\|P_{\boldsymbol{m}'}^{\perp}\dot{\boldsymbol{v}}\|}{\sqrt{N}}Z$$
$$+\frac{\|P_{\boldsymbol{n}'}\widehat{\boldsymbol{v}}\|\langle\dot{\boldsymbol{h}}'+\varepsilon^{1/2}P_{\boldsymbol{m}'}^{\perp}\dot{\boldsymbol{\xi}}',\dot{\boldsymbol{v}}\rangle}{\widetilde{\psi}_\varepsilon^{1/2}\sqrt{N}}+\frac{\|P_{\boldsymbol{m}'}\dot{\boldsymbol{v}}\|\langle\widehat{\boldsymbol{h}}'+\varepsilon^{1/2}P_{\boldsymbol{n}'}^{\perp}\widehat{\boldsymbol{\xi}}',\widehat{\boldsymbol{v}}\rangle}{\widetilde{q}_\varepsilon^{1/2}\sqrt{N}}$$
$$+\frac{\varepsilon^{1/2}\|P_{\boldsymbol{n}'}\widehat{\boldsymbol{v}}\|\|P_{\boldsymbol{m}'}\dot{\boldsymbol{v}}\|}{(q_\varepsilon+\psi_\varepsilon+\varepsilon)^{1/2}\sqrt{N}}Z'+o_\upsilon(1).$$

The result now follows by using Gordon's inequality to compare $\frac{1}{\sqrt{N}}\langle \widetilde{\boldsymbol{G}}P_{\boldsymbol{m}'}^{\perp}\dot{\boldsymbol{v}},P_{\boldsymbol{n}'}^{\perp}\widehat{\boldsymbol{v}}\rangle+\frac{\|P_{\boldsymbol{n}'}^{\perp}\widehat{\boldsymbol{v}}\|\|P_{\boldsymbol{m}'}^{\perp}\dot{\boldsymbol{v}}\|}{\sqrt{N}}Z$ to $\frac{1}{\sqrt{N}}\|P_{\boldsymbol{n}'}^{\perp}\widehat{\boldsymbol{v}}\|\langle\dot{\boldsymbol{v}},P_{\boldsymbol{m}'}^{\perp}\dot{\boldsymbol{\xi}}\rangle+\frac{1}{\sqrt{N}}\|P_{\boldsymbol{m}'}^{\perp}\dot{\boldsymbol{v}}\|\langle\widehat{\boldsymbol{v}},P_{\boldsymbol{n}'}^{\perp}\widehat{\boldsymbol{\xi}}\rangle$. □

Let

$$\dot{\boldsymbol{g}}_{\mathsf{PI}}(\widehat{\boldsymbol{v}})=\frac{\|P_{\boldsymbol{n}'}\widehat{\boldsymbol{v}}\|(\dot{\boldsymbol{h}}'+\varepsilon^{1/2}\dot{\boldsymbol{\xi}}')}{\widetilde{\psi}_\varepsilon^{1/2}}+\|P_{\boldsymbol{n}'}^{\perp}\widehat{\boldsymbol{v}}\|\dot{\boldsymbol{\xi}},\qquad \widehat{\boldsymbol{g}}_{\mathsf{PI}}(\dot{\boldsymbol{v}})=\frac{\|P_{\boldsymbol{m}'}\dot{\boldsymbol{v}}\|(\widehat{\boldsymbol{h}}'+\varepsilon^{1/2}\widehat{\boldsymbol{\xi}}')}{\widetilde{q}_\varepsilon^{1/2}}+\|P_{\boldsymbol{m}'}^{\perp}\dot{\boldsymbol{v}}\|\widehat{\boldsymbol{\xi}}.$$

As argued above (2.60), with high probability,

$$\frac{1}{\sqrt{N}}|Z|,\frac{1}{\sqrt{N}}|Z'|,\frac{1}{\sqrt{N}}\sup_{\|\widehat{\boldsymbol{v}}\|=r_\varepsilon}\|\dot{\boldsymbol{g}}_{\mathsf{PI}}(\widehat{\boldsymbol{v}})-\dot{\boldsymbol{g}}'_{\mathsf{PI}}(\widehat{\boldsymbol{v}})\|,\frac{1}{\sqrt{N}}\sup_{\|\dot{\boldsymbol{v}}\|=1}\|\widehat{\boldsymbol{g}}_{\mathsf{PI}}(\dot{\boldsymbol{v}})-\widehat{\boldsymbol{g}}'_{\mathsf{PI}}(\dot{\boldsymbol{v}})\|\le\upsilon.$$

So it suffices to show that with high probability,

$$\sup_{\substack{(\boldsymbol{m},\boldsymbol{n})\in U(r_0)\\ \acute{\boldsymbol{h}}\in U'(r_0)}}\sup_{\substack{\|\dot{\boldsymbol{v}}\|=1\\ \dot{\boldsymbol{v}}\perp\boldsymbol{m}}}\inf_{\substack{\|\widehat{\boldsymbol{v}}\|=r_\varepsilon,\\ \widehat{\boldsymbol{v}}\perp\boldsymbol{n}}}\Bigg\{-\langle\boldsymbol{D}_1\dot{\boldsymbol{v}},\dot{\boldsymbol{v}}\rangle+\langle\boldsymbol{D}_2(\acute{\boldsymbol{h}})^{-1}\widehat{\boldsymbol{v}},\widehat{\boldsymbol{v}}\rangle$$
$$+\frac{2}{\sqrt{N}}\langle\dot{\boldsymbol{v}},\dot{\boldsymbol{g}}_{\mathsf{PI}}(\widehat{\boldsymbol{v}})\rangle+\frac{2}{\sqrt{N}}\langle\widehat{\boldsymbol{v}},\widehat{\boldsymbol{g}}_{\mathsf{PI}}(\dot{\boldsymbol{v}})\rangle\Bigg\}\le\lambda_\varepsilon+d_\varepsilon+\mathsf{err}.\tag{2.71}$$

Lemma 2.6.18. *For all $(\boldsymbol{m}',\boldsymbol{n}')\in\mathcal{S}_{\varepsilon,\upsilon}$, the following holds with high probability. Uniformly over $(\boldsymbol{m},\boldsymbol{n})\in U(r_0)$, $\acute{\boldsymbol{h}}\in U'(r_0)$, $\dot{\boldsymbol{v}}\in\{\|\dot{\boldsymbol{v}}\|=1,\dot{\boldsymbol{v}}\perp\boldsymbol{m}\}$,*

$$\mathbb{W}_2\left(\frac{1}{M}\sum_{a=1}^M\delta(\widehat{h}'_a,\acute{h}_a,n'_a,\widehat{g}_{\mathsf{PI}}(\dot{\boldsymbol{v}})_a),(\widetilde{q}_\varepsilon^{1/2}Z,\widetilde{q}_\varepsilon^{1/2}Z,F_{\varepsilon,\varrho_\varepsilon}(\widetilde{q}_\varepsilon^{1/2}Z),Z')\right)\le\mathsf{err}.\tag{2.72}$$

Similarly, uniformly over $(\boldsymbol{m}',\boldsymbol{n}')\in\mathcal{S}_{\varepsilon,\upsilon}$, $(\boldsymbol{m},\boldsymbol{n})\in U(r_0)$, $\widehat{\boldsymbol{v}}\in\{\|\widehat{\boldsymbol{v}}\|=r_\varepsilon,\widehat{\boldsymbol{v}}\perp\boldsymbol{n}\}$,

$$\mathbb{W}_2\left(\frac{1}{N}\sum_{i=1}^N\delta(\dot{h}'_i,m'_i,\dot{g}_{\mathsf{PI}}(\widehat{\boldsymbol{v}})_i),(\widetilde{\psi}_\varepsilon^{1/2}Z,\mathrm{th}_\varepsilon(\widetilde{\psi}_\varepsilon^{1/2}Z),r_\varepsilon Z')\right)\le\mathsf{err}.\tag{2.73}$$

Proof. Let $\widehat{\boldsymbol{h}}''=F_{\varepsilon,\varrho_\varepsilon}^{-1}(\boldsymbol{n}')$. Consider first $\dot{\boldsymbol{v}}'\in\{\|\dot{\boldsymbol{v}}'\|=1,\dot{\boldsymbol{v}}'\perp\boldsymbol{m}\}$, Then $\widehat{\boldsymbol{g}}_{\mathsf{PI}}(\dot{\boldsymbol{v}}')=\widehat{\boldsymbol{\xi}}$, so clearly

$$\mathbb{W}_2\left(\frac{1}{M}\sum_{a=1}^M\delta(\widehat{h}''_a,\widehat{g}_{\mathsf{PI}}(\dot{\boldsymbol{v}}')_a),(\widetilde{q}_\varepsilon^{1/2}Z,Z')\right)=o_\upsilon(1).$$

For $(\boldsymbol{m},\boldsymbol{n})\in U(r_0)$, let T be a rotation operator mapping $\boldsymbol{m}/\|\boldsymbol{m}\|$ to $\boldsymbol{m}'/\|\boldsymbol{m}'\|$. Note that $\|T-I\|_{\mathsf{op}}=o_{r_0}(1)$. Consider any $\dot{\boldsymbol{v}}\in\{\|\dot{\boldsymbol{v}}\|=1,\dot{\boldsymbol{v}}\perp\boldsymbol{m}\}$, and let $\dot{\boldsymbol{v}}'=T\dot{\boldsymbol{v}}$, so $\|\dot{\boldsymbol{v}}-\dot{\boldsymbol{v}}'\|=o_{r_0}(1)$. Then

$$\|\widehat{\boldsymbol{g}}_{\mathsf{PI}}(\dot{\boldsymbol{v}}')-\widehat{\boldsymbol{g}}_{\mathsf{PI}}(\dot{\boldsymbol{v}})\|\le O(1)\left(\|\widehat{\boldsymbol{h}}'\|+\|\widehat{\boldsymbol{\xi}}'\|+\|\widehat{\boldsymbol{\xi}}\|\right)\|\dot{\boldsymbol{v}}-\dot{\boldsymbol{v}}'\|.$$

With high probability over $\widehat{\boldsymbol{\xi}},\widehat{\boldsymbol{\xi}}'$, this is bounded by $o_{r_0}(1)\sqrt{N}$. Thus

$$\mathbb{W}_2\left(\frac{1}{M}\sum_{a=1}^M\delta(\widehat{h}''_a,\widehat{g}_{\mathsf{PI}}(\dot{\boldsymbol{v}})_a),(\widetilde{q}_\varepsilon^{1/2}Z,Z')\right)=o_{r_0}(1)+o_\upsilon(1).\tag{2.74}$$

Note that

$$\|\widehat{\boldsymbol{h}}' - \widehat{\boldsymbol{h}}''\| = \|F^{-1}_{\varepsilon,\rho_\varepsilon(q(\boldsymbol{m}))}(\boldsymbol{n}') - F^{-1}_{\varepsilon,\varrho_\varepsilon}(\boldsymbol{n}')\| \le \mathsf{err}\sqrt{N}.$$

Identically to (2.66) and (2.67), we can show

$$\|\widehat{\boldsymbol{h}}' - \acute{\boldsymbol{h}}\|, \|F_{\varepsilon,\varrho_\varepsilon}(\widehat{\boldsymbol{h}}'') - \boldsymbol{n}\| \le \mathsf{err}\sqrt{N}.$$

Combined with (2.74), this proves (2.72). The proof of (2.73) is analogous. □

The following two propositions are proved identically to Propositions 2.6.13 and 2.6.14, with $\widehat{\boldsymbol{g}}_{\mathsf{Pl}}$, $\dot{\boldsymbol{g}}_{\mathsf{Pl}}$, and Lemma 2.6.18 playing the roles of $\widehat{\boldsymbol{g}}_{\mathrm{AMP}}$, $\dot{\boldsymbol{g}}_{\mathrm{AMP}}$, and Lemma 2.6.12.

Proposition 2.6.19. *For all $(\boldsymbol{m}', \boldsymbol{n}') \in \mathcal{S}_{\varepsilon,\upsilon}$, the following holds with high probability. Uniformly over $(\boldsymbol{m}, \boldsymbol{n}) \in U(r_0)$, $\acute{\boldsymbol{h}} \in U'(r_0)$, $\dot{\boldsymbol{v}} \in \{\|\dot{\boldsymbol{v}}\| = 1, \dot{\boldsymbol{v}} \perp \boldsymbol{m}\}$, we have*

$$\inf_{\substack{\|\widehat{\boldsymbol{v}}\|=r_\varepsilon,\\ \widehat{\boldsymbol{v}}\perp\boldsymbol{n}}} \langle \boldsymbol{D}_2(\acute{\boldsymbol{h}})^{-1}\widehat{\boldsymbol{v}}, \widehat{\boldsymbol{v}}\rangle + \frac{2}{\sqrt{N}}\langle \widehat{\boldsymbol{v}}, \widehat{\boldsymbol{g}}_{\mathsf{Pl}}(\dot{\boldsymbol{v}})\rangle \le -\alpha_\star \mathbb{E}\left[\frac{\widehat{f}_\varepsilon(\widetilde{q}_\varepsilon^{1/2}Z)}{1 + m_\varepsilon(z_\varepsilon)\widehat{f}_\varepsilon(\widetilde{q}_\varepsilon^{1/2}Z)}\right] - m_\varepsilon(z_\varepsilon) r_\varepsilon^2 + \mathsf{err}.$$

Proposition 2.6.20. *For all $(\boldsymbol{m}', \boldsymbol{n}') \in \mathcal{S}_{\varepsilon,\upsilon}$, the following holds with high probability. Uniformly over $(\boldsymbol{m}, \boldsymbol{n}) \in U(r_0)$, $\widehat{\boldsymbol{v}} \in \{\|\widehat{\boldsymbol{v}}\| = r_\varepsilon, \widehat{\boldsymbol{v}} \perp \boldsymbol{n}\}$, we have*

$$\sup_{\substack{\|\dot{\boldsymbol{v}}\|=1\\ \dot{\boldsymbol{v}}\perp\boldsymbol{m}}} -\langle \boldsymbol{D}_1\dot{\boldsymbol{v}}, \dot{\boldsymbol{v}}\rangle + \frac{2}{\sqrt{N}}\langle \dot{\boldsymbol{v}}, \dot{\boldsymbol{g}}_{\mathsf{Pl}}(\widehat{\boldsymbol{v}})\rangle \le z_\varepsilon + m_\varepsilon(z_\varepsilon) r_\varepsilon^2 + \mathsf{err}.$$

Proof of Proposition 2.6.16. Adding Propositions 2.6.19 and 2.6.20 shows that (2.71) holds with high probability. The result follows from the discussion leading to (2.71). □

Proof of Proposition 2.4.8(c), under $\mathbb{P}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{Pl}}$. By Proposition 2.4.8(d), $\|(\boldsymbol{m}^k, \boldsymbol{n}^k) - (\boldsymbol{m}, \boldsymbol{n})\| = \upsilon_0\sqrt{N}$ with high probability. We set $\upsilon_0 < r_0$. Since we defined

$$U(r_0) = \{(\boldsymbol{m}, \boldsymbol{n}) : \|(\boldsymbol{m}, \boldsymbol{n}) - (\boldsymbol{m}', \boldsymbol{n}')\| \le 2r_0\sqrt{N}\} \supseteq \{(\boldsymbol{m}, \boldsymbol{n}) : \|(\boldsymbol{m}, \boldsymbol{n}) - (\boldsymbol{m}^k, \boldsymbol{n}^k)\| \le r_0\sqrt{N}\},$$

the conclusion of Proposition 2.6.16 holds for all $\|(\boldsymbol{m}, \boldsymbol{n}) - (\boldsymbol{m}^k, \boldsymbol{n}^k)\| \le r_0\sqrt{N}$. Identically to (2.70), we have

$$\{(\boldsymbol{m}, \boldsymbol{n}) : \|(\boldsymbol{m}, \boldsymbol{n}) - (\boldsymbol{m}^k, \boldsymbol{n}^k)\| \le r_0\sqrt{N}\} \subseteq \mathcal{S}_{\varepsilon, 2C_\varepsilon r_0}$$

for some $C_\varepsilon = O_\varepsilon(1)$. Since $\|\boldsymbol{G}\|_{\mathsf{op}}, \|\widehat{\boldsymbol{g}}\| \le C\sqrt{N}$ holds with high probability under $\mathbb{P}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{Pl}}$, Lemma 2.6.4 holds. Applying this lemma (with $2C_\varepsilon r_0$ in place of r_0) gives that for all $\|(\boldsymbol{m}, \boldsymbol{n}) - (\boldsymbol{m}^k, \boldsymbol{n}^k)\| \le r_0\sqrt{N}$,

$$\begin{aligned}
\nabla^2_\diamond \mathcal{F}^\varepsilon_{\mathsf{TAP}}(\boldsymbol{m}, \boldsymbol{n}) &\preceq \boldsymbol{R}(\boldsymbol{m}, \boldsymbol{n}) + \lambda_\varepsilon P_{\boldsymbol{m}} + (o_{C_{\mathsf{cvx}}}(1) + o_{r_0}(1))\boldsymbol{I}_N \\
&\preceq (\lambda_\varepsilon + o_{C_{\mathsf{cvx}}}(1) + o_{r_0}(1) + o_k(1))\boldsymbol{I}_N \\
&\preceq (\lambda_0 + o_\varepsilon(1) + o_{C_{\mathsf{cvx}}}(1) + o_{r_0}(1) + o_k(1))\boldsymbol{I}_N.
\end{aligned}$$

Under Condition 2.3.4, $\lambda_0 < 0$, and the result follows by setting the error terms small. □

2.6.5 Determinant concentration

In this subsection, we prove Lemma 2.4.9. We fix some $(\boldsymbol{m}, \boldsymbol{n}) \in \mathcal{S}_{\varepsilon,\upsilon}$ and work under the measure $\mathbb{P}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{Pl}}$. Define, as in Lemma 2.4.16,

$$\dot{\boldsymbol{h}} = \mathrm{th}_\varepsilon^{-1}(\boldsymbol{m}), \qquad \widehat{\boldsymbol{h}} = F^{-1}_{\varepsilon,\rho_\varepsilon(\boldsymbol{m})}(\boldsymbol{n}), \qquad \acute{\boldsymbol{h}} = \frac{\boldsymbol{G}\boldsymbol{m}}{\sqrt{N}} + \varepsilon^{1/2}\widehat{\boldsymbol{g}} - \rho_\varepsilon(q(\boldsymbol{m}))\boldsymbol{n}.$$

Recall from Lemma 2.4.16 that under $\mathbb{P}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{Pl}}$, we have $\acute{\boldsymbol{h}} = \widehat{\boldsymbol{h}}$ deterministically. We computed $\nabla^2\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m}, \boldsymbol{n})$ in Fact 2.6.5, and under $\mathbb{P}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{Pl}}$ the matrices $\boldsymbol{D}_1, \widetilde{\boldsymbol{D}}_2, \boldsymbol{D}_3, \boldsymbol{D}_4$ therein are all nonrandom. By Schur's lemma,

$$|\det \nabla^2\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m}, \boldsymbol{n})| = |\det \nabla^2_{\boldsymbol{n},\boldsymbol{n}}\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m}, \boldsymbol{n})||\det \nabla^2_\diamond\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m}, \boldsymbol{n})|, \tag{2.75}$$

and $\nabla^2_{\boldsymbol{n},\boldsymbol{n}}\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n})$ is nonrandom. By Fact 2.6.5,

$$\nabla^2_{\diamond}\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n}) = -\boldsymbol{D}_1 - \frac{1}{N}\boldsymbol{G}^\top \widetilde{\boldsymbol{D}}_2\boldsymbol{G} + \rho_\varepsilon'(q(\boldsymbol{m}))d_\varepsilon(\boldsymbol{m},\boldsymbol{n})\boldsymbol{I}_N + \frac{C}{N}\boldsymbol{m}\boldsymbol{m}^\top + \frac{1}{N}(\boldsymbol{G}^\top\boldsymbol{v}\boldsymbol{m}^\top + \boldsymbol{m}\boldsymbol{v}^\top\boldsymbol{G})$$

for some nonrandom $C \in \mathbb{R}$, $\boldsymbol{v} \in \mathbb{R}^M$ depending on $(\boldsymbol{m},\boldsymbol{n})$. By Lemma 2.6.6, $|C|, \|\boldsymbol{v}\|$ are uniformly bounded over $(\boldsymbol{m},\boldsymbol{n}) \in \mathcal{S}_{\varepsilon,\upsilon}$, with bound depending on $\varepsilon, C_{\mathsf{cvx}}$. Define for convenience the nonrandom matrix

$$\boldsymbol{A} = \boldsymbol{D}_1 - \rho_\varepsilon'(q(\boldsymbol{m}))d_\varepsilon(\boldsymbol{m},\boldsymbol{n})\boldsymbol{I}_N - \frac{C}{N}\boldsymbol{m}\boldsymbol{m}^\top$$

and note that $\|\boldsymbol{A}\|_{\mathsf{op}}$ is uniformly bounded (depending on $\varepsilon, C_{\mathsf{cvx}}$) over $(\boldsymbol{m},\boldsymbol{n}) \in \mathcal{S}_{\varepsilon,\upsilon}$. Then let

$$\boldsymbol{X} = \begin{bmatrix} \boldsymbol{A} & \frac{1}{\sqrt{N}}\boldsymbol{m}\boldsymbol{v}^\top & \frac{1}{\sqrt{N}}\boldsymbol{G}^\top \\ \frac{1}{\sqrt{N}}\boldsymbol{v}\boldsymbol{m}^\top & \widetilde{\boldsymbol{D}}_2 & \boldsymbol{I}_M \\ \frac{1}{\sqrt{N}}\boldsymbol{G} & \boldsymbol{I}_M & \boldsymbol{0} \end{bmatrix} \in \mathbb{R}^{(N+2M)\times(N+2M)}. \tag{2.76}$$

Lemma 2.6.21. *We have* $|\det \nabla^2_\diamond \mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n})| = |\det \boldsymbol{X}|$.

Proof. Let $\boldsymbol{Y} = \begin{bmatrix} \widetilde{\boldsymbol{D}}_2 & \boldsymbol{I}_M \\ \boldsymbol{I}_M & \boldsymbol{0}\end{bmatrix}$. Note that $|\det \boldsymbol{Y}| = 1$ and $\boldsymbol{Y}^{-1} = \begin{bmatrix} \boldsymbol{0} & \boldsymbol{I}_M \\ \boldsymbol{I}_M & -\widetilde{\boldsymbol{D}}_2\end{bmatrix}$. By Schur's lemma,

$$|\det \boldsymbol{X}| = \left|\det\left(\boldsymbol{A} - \frac{1}{N}\begin{bmatrix}\boldsymbol{m}\boldsymbol{v}^\top & \boldsymbol{G}^\top\end{bmatrix}\boldsymbol{Y}^{-1}\begin{bmatrix}\boldsymbol{v}\boldsymbol{m}^\top \\ \boldsymbol{G}\end{bmatrix}\right)\right| = |\det \nabla^2_\diamond\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n})|. \qquad \square$$

It therefore suffices to study $|\det \boldsymbol{X}|$. This formulation has the benefit that the only randomness in $\boldsymbol{X}$ is from $\boldsymbol{G}$, and by Lemma 2.4.17 (in a suitable orthonormal basis) $\boldsymbol{G}$ is a matrix of independent (noncentered) gaussians. This structure will enable us to prove Lemma 2.4.9 using the spectral concentration results of [GZ00]. Before carrying out this argument, we first prove a preliminary lemma.

Lemma 2.6.22. *There exists $\tau > 0$ depending on $\varepsilon, C_{\mathsf{cvx}}$ such that, for all $(\boldsymbol{m},\boldsymbol{n}) \in \mathcal{S}_{\varepsilon,\upsilon}$, $\boldsymbol{X}$ has no eigenvalues in $[-\tau,\tau]$ with high probability under $\mathbb{P}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{Pl}}$.*

Proof. We will show that $\det(z\boldsymbol{I}_{N+2M} - \boldsymbol{X})$ has no zeros in $[-\tau,\tau]$. By Schur's lemma, for any $z \neq 0$,

$$|\det(z\boldsymbol{I}_{2M} - \boldsymbol{Y})| = |\det(z\boldsymbol{I}_M - \widetilde{\boldsymbol{D}}_2)||\det(z\boldsymbol{I}_M - (z\boldsymbol{I}_M - \widetilde{\boldsymbol{D}}_2)^{-1})| = |\det(z(z\boldsymbol{I}_M - \widetilde{\boldsymbol{D}}_2) - \boldsymbol{I}_M)|$$

Let τ_1 be the smallest positive solution to $\tau_1|\max(\widehat{f}_\varepsilon) + \tau| \le \frac{1}{2}$. Note that τ_1 depends only on ε, and the above determinant is nonzero for any $|z| \le \tau_1$. Further, note that

$$(z\boldsymbol{I}_{2M} - \boldsymbol{Y})^{-1} = \begin{bmatrix} -z(\boldsymbol{I}_M - z(z\boldsymbol{I}_M - \widetilde{\boldsymbol{D}}_2))^{-1} & (\boldsymbol{I}_M - z(z\boldsymbol{I}_M - \widetilde{\boldsymbol{D}}_2))^{-1} \\ (\boldsymbol{I}_M - z(z\boldsymbol{I}_M - \widetilde{\boldsymbol{D}}_2))^{-1} & -(z\boldsymbol{I}_M - \widetilde{\boldsymbol{D}}_2)(\boldsymbol{I}_M - z(z\boldsymbol{I}_M - \widetilde{\boldsymbol{D}}_2))^{-1}\end{bmatrix}.$$

From this, we see that there exists $C_\varepsilon > 0$ such that for all $|z| \le \tau_1$,

$$\|(z\boldsymbol{I}_{2M} - \boldsymbol{Y})^{-1} + \boldsymbol{Y}^{-1}\|_{\mathsf{op}} \le C_\varepsilon |z|.$$

By Schur's lemma, for all $|z| \le \tau_1$,

$$|\det(z\boldsymbol{I}_{N+2M} - \boldsymbol{X})| = |\det(z\boldsymbol{I}_{2M} - \boldsymbol{Y})||\det \boldsymbol{B}(z)|,$$

for

$$\boldsymbol{B}(z) = z\boldsymbol{I}_N - \boldsymbol{A} - \frac{1}{N}\begin{bmatrix}\boldsymbol{m}\boldsymbol{v}^\top & \boldsymbol{G}^\top\end{bmatrix}(z\boldsymbol{I}_{2M} - \boldsymbol{Y})^{-1}\begin{bmatrix}\boldsymbol{v}\boldsymbol{m}^\top \\ \boldsymbol{G}\end{bmatrix}.$$

It follows that for all $|z| \le \tau_1$,

$$\|\boldsymbol{B}(z) - \nabla^2_\diamond\mathcal{F}^\varepsilon_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n})\|_{\mathsf{op}} \le |z| + C_\varepsilon|z|\left(\frac{\|\boldsymbol{v}\boldsymbol{m}^\top\|_{\mathsf{op}}}{\sqrt{N}} + \frac{\|\boldsymbol{G}\|_{\mathsf{op}}}{\sqrt{N}}\right)^2.$$

As shown in Proposition 2.4.8(c), $\nabla^2_\diamond\mathcal{F}^\varepsilon_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n}) \preceq -C_{\mathsf{spec}}\boldsymbol{I}_N$ with high probability under $\mathbb{P}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{Pl}}$. Furthermore, $\frac{\|\boldsymbol{v}\boldsymbol{m}^\top\|_{\mathsf{op}}}{\sqrt{N}} = \frac{1}{\sqrt{N}}\|\boldsymbol{v}\|\|\boldsymbol{m}\|$ is bounded, with bound depending on $\varepsilon, C_{\mathsf{cvx}}$, and with high probability, $\frac{\|\boldsymbol{G}\|_{\mathsf{op}}}{\sqrt{N}}$ is bounded by an absolute constant. It follows that for $|z|$ small enough depending on $\varepsilon, C_{\mathsf{cvx}}$, $\boldsymbol{B}(z) \preceq -C_{\mathsf{spec}}\boldsymbol{I}_N/2$, and thus $|\det \boldsymbol{B}(z)| \neq 0$. $\square$

The core of the proof of Lemma 2.4.9 is the following spectral concentration inequality, which adapts [GZ00, Theorem 1.1(b)]. For any $f : \mathbb{R} \to \mathbb{R}$, let

$$\mathrm{tr} f(\boldsymbol{X}) = \sum_{i=1}^{N+2M} f(\lambda_i(\boldsymbol{X})),$$

where $\lambda_1(\boldsymbol{X}), \ldots, \lambda_{N+2M}(\boldsymbol{X})$ are the eigenvalues of $\boldsymbol{X}$.

Lemma 2.6.23. *If f is L-Lipschitz, then for any $t \ge 0$,*

$$\mathbb{P}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{PI}}(|\mathrm{tr} f(\boldsymbol{X}) - \mathbb{E}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{PI}} \mathrm{tr} f(\boldsymbol{X})| \ge t) \le 2e^{-t^2/8L^2}.$$

Proof. Let $\{\omega_{a,i} : a \in [M], i \in [N]\}$ be i.i.d. standard gaussians, and let $\dot{\boldsymbol{e}}_1, \ldots, \dot{\boldsymbol{e}}_N$ and $\widehat{\boldsymbol{e}}_1, \ldots, \widehat{\boldsymbol{e}}_M$ be orthonormal bases of $\mathbb{R}^N$ and $\mathbb{R}^M$ as in Lemma 2.4.17. By (2.40), we can sample $\widetilde{\boldsymbol{G}}$ by

$$\widetilde{\boldsymbol{G}} = \sum_{a=1}^{M} \sum_{i=1}^{N} w_{a,i} \omega_{a,i} \widehat{\boldsymbol{e}}_a \dot{\boldsymbol{e}}_i^\top, \qquad w_{a,i} = \begin{cases} \sqrt{\varepsilon/(q(\boldsymbol{m}) + \psi(\boldsymbol{n}) + \varepsilon)} & i = j = 1, \\ \sqrt{\varepsilon/(q(\boldsymbol{m}) + \varepsilon)} & i = 1, j \neq 1, \\ \sqrt{\varepsilon/(\psi(\boldsymbol{n}) + \varepsilon)} & i \neq 1, j = 1, \\ 1 & i \neq 1, j \neq 1. \end{cases}$$

By [GZ00, Lemma 1.2(b)], the map $\{\omega_{a,i} : a \in [M], i \in [N]\} \mapsto \mathrm{tr} f(\boldsymbol{X})$ is $2L$-Lipschitz. The result follows from the gaussian concentration inequality. $\square$

Proof of Lemma 2.4.9. Define $f(x) = \log \max(|x|, \tau)$, which is τ^{-1}-Lipschitz. Lemma 2.6.23 implies that

$$\mathbb{P}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{PI}}(|\mathrm{tr} f(\boldsymbol{X}) - \mathbb{E}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{PI}} \mathrm{tr} f(\boldsymbol{X})| \ge t) \le 2e^{-\tau^2 t^2/8}. \tag{2.77}$$

Let $\widetilde{\det}(\boldsymbol{X}) = \exp \mathrm{tr} f(\boldsymbol{X})$. Also let

$$\mathcal{E}_{\mathsf{spec}}(\boldsymbol{X}) = \{\mathsf{spec}(\boldsymbol{X}) \cap [-\tau, \tau] = \emptyset\},$$

so that $\mathbb{P}(\mathcal{E}_{\mathsf{spec}}) \ge 1 - \iota$ for some $\iota = o_N(1)$ by Lemma 2.6.22. Note that $|\det(\boldsymbol{X})| \le \widetilde{\det}(\boldsymbol{X})$ for all $\boldsymbol{X}$, with equality for all $\boldsymbol{X} \in \mathscr{E}_{\mathsf{spec}}$. Thus

$$\mathbb{E}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{PI}}[|\det(\boldsymbol{X})|^2] \le \mathbb{E}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{PI}}[\widetilde{\det}(\boldsymbol{X})^2], \qquad \mathbb{E}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{PI}}[|\det(\boldsymbol{X})|] \ge \mathbb{E}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{PI}}[\widetilde{\det}(\boldsymbol{X}) \mathbf{1}\{\mathcal{E}_{\mathsf{spec}}\}]. \tag{2.78}$$

By the concentration (2.77), there exists C depending on $\varepsilon, C_{\mathsf{cvx}}$ such that

$$\mathbb{E}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{PI}}[\widetilde{\det}(\boldsymbol{X})^2] \le C \exp(2\mathbb{E}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{PI}} \mathrm{tr} f(\boldsymbol{X})).$$

Furthermore, by Jensen's inequality $\mathbb{E}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{PI}}[\widetilde{\det}(\boldsymbol{X})] \ge \exp(\mathbb{E}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{PI}} \mathrm{tr} f(\boldsymbol{X}))$. Thus,

$$\mathbb{E}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{PI}}[\widetilde{\det}(\boldsymbol{X})^2] \le C\mathbb{E}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{PI}}[\widetilde{\det}(\boldsymbol{X})]^2. \tag{2.79}$$

By Cauchy–Schwarz,

$$\mathbb{E}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{PI}}[\widetilde{\det}(\boldsymbol{X}) \mathbf{1}\{\mathcal{E}^c_{\mathsf{spec}}\}] \le \mathbb{E}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{PI}}[\widetilde{\det}(\boldsymbol{X})^2]^{1/2} \mathbb{P}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{PI}}(\mathcal{E}^c_{\mathsf{spec}})^{1/2} \le C^{1/2} \iota^{1/2} \mathbb{E}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{PI}}[\widetilde{\det}(\boldsymbol{X})].$$

It follows that

$$\mathbb{E}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{PI}}[\widetilde{\det}(\boldsymbol{X}) \mathbf{1}\{\mathcal{E}_{\mathsf{spec}}\}] \ge (1 - C^{1/2}\iota^{1/2}) \mathbb{E}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{PI}}[\widetilde{\det}(\boldsymbol{X})].$$

Combining with (2.78), (2.79) shows that

$$\mathbb{E}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{PI}}[|\det(\boldsymbol{X})|^2]^{1/2} \le C^{1/2}(1 - C^{1/2}\iota^{1/2})^{-1} \mathbb{E}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{PI}}[|\det(\boldsymbol{X})|],$$

which implies the result after adjusting C. $\square$

2.7 First moment in planted model

In this section, we prove Proposition 2.3.9, bounding the first moment of $Z_N(\boldsymbol{G})$ in the planted model. The proof is structured as follows. In Subsection 2.7.1, we show this moment is bounded by a optimization problem over $\boldsymbol{\Lambda} : \mathbb{R} \to \mathbb{R}$ encoding subsets of Σ_N with a certain coordinate profile (heuristically described in (2.9)). Subsection 2.7.2 reduces this optimization to two dimensions by showing the maximizer is attained in a two-parameter family. For technical reasons, the functional in this optimization problem is not the $\mathscr{S}_\star$ defined in (2.8), but a variant $\mathscr{S}_\star^{s_{\max}}$ where s is minimized over $[0, s_{\max}]$ instead of $[0, +\infty)$ see (2.80). Subsection 2.7.3 and Subsection 2.7.4 show that we recover the optimization of $\mathscr{S}_\star$ when $s_{\max} \to \infty$, completing the proof of Proposition 2.3.9. Subsection 2.7.5 proves Lemma 2.2.5, on the local behavior of the first moment functional $\mathscr{S}_\star(\lambda_1, \lambda_2)$ near $(1, 0)$.

2.7.1 Reduction to functional optimization

Recall that (q_0, ψ_0) are given by Condition 2.3.1. Let $\dot{\boldsymbol{H}} \sim \mathcal{N}(0, \psi_0)$, $\boldsymbol{M} = \operatorname{th}(\dot{\boldsymbol{H}})$, and $\hat{\boldsymbol{H}} \sim \mathcal{N}(0, q_0)$, $\boldsymbol{N} = F_{1-q_0}(\hat{\boldsymbol{H}})$, for F_{1-q_0} given by (2.13). Let $\mathscr{L} = L^2(\mathbb{R}, \mathcal{N}(0, \psi_0))$ denote the space of measurable functions $\boldsymbol{\Lambda} : \mathbb{R} \to \mathbb{R}$, equipped with the inner product

$$\langle \boldsymbol{\Lambda}_1, \boldsymbol{\Lambda}_2 \rangle = \mathbb{E}[\boldsymbol{\Lambda}_1(\dot{\boldsymbol{H}})\boldsymbol{\Lambda}_2(\dot{\boldsymbol{H}})]$$

and square-integrable w.r.t. the associated norm. Let $\mathscr{K} \subseteq \mathscr{L}$ denote the set of functions with image in $[-1, 1]$. For $s_{\max} > 0$, define

$$\mathscr{S}_\star^{s_{\max}}(\boldsymbol{\Lambda}) = \inf_{0 \le s \le s_{\max}} \mathscr{S}_\star(\boldsymbol{\Lambda}, s), \tag{2.80}$$

where $\mathscr{S}_\star : \mathscr{K} \times [0, +\infty) \to \mathbb{R}$ is defined by (2.7). The following proposition bounds the first moment by the maximum of an optimization problem over functions $\boldsymbol{\Lambda}$, and is the starting point of the proof of Proposition 2.3.9.

Proposition 2.7.1. *For any $s_{\max} > 0$, $(\boldsymbol{m}, \boldsymbol{n}) \in \mathcal{S}_{\varepsilon,\upsilon}$, we have $\frac{1}{N} \log \mathbb{E}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{Pl}}[Z_N(\boldsymbol{G})] \le \sup_{\boldsymbol{\Lambda} \in \mathscr{K}} \mathscr{S}_\star^{s_{\max}}(\boldsymbol{\Lambda}) + o_{\varepsilon,\upsilon}(1)$.*

Here $o_{\varepsilon,\upsilon}(1)$ denotes a term vanishing as $\varepsilon, \upsilon \to 0$, which can depend on $s_{\max}$; we send $s_{\max} \to \infty$ after $\varepsilon, \upsilon \to 0$ in the end.

Before proving Proposition 2.7.1, we state a few facts that will be useful below. Lemma 2.7.2 ensures that the denominator of $\mathscr{S}_\star(\boldsymbol{\Lambda}, s)$ is well-behaved, while Lemmas 2.7.3 and 2.7.4 are useful in approximation arguments.

Lemma 2.7.2. *There exists $\iota > 0$ such that $\mathbb{E}[\boldsymbol{M}\boldsymbol{\Lambda}(\dot{\boldsymbol{H}})]^2 < (1 - \iota)q_0$ for all $\boldsymbol{\Lambda} \in \mathscr{K}$.*

Proof. Since $|\boldsymbol{\Lambda}(\dot{\boldsymbol{H}})| \le 1$, by Cauchy–Schwarz,

$$\mathbb{E}[\boldsymbol{M}\boldsymbol{\Lambda}(\dot{\boldsymbol{H}})]^2 \le \mathbb{E}[|\boldsymbol{M}|]^2 < \mathbb{E}[\boldsymbol{M}^2].$$

The inequality is strict because $|\boldsymbol{M}|$ has nonzero variance. Since $\mathbb{E}[\boldsymbol{M}^2] = P(\psi_0) = q_0$ (recall Condition 2.3.1), the result follows. $\square$

Lemma 2.7.3. *The function $\log \Psi(x)$ is $(2, 1)$-pseudo-Lipschitz (recall Definition 2.4.19).*

Proof. Note that $(\log \Psi)'(x) = -\mathcal{E}(x)$. Recall from Lemma 2.4.21(a) that $0 \le \mathcal{E}(x) \le 1 + |x|$. Thus,

$$|\log \Psi(x) - \log \Psi(y)| = \left| \int_x^y \mathcal{E}(s) \, \mathsf{d}s \right| \le |x - y|(1 + |x| + |y|).$$

$\square$

Lemma 2.7.4 (Proved in Appendix 2.A)**.** *There exists $C > 0$ such that for all $a_1, a_2, b_1, b_2, c_1, c_2 > 0$,*

$$\begin{aligned} &\left| \mathbb{E} \log \Psi \left\{ \frac{\kappa - a_1\hat{\boldsymbol{H}} - b_1\boldsymbol{N}}{c_1} \right\} - \log \Psi \left\{ \frac{\kappa - a_2\hat{\boldsymbol{H}} - b_2\boldsymbol{N}}{c_2} \right\} \right| \\ &\le \frac{C \max(a_1, a_2, b_1, b_2, c_1, c_2, 1)^3}{\min(c_1, c_2)^2} \left(|a_1 - a_2| + |b_1 - b_2| + |c_1 - c_2| \right). \end{aligned}$$

We turn to the proof of Proposition 2.7.1. The main step will be Proposition 2.7.5 below, where we show the bound in Proposition 2.7.1 holds for piecewise-constant $\mathbf{\Lambda}$ with finitely many parts. This case follows from a direct moment calculation, and Proposition 2.7.1 follows by approximation.

For any $\vec{r} = (r_1, \dots, r_{n-1})$ with $-\infty < r_1 < r_2 < \cdots < r_{n-1} < +\infty$, let $\mathscr{K}_{\mathrm{elt}}(\vec{r}) \subseteq \mathscr{K}$ denote the set of right-continuous functions which are constant on each interval $[r_{k-1}, r_k)$, $1 \le k \le n$. Here we take as convention $r_0 = -\infty$, $r_n = +\infty$. Define the quantiles $\vec{p} = (p_0, \dots, p_n)$ by $p_k = \mathbb{P}(\dot{\boldsymbol{H}} < r_k)$, and let

$$\mathsf{mesh}(\vec{p}) = \min_{1 \le k \le n} (p_k - p_{k-1}).$$

Let $o_{\varepsilon,\upsilon,\vec{p}}(1)$ denote a term vanishing as $\varepsilon, \upsilon, \mathsf{mesh}(\vec{p}) \to 0$, where (like before) this limit is taken after $N \to \infty$ for fixed $s_{\max}$. We will show the following.

Proposition 2.7.5. *Suppose $s_{\max} > 0$, $(\boldsymbol{m}, \boldsymbol{n}) \in \mathcal{S}_{\varepsilon,\upsilon}$, and $\vec{r} = (r_1, \dots, r_{n-1})$ is as above. We have that $\frac{1}{N} \log \mathbb{E}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{Pl}}[Z_N(\boldsymbol{G})] \le \sup_{\mathbf{\Lambda} \in \mathscr{K}_{\mathrm{elt}}(\vec{r})} \mathscr{S}^{s_{\max}}_{\star}(\mathbf{\Lambda}) + o_{\varepsilon,\upsilon,\vec{p}}(1)$.*

For the rest of this subsection, fix $s_{\max}, \varepsilon, \upsilon, \vec{r}$ and $(\boldsymbol{m}, \boldsymbol{n})$ as in Proposition 2.7.5. Let $\dot{\boldsymbol{h}} = \mathrm{th}_\varepsilon^{-1}(\boldsymbol{m})$ and $\widehat{\boldsymbol{h}} = F^{-1}_{\varepsilon,\varrho_\varepsilon}(\boldsymbol{n})$, so that $(\dot{\boldsymbol{h}}, \widehat{\boldsymbol{h}}) \in \mathcal{T}_{\varepsilon,\upsilon}$. Fix a partition $[N] = \mathcal{I}_1 \cup \cdots \cup \mathcal{I}_n$ satisfying

$$\begin{aligned} |\mathcal{I}_k| &= \lfloor p_k N \rfloor - \lfloor p_{k-1} N \rfloor, && \forall 1 \le k \le n, \\ \max\{\dot{h}_i : i \in \mathcal{I}_k\} &\le \min\{\dot{h}_i : i \in \mathcal{I}_{k+1}\}, && \forall 1 \le k \le n-1. \end{aligned}$$

(In words, $\mathcal{I}_k$ is the set of coordinates $i \in [N]$ such that the quantile of $\dot{h}_i$ among the entries of $\dot{\boldsymbol{h}}$, breaking ties in an arbitrary but fixed order, lies in $[p_{k-1}, p_k)$.) Then, partition Σ_N into sets

$$\Sigma_N(\vec{a}) = \left\{ \boldsymbol{x} \in \Sigma_N : \sum_{i \in \mathcal{I}_k} x_i = a_k, \ \forall 1 \le k \le n \right\}. \tag{2.81}$$

indexed by $\vec{a} = (a_1, \dots, a_n) \in \mathbb{Z}^n$. Let $\mathcal{J}$ be the set of $\vec{a}$ such that $\Sigma_N(\vec{a})$ is nonempty, and note that $|\mathcal{J}| \le N^n$. Thus

$$\begin{aligned} \frac{1}{N} \log \mathbb{E}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{Pl}}[Z_N(\boldsymbol{G})] &= \frac{1}{N} \log \sum_{\vec{a} \in \mathcal{J}} \sum_{\boldsymbol{x} \in \Sigma_N(\vec{a})} \mathbb{P}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{Pl}} \left(\frac{\boldsymbol{G}\boldsymbol{x}}{\sqrt{N}} \ge \kappa \right) \\ &= \sup_{\vec{a} \in \mathcal{J}} \left\{ \frac{1}{N} \log |\Sigma_N(\vec{a})| + \sup_{\boldsymbol{x} \in \Sigma_N(\vec{a})} \frac{1}{N} \log \mathbb{P}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{Pl}} \left(\frac{\boldsymbol{G}\boldsymbol{x}}{\sqrt{N}} \ge \kappa \right) \right\} + o_N(1). \end{aligned} \tag{2.82}$$

Associate to each $\vec{a} \in \mathcal{J}$ a function $\mathbf{\Lambda}^{\vec{a}} \in \mathscr{K}_{\mathrm{elt}}(r_1, \dots, r_{n-1})$ defined by

$$\mathbf{\Lambda}^{\vec{a}}(x) = \frac{a_k}{|\mathcal{I}_k|}, \qquad x \in [r_{k-1}, r_k), 1 \le k \le n.$$

Recall the function $\mathsf{ent} : \mathscr{K} \to \mathbb{R}$ defined in (2.6).

Lemma 2.7.6. *We have $\frac{1}{N} \log |\Sigma_N(\vec{a})| = \mathsf{ent}(\mathbf{\Lambda}^{\vec{a}}) + o_N(1)$ for an error $o_N(1)$ uniform over $\vec{a} \in \mathcal{J}$.*

Proof. By direct counting,

$$|\Sigma_N(\vec{a})| = \prod_{k=1}^n \binom{|\mathcal{I}_k|}{\frac{1}{2}(|\mathcal{I}_k| + a_k)}.$$

Stirling's approximation yields

$$\frac{1}{N} \log |\Sigma_N(\vec{a})| = \sum_{k=1}^n \left\{ (p_k - p_{k-1}) \mathcal{H} \left(\frac{1 + \frac{a_k}{(p_k - p_{k-1})N}}{2} \right) \right\} + o_N(1) = \mathbb{E}\, \mathcal{H} \left(\frac{1 + \mathbf{\Lambda}^{\vec{a}}(\dot{\boldsymbol{H}})}{2} \right) + o_N(1),$$

where the last equality holds because $\mathbb{P}(\dot{\boldsymbol{H}} \in [r_{k-1}, r_k)) = p_k - p_{k-1}$. $\square$

Lemma 2.7.7. *For all $\vec{a} \in \mathcal{J}$ and $\boldsymbol{x} \in \Sigma_N(\vec{a})$,*

$$\frac{1}{N}\langle \dot{\boldsymbol{h}}, \boldsymbol{x}\rangle = \mathbb{E}[\dot{\boldsymbol{H}}\boldsymbol{\Lambda}^{\vec{a}}(\dot{\boldsymbol{H}})] + o_{\varepsilon,\upsilon,\vec{p}}(1), \qquad\qquad \frac{1}{N}\langle \boldsymbol{m}, \boldsymbol{x}\rangle = \mathbb{E}[\boldsymbol{M}\boldsymbol{\Lambda}^{\vec{a}}(\dot{\boldsymbol{H}})] + o_{\varepsilon,\upsilon,\vec{p}}(1),$$

for error terms $o_{\varepsilon,\upsilon,\vec{p}}(1)$ uniform over $\vec{a}, \boldsymbol{x}$.

Proof. We will only show the proof for $\frac{1}{N}\langle \dot{\boldsymbol{h}}, \boldsymbol{x}\rangle$, as the other estimate is analogous. Let $\boldsymbol{x} \in \Sigma_N(\vec{a})$ be fixed, and let $\boldsymbol{y} \in [-1,1]^N$ be defined by $y_i = \frac{a_k}{|\mathcal{I}_k|}$ for all $i \in \mathcal{I}_k$. We write $(\dot{\boldsymbol{H}}', \boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{K})$ for the random variable with value $(\dot{h}_i, x_i, y_i, k)$, where $i \sim \mathsf{unif}([N])$ and $k \in [n]$ is the index of the set $\mathcal{I}_k$ containing i. Recall that $\dot{\boldsymbol{H}} \sim \mathcal{N}(0, \psi_0)$. Note that

$$\mathbb{W}_2(\mathcal{L}(\dot{\boldsymbol{H}}'), \mathcal{L}(\dot{\boldsymbol{H}})) \le \mathbb{W}_2(\mu_{\dot{\boldsymbol{h}}}, \mathcal{N}(0, \psi_\varepsilon + \varepsilon)) + \mathbb{W}_2(\mathcal{N}(0, \psi_\varepsilon + \varepsilon), \mathcal{N}(0, \psi_0)) = o_{\varepsilon,\upsilon}(1),$$

where the latter two distances are bounded by definition of $\mathcal{T}_\upsilon$ and Proposition 2.4.1, respectively. We couple $(\dot{\boldsymbol{H}}', \dot{\boldsymbol{H}})$ monotonically (which is the $\mathbb{W}_2$-optimal coupling) and write

$$\frac{1}{N}\langle \dot{\boldsymbol{h}}, \boldsymbol{x}\rangle = \mathbb{E}[\dot{\boldsymbol{H}}'\boldsymbol{X}] = \mathbb{E}[\dot{\boldsymbol{H}}\boldsymbol{Y}] + \mathbb{E}[(\dot{\boldsymbol{H}}' - \dot{\boldsymbol{H}})\boldsymbol{X}] + \mathbb{E}[\dot{\boldsymbol{H}}(\boldsymbol{X} - \boldsymbol{Y})].$$

We now estimate each of these terms. Because $(\dot{\boldsymbol{H}}', \dot{\boldsymbol{H}})$ are coupled monotonically, $\boldsymbol{K} = k$ if and only if the quantile of $\dot{\boldsymbol{H}}$ lies in $[p'_{k-1}, p'_k)$, where $p'_k = \frac{1}{N}\lfloor p_k N\rfloor = p_k + O(N^{-1})$. Thus, on an event with probability $1 - O(N^{-1})$, $\boldsymbol{K} = k$ if and only if $\dot{\boldsymbol{H}} \in [r_{k-1}, r_k)$. On this event, $\boldsymbol{Y} = \boldsymbol{\Lambda}^{\vec{a}}(\dot{\boldsymbol{H}})$. Thus

$$\mathbb{E}[\dot{\boldsymbol{H}}\boldsymbol{Y}] = \mathbb{E}[\dot{\boldsymbol{H}}\boldsymbol{\Lambda}^{\vec{a}}(\dot{\boldsymbol{H}})] + o_N(1).$$

Moreover,

$$|\,\mathbb{E}[(\dot{\boldsymbol{H}}' - \dot{\boldsymbol{H}})\boldsymbol{X}]| \le \mathbb{E}[(\dot{\boldsymbol{H}}' - \dot{\boldsymbol{H}})^2]^{1/2} = \mathbb{W}_2(\mathcal{L}(\dot{\boldsymbol{H}}'), \mathcal{L}(\dot{\boldsymbol{H}})) = o_{\varepsilon,\upsilon}(1).$$

Finally, note that $\boldsymbol{Y} = \mathbb{E}[\boldsymbol{X}|\boldsymbol{K}]$, so

$$\mathbb{E}[\mathbb{E}[\dot{\boldsymbol{H}}|\boldsymbol{K}](\boldsymbol{X} - \boldsymbol{Y})] = \mathbb{E}[\mathbb{E}[\dot{\boldsymbol{H}}|\boldsymbol{K}]\,\mathbb{E}[\boldsymbol{X} - \boldsymbol{Y}|\boldsymbol{K}]] = 0.$$

Thus

$$|\,\mathbb{E}[\dot{\boldsymbol{H}}(\boldsymbol{X} - \boldsymbol{Y})]| = |\,\mathbb{E}[(\dot{\boldsymbol{H}} - \mathbb{E}[\dot{\boldsymbol{H}}|\boldsymbol{K}])(\boldsymbol{X} - \boldsymbol{Y})]| \le \mathbb{E}[(\dot{\boldsymbol{H}} - \mathbb{E}[\dot{\boldsymbol{H}}|\boldsymbol{K}])^2]^{1/2}.$$

Recall from the above discussion that conditioning on $\boldsymbol{K}$ reveals the interval $[p'_{k-1}, p'_k)$ containing the quantile of $\dot{\boldsymbol{H}}$. It follows that $\mathbb{E}[(\dot{\boldsymbol{H}} - \mathbb{E}[\dot{\boldsymbol{H}}|\boldsymbol{K}])^2] = o_{\varepsilon,\upsilon,\vec{p}}(1)$. $\square$

Lemma 2.7.8. *For all $\vec{a} \in \mathcal{J}$, $\boldsymbol{x} \in \Sigma_N(\vec{a})$, and $s \in [0, s_{\max}]$,*

$$\frac{1}{N}\log \mathbb{P}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{Pl}}\left(\frac{\boldsymbol{G}\boldsymbol{x}}{\sqrt{N}} \ge \kappa\right) \le \frac{1}{2}s^2\psi_0 + \alpha_\star\,\mathbb{E}\log\Psi\left\{\frac{\kappa - \frac{\mathbb{E}[\boldsymbol{M}\boldsymbol{\Lambda}^{\vec{a}}(\dot{\boldsymbol{H}})]}{q_0}\hat{\boldsymbol{H}} - \frac{\mathbb{E}[\dot{\boldsymbol{H}}\boldsymbol{\Lambda}^{\vec{a}}(\dot{\boldsymbol{H}})]}{\psi_0}\boldsymbol{N}}{\sqrt{1 - \frac{\mathbb{E}[\boldsymbol{M}\boldsymbol{\Lambda}^{\vec{a}}(\dot{\boldsymbol{H}})]^2}{q_0}}} + s\boldsymbol{N}\right\} + o_{\varepsilon,\upsilon,\vec{p}}(1),$$

where the $o_{\varepsilon,\upsilon,\vec{p}}(1)$ is uniform over $\vec{a}, \boldsymbol{x}, s$ (but can depend on $s_{\max}$).

Proof. Let $\widetilde{\boldsymbol{G}}$ be defined in Corollary 2.4.18. By Corollary 2.4.18 and Lemma 2.7.7,

$$\begin{aligned}\frac{\boldsymbol{G}\boldsymbol{x}}{\sqrt{N}} &\overset{d}{=} \left(\frac{(1 + o_{\varepsilon,\upsilon}(1))}{q_0}\widehat{\boldsymbol{h}} + o_{\varepsilon,\upsilon}(1)\boldsymbol{n}\right)\frac{1}{N}\langle \boldsymbol{m}, \boldsymbol{x}\rangle + \frac{(1 + o_{\varepsilon,\upsilon}(1))}{\psi_0}\boldsymbol{n}\cdot\frac{1}{N}\langle \dot{\boldsymbol{h}}, \boldsymbol{x}\rangle + \frac{\widetilde{\boldsymbol{G}}\boldsymbol{x}}{\sqrt{N}} \\ &= \frac{\mathbb{E}[\boldsymbol{M}\boldsymbol{\Lambda}^{\vec{a}}(\dot{\boldsymbol{H}})] + o_{\varepsilon,\upsilon,\vec{p}}(1)}{q_0}\widehat{\boldsymbol{h}} + \frac{\mathbb{E}[\dot{\boldsymbol{H}}\boldsymbol{\Lambda}^{\vec{a}}(\dot{\boldsymbol{H}})] + o_{\varepsilon,\upsilon,\vec{p}}(1)}{\psi_0}\boldsymbol{n} + \frac{\widetilde{\boldsymbol{G}}\boldsymbol{x}}{\sqrt{N}}.\end{aligned}$$

Let $\widehat{\boldsymbol{n}} = \boldsymbol{n}/\|\boldsymbol{n}\|$. By inspecting (2.40), we see that for independent $\widetilde{\boldsymbol{g}} \sim \mathcal{N}(0, P^\perp_{\boldsymbol{n}})$ and $Z \sim \mathcal{N}(0,1)$,

$$\frac{\widetilde{\boldsymbol{G}}\boldsymbol{x}}{\sqrt{N}} \overset{d}{=} \left(\frac{\|P^\perp_{\boldsymbol{m}}(\boldsymbol{x})\|^2}{N} + o_\varepsilon(1)\right)^{1/2}\widetilde{\boldsymbol{g}} + o_\varepsilon(1)Z\widehat{\boldsymbol{n}} = t^{1/2}\widetilde{\boldsymbol{g}} + \iota_1^{1/2}Z\widehat{\boldsymbol{n}},$$

where $t = 1 - \frac{\mathbb{E}[\boldsymbol{M}\boldsymbol{\Lambda}^{\vec{a}}(\dot{\boldsymbol{H}})]^2}{q_0} + \iota_2$ and $\iota_1, \iota_2 = o_{\varepsilon,\upsilon,\vec{p}}(1)$. For $Z' \sim \mathcal{N}(0,1)$ independent of $\widetilde{\boldsymbol{g}}, Z$, let

$$\widehat{\boldsymbol{g}} = \widetilde{\boldsymbol{g}} + Z'\widehat{\boldsymbol{n}} + s\boldsymbol{n}$$

so that $\widehat{\boldsymbol{g}} \sim \mathcal{N}(s\boldsymbol{n}, \boldsymbol{I}_N)$. Then, for any measurable $S \subseteq \mathbb{R}^N$,

$$\begin{aligned}
\frac{\mathbb{P}(t^{1/2}\widetilde{\boldsymbol{g}} + \iota_1^{1/2} Z\widehat{\boldsymbol{n}} \in S)}{\mathbb{P}(t^{1/2}\widehat{\boldsymbol{g}} \in S)} &\le \sup_{T\subseteq\mathbb{R}} \frac{\mathbb{P}(\iota_1^{1/2} Z \in T)}{\mathbb{P}(st^{1/2}\|\boldsymbol{n}\| + t^{1/2}Z' \in T)} \\
&\le \sup_{x\in\mathbb{R}} \frac{\iota_1^{-1/2}\exp(-\frac{1}{2\iota_1}x^2)}{t^{-1/2}\exp(-\frac{1}{2t}(x - st^{1/2}\|\boldsymbol{n}\|)^2)} = \sqrt{\frac{t}{\iota_1}} \exp\left(\frac{s^2\|\boldsymbol{n}\|^2}{2(1-\iota_1/t)}\right).
\end{aligned}$$

Thus,

$$\begin{aligned}
&\frac{1}{N}\log \mathbb{P}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{PI}}\left(\frac{\boldsymbol{G}\boldsymbol{x}}{\sqrt{N}} \ge \kappa\right) \le \frac{s^2\psi(\boldsymbol{n})}{2(1-\iota_1/t)} + o_N(1) \\
&\quad + \frac{1}{N}\log \mathbb{P}\left\{\frac{\mathbb{E}[\boldsymbol{M}\boldsymbol{\Lambda}^{\vec{a}}(\dot{\boldsymbol{H}})] + o_{\varepsilon,\upsilon,\vec{p}}(1)}{q_0}\widehat{\boldsymbol{h}} + \frac{\mathbb{E}[\dot{\boldsymbol{H}}\boldsymbol{\Lambda}^{\vec{a}}(\dot{\boldsymbol{H}})] + o_{\varepsilon,\upsilon,\vec{p}}(1)}{\psi_0}\boldsymbol{n} + t^{1/2}\widehat{\boldsymbol{g}} \ge \kappa\right\}. \qquad (2.83)
\end{aligned}$$

By Lemma 2.7.2, t is bounded away from 0. Since $\psi(\boldsymbol{n}) = \psi_0 + o_\varepsilon(1)$, we have

$$\frac{s^2\psi(\boldsymbol{n})}{2(1-\iota_1/t)} = (1 + o_{\varepsilon,\upsilon,\vec{p}}(1))\frac{1}{2}s^2\psi_0 = \frac{1}{2}s^2\psi_0 + o_{\varepsilon,\upsilon,\vec{p}}(1).$$

The last estimate holds uniformly over $s \in [0, s_{\max}]$. The last term of (2.83) equals

$$\frac{1}{N}\sum_{a=1}^{M} \log\Psi\left\{\frac{\kappa - \frac{\mathbb{E}[\boldsymbol{M}\boldsymbol{\Lambda}^{\vec{a}}(\dot{\boldsymbol{H}})] + o_{\varepsilon,\upsilon,\vec{p}}(1)}{q_0}\widehat{h}_a - \frac{\mathbb{E}[\dot{\boldsymbol{H}}\boldsymbol{\Lambda}^{\vec{a}}(\dot{\boldsymbol{H}})] + o_{\varepsilon,\upsilon,\vec{p}}(1)}{\psi_0}n_a}{\sqrt{1 - \frac{\mathbb{E}[\boldsymbol{M}\boldsymbol{\Lambda}^{\vec{a}}(\dot{\boldsymbol{H}})]^2}{q_0} + o_{\varepsilon,\upsilon,\vec{p}}(1)}} + sn_a\right\} + o_N(1).$$

By Lemma 2.7.3, $\log\Psi$ is $(2,1)$-pseudo-Lipschitz. By Fact 2.4.20 and Lemma 2.7.4 (using again that the denominator is bounded away from 0), the last display equals

$$\alpha_\star \mathbb{E}\log\Psi\left\{\frac{\kappa - \frac{\mathbb{E}[\boldsymbol{M}\boldsymbol{\Lambda}^{\vec{a}}(\dot{\boldsymbol{H}})]}{q_0}\hat{\boldsymbol{H}} - \frac{\mathbb{E}[\dot{\boldsymbol{H}}\boldsymbol{\Lambda}^{\vec{a}}(\dot{\boldsymbol{H}})]}{\psi_0}\boldsymbol{N}}{\sqrt{1 - \frac{\mathbb{E}[\boldsymbol{M}\boldsymbol{\Lambda}^{\vec{a}}(\dot{\boldsymbol{H}})]^2}{q_0}}} + s\boldsymbol{N}\right\} + o_{\varepsilon,\upsilon,\vec{p}}(1).$$

Combining the above concludes the proof. □

Proof of Proposition 2.7.5. Follows from equation (2.82) and Lemmas 2.7.6 and 2.7.8. □

Proof of Proposition 2.7.1. Set $\vec{r}$ such that $\mathsf{mesh}(\vec{p})$ is suitably small depending on (ε, υ). Then

$$\frac{1}{N}\log \mathbb{E}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{PI}}[Z_N(\boldsymbol{G})] \le \sup_{\boldsymbol{\Lambda}\in\mathscr{K}_{\mathrm{elt}}(\vec{r})} \mathscr{S}_\star^{s_{\max}}(\boldsymbol{\Lambda}) + o_{\varepsilon,\upsilon}(1) \le \sup_{\boldsymbol{\Lambda}\in\mathscr{K}} \mathscr{S}_\star^{s_{\max}}(\boldsymbol{\Lambda}) + o_{\varepsilon,\upsilon}(1).$$

□

2.7.2 Reduction to two parameters

Let $\mathscr{K}_* \subseteq \mathscr{K}$ denote the set of functions of the form $\boldsymbol{\Lambda}_{\lambda_1,\lambda_2}$ defined above (2.8). Let $\overline{\mathscr{K}}_*$ denote the closure of this set in the topology of $\mathscr{L}$. We next prove the following, which reduces the functional optimization problem in Proposition 2.7.1 to an optimization over $\overline{\mathscr{K}}_*$.

Proposition 2.7.9. *For any $s_{\max} > 0$, we have $\sup_{\boldsymbol{\Lambda}\in\mathscr{K}} \mathscr{S}_\star^{s_{\max}}(\boldsymbol{\Lambda}) = \sup_{\boldsymbol{\Lambda}\in\overline{\mathscr{K}}_*} \mathscr{S}_\star^{s_{\max}}(\boldsymbol{\Lambda})$. Similarly, $\sup_{\boldsymbol{\Lambda}\in\mathscr{K}} \mathscr{S}_\star(\boldsymbol{\Lambda}) = \sup_{\boldsymbol{\Lambda}\in\overline{\mathscr{K}}_*} \mathscr{S}_\star(\boldsymbol{\Lambda})$ for $\mathscr{S}_\star(\boldsymbol{\Lambda})$ defined in (2.8).*

Lemma 2.7.10. *Let $a_1, a_2 \in \mathbb{R}$ be such that there exists $\mathbf{\Lambda} \in \mathscr{K}$ with $\mathbb{E}[\dot{\boldsymbol{H}}\mathbf{\Lambda}(\dot{\boldsymbol{H}})] = a_1$, $\mathbb{E}[\boldsymbol{M}\mathbf{\Lambda}(\dot{\boldsymbol{H}})] = a_2$. Then, the concave optimization problem*

$$\textit{maximize } \mathsf{ent}(\mathbf{\Lambda}) \qquad \textit{subject to} \quad \mathbf{\Lambda} \in \mathscr{K}, \quad \mathbb{E}[\dot{\boldsymbol{H}}\mathbf{\Lambda}(\dot{\boldsymbol{H}}))] = a_1, \quad \mathbb{E}[\boldsymbol{M}\mathbf{\Lambda}(\dot{\boldsymbol{H}}))] = a_2$$

has a maximizer in $\overline{\mathscr{K}}_*$.

Proof. Introduce Lagrange multipliers $\lambda_1, \lambda_2 \in \mathbb{R}$. The Lagrangian is

$$L(\mathbf{\Lambda}; \lambda_1, \lambda_2) = \mathbb{E}\left\{\mathcal{H}\left(\frac{1+\mathbf{\Lambda}(\dot{\boldsymbol{H}})}{2}\right) + \lambda_1\dot{\boldsymbol{H}}\mathbf{\Lambda}(\dot{\boldsymbol{H}}) + \lambda_2\boldsymbol{M}\mathbf{\Lambda}(\dot{\boldsymbol{H}})\right\} - \lambda_1 a_1 - \lambda_2 a_2.$$

The quantity inside the expectation is concave in $\mathbf{\Lambda}(\dot{\boldsymbol{H}})$, with derivative

$$-\mathrm{th}^{-1}(\mathbf{\Lambda}(\dot{\boldsymbol{H}})) + \lambda_1\dot{\boldsymbol{H}} + \lambda_2\boldsymbol{M}.$$

This is pointwise maximized by $\mathbf{\Lambda}(\dot{\boldsymbol{H}}) = \mathrm{th}(\lambda_1\dot{\boldsymbol{H}} + \lambda_2\boldsymbol{M})$, i.e. $\mathbf{\Lambda} = \mathbf{\Lambda}_{\lambda_1,\lambda_2}$. □

Proof of Proposition 2.7.9. Note that $\mathscr{S}_\star^{s_{\max}}(\mathbf{\Lambda})$ is the sum of $\mathsf{ent}(\mathbf{\Lambda})$ and a term depending on $\mathbf{\Lambda}$ only through $\mathbb{E}[\dot{\boldsymbol{H}}\mathbf{\Lambda}(\dot{\boldsymbol{H}})]$ and $\mathbb{E}[\boldsymbol{M}\mathbf{\Lambda}(\dot{\boldsymbol{H}})]$. Let $\mathbf{\Lambda} \in \mathscr{K}$ be arbitrary. By Lemma 2.7.10, the maximum of $\mathsf{ent}(\widetilde{\mathbf{\Lambda}})$ subject to $\widetilde{\mathbf{\Lambda}} \in \mathscr{K}$, $\mathbb{E}[\dot{\boldsymbol{H}}\widetilde{\mathbf{\Lambda}}(\dot{\boldsymbol{H}})] = \mathbb{E}[\dot{\boldsymbol{H}}\mathbf{\Lambda}(\dot{\boldsymbol{H}})]$, $\mathbb{E}[\boldsymbol{M}\widetilde{\mathbf{\Lambda}}(\dot{\boldsymbol{H}})] = \mathbb{E}[\boldsymbol{M}\mathbf{\Lambda}(\dot{\boldsymbol{H}})]$ is attained by some $\widetilde{\mathbf{\Lambda}} \in \overline{\mathscr{K}}_*$. Thus $\mathscr{S}_\star^{s_{\max}}(\mathbf{\Lambda}) \le \mathscr{S}_\star^{s_{\max}}(\widetilde{\mathbf{\Lambda}})$, which implies the conclusion for $\mathscr{S}^{s_{\max}}$. The proof for $\mathscr{S}_\star$ is identical. □

2.7.3 The $s_{\max} \to \infty$ limit

In this subsection, we prove the following proposition, which shows that the optimization problem derived in Proposition 2.7.9 has a well-behaved limit when we take $s_{\max} \to \infty$. This allows us to remove the parameter $s_{\max}$, replacing the constrained optimization $\mathscr{S}_\star^{s_{\max}}$ defined in (2.80) with the $\mathscr{S}_\star$ defined in (2.8).

Proposition 2.7.11. *We have $\lim_{s_{\max}\to\infty} \sup_{\mathbf{\Lambda}\in\overline{\mathscr{K}}_*} \mathscr{S}_\star^{s_{\max}}(\mathbf{\Lambda}) = \sup_{\mathbf{\Lambda}\in\overline{\mathscr{K}}_*} \mathscr{S}_\star(\mathbf{\Lambda})$, and moreover $\mathscr{S}_\star$ attains its supremum on $\overline{\mathscr{K}}_*$.*

Lemma 2.7.12. *The function $\mathscr{S}_\star : \mathscr{K} \times \mathbb{R} \to \mathbb{R}$ (recall (2.7)) is continuous.*

Proof. Note that $s \mapsto \frac{1}{2}s^2\psi_0$ is manifestly continuous. By concavity of $\mathcal{H}$, $|\mathcal{H}(x) - \mathcal{H}(y)| \le \mathcal{H}(|x-y|)$ for all $x, y \in [0,1]$. By concavity of $x \mapsto \mathcal{H}(\sqrt{x}/2)$ and Jensen's inequality,

$$\begin{aligned}|\mathsf{ent}(\mathbf{\Lambda}) - \mathsf{ent}(\mathbf{\Lambda}')| &\le \mathbb{E}\left|\mathcal{H}\left(\frac{1+\mathbf{\Lambda}(\dot{\boldsymbol{H}})}{2}\right) - \mathcal{H}\left(\frac{1+\mathbf{\Lambda}'(\dot{\boldsymbol{H}})}{2}\right)\right| \le \mathbb{E}\left|\mathcal{H}\left(\frac{|\mathbf{\Lambda}(\dot{\boldsymbol{H}}) - \mathbf{\Lambda}'(\dot{\boldsymbol{H}})|}{2}\right)\right| \\ &\le \mathcal{H}\left(\frac{\mathbb{E}[|\mathbf{\Lambda}(\dot{\boldsymbol{H}}) - \mathbf{\Lambda}'(\dot{\boldsymbol{H}})|^2]^{1/2}}{2}\right) = \mathcal{H}\left(\frac{\|\mathbf{\Lambda} - \mathbf{\Lambda}'\|}{2}\right).\end{aligned}$$

Thus ent is continuous. By Cauchy–Schwarz,

$$|\mathbb{E}[\dot{\boldsymbol{H}}\mathbf{\Lambda}] - \mathbb{E}[\dot{\boldsymbol{H}}\mathbf{\Lambda}']| \le \mathbb{E}[\dot{\boldsymbol{H}}^2]^{1/2}\|\mathbf{\Lambda} - \mathbf{\Lambda}'\| = \psi_0^{1/2}\|\mathbf{\Lambda} - \mathbf{\Lambda}'\|$$

and similarly $|\mathbb{E}[\boldsymbol{M}\mathbf{\Lambda}] - \mathbb{E}[\boldsymbol{M}\mathbf{\Lambda}']| \le q_0^{1/2}\|\mathbf{\Lambda} - \mathbf{\Lambda}'\|$. Since the denominator $1 - \frac{\mathbb{E}[\boldsymbol{M}\mathbf{\Lambda}(\dot{\boldsymbol{H}})]^2}{q_0}$ is bounded away from 0 by Lemma 2.7.2, the final term of $\mathscr{S}_\star$ is continuous by Lemma 2.7.4. Thus $\mathscr{S}_\star$ is continous. □

We will need the following analytical lemma, which is a simple adaptation of Dini's Theorem [Rud76, Theorem 7.13]. We provide a proof for completeness.

Lemma 2.7.13. *Suppose $f_1, f_2, \ldots : K \to \mathbb{R}$ are a decreasing sequence of continuous functions on a compact space K. Let $f : K \to \mathbb{R}\cup\{-\infty\}$ denote their (not necessarily continuous) pointwise limit, which we assume is not $-\infty$ everywhere. Then $\lim_{n\to\infty} \sup f_n = \sup f$, and furthermore f attains its supremum.*

Proof. Without loss of generality assume $\sup f = 0$. For $\iota > 0$, let $E_n = \{x \in K : f_n(x) < \iota\}$. Then E_n is open and $E_n \subseteq E_{n+1}$. Since the f_n converge pointwise to f, $\cup_n E_n = K$. By compactness of K, $E_n = K$ for some finite n, and thus $\sup f_n < \iota$. As this holds for any ι, $\lim_{n\to\infty} \sup f_n = 0$. Finally, f, as the decreasing limit of (upper-semi)continuous functions, is upper-semicontinuous. Therefore f attains its supremum. $\square$

To apply Lemma 2.7.13, we verify that $\mathscr{S}_\star$ is not $-\infty$ everywhere by calculating its value at $\boldsymbol{\Lambda}_{1,0}(x) = \mathrm{th}(x)$ in Lemma 2.7.15 below. Recalling Subsection 2.2.6, we expect this to be the maximizer of $\mathscr{S}_\star$.

Lemma 2.7.14. *For any* $\boldsymbol{\Lambda} \in \mathscr{K}$, $s \geq 0$, *we have* $\frac{\partial^2}{\partial s^2}\mathscr{S}_\star(\boldsymbol{\Lambda}, s) > 0$.

Proof. Since $(\log \Psi)' = -\mathcal{E}$, we have

$$\begin{aligned}\frac{\partial^2}{\partial s^2}\mathscr{S}_\star(\boldsymbol{\Lambda}, s) &= \psi_0 - \alpha_\star \mathbb{E}\left\{\mathcal{E}'\left(\frac{\kappa - \frac{\mathbb{E}[\boldsymbol{M}\boldsymbol{\Lambda}(\dot{\boldsymbol{H}})]}{q_0}\hat{\boldsymbol{H}} - \frac{\mathbb{E}[\dot{\boldsymbol{H}}\boldsymbol{\Lambda}(\dot{\boldsymbol{H}})]}{\psi_0}\boldsymbol{N}}{\sqrt{1 - \frac{\mathbb{E}[\boldsymbol{M}\boldsymbol{\Lambda}(\dot{\boldsymbol{H}})]^2}{q_0}}} + s\boldsymbol{N}\right)\boldsymbol{N}^2\right\} \\ &\overset{Lem.\ 2.4.21(b)}{>} \psi_0 - \alpha_\star \mathbb{E}[\boldsymbol{N}^2] = 0.\end{aligned}$$

$\square$

Lemma 2.7.15. *We have* $\mathscr{S}_\star(\boldsymbol{\Lambda}_{1,0}) = \mathscr{S}_\star(\boldsymbol{\Lambda}_{1,0}, \sqrt{1-q_0}) = 0$.

Proof. Let $\boldsymbol{\Lambda} = \boldsymbol{\Lambda}_{1,0}$. Note that $\boldsymbol{\Lambda}(\dot{\boldsymbol{H}}) = \mathrm{th}(\dot{\boldsymbol{H}}) = \boldsymbol{M}$. Thus $\mathbb{E}[\boldsymbol{M}\boldsymbol{\Lambda}(\dot{\boldsymbol{H}})] = q_0$ and, by gaussian integration by parts, $\mathbb{E}[\dot{\boldsymbol{H}}\boldsymbol{\Lambda}(\dot{\boldsymbol{H}})] = (1-q_0)\psi_0$. So

$$\frac{\kappa - \frac{\mathbb{E}[\boldsymbol{M}\boldsymbol{\Lambda}(\dot{\boldsymbol{H}})]}{q_0}\hat{\boldsymbol{H}} - \frac{\mathbb{E}[\dot{\boldsymbol{H}}\boldsymbol{\Lambda}(\dot{\boldsymbol{H}})]}{\psi_0}\boldsymbol{N}}{\sqrt{1 - \frac{\mathbb{E}[\boldsymbol{M}\boldsymbol{\Lambda}(\dot{\boldsymbol{H}})]^2}{q_0}}} + \sqrt{1-q_0}\boldsymbol{N} = \frac{\kappa - \hat{\boldsymbol{H}}}{\sqrt{1-q_0}}.$$

By the identity $\mathcal{H}(\frac{1+\mathrm{th}x}{2}) = \log(2\mathrm{ch}x) - x\mathrm{th}x$,

$$\mathbb{E}\,\mathcal{H}\left(\frac{1+\boldsymbol{\Lambda}}{2}\right) = \mathbb{E}\log(2\mathrm{ch}\dot{\boldsymbol{H}}) - \mathbb{E}[\dot{\boldsymbol{H}}\boldsymbol{\Lambda}] = \mathbb{E}\log(2\mathrm{ch}\dot{\boldsymbol{H}}) - (1-q_0)\psi_0.$$

Thus

$$\mathscr{S}_\star(\boldsymbol{\Lambda}, \sqrt{1-q_0}) = -\frac{1}{2}(1-q_0)\psi_0 + \mathbb{E}\log(2\mathrm{ch}\dot{\boldsymbol{H}}) + \alpha\,\mathbb{E}\log\Psi\left(\frac{\kappa - \hat{\boldsymbol{H}}}{\sqrt{1-q_0}}\right) = \mathscr{G}(\alpha_\star, q_0, \psi_0),$$

which equals 0 by definition of $\alpha_\star$. Furthermore,

$$\begin{aligned}\frac{\partial}{\partial s}\mathscr{S}_\star(\boldsymbol{\Lambda}, s)\big|_{s=\sqrt{1-q_0}} &= \sqrt{1-q_0}\psi_0 - \alpha_\star \mathbb{E}\left\{\mathcal{E}\left(\frac{\kappa - \hat{\boldsymbol{H}}}{\sqrt{1-q_0}}\right)\boldsymbol{N}\right\} \\ &= \sqrt{1-q_0}\left(\psi_0 - \alpha_\star \mathbb{E}[\boldsymbol{N}^2]\right) = 0.\end{aligned}$$

By Lemma 2.7.14, this implies $s = \sqrt{1-q_0}$ minimizes $\mathscr{S}_\star(\boldsymbol{\Lambda}, s)$, and thus $\mathscr{S}_\star(\boldsymbol{\Lambda}) = \mathscr{S}_\star(\boldsymbol{\Lambda}, \sqrt{1-q_0})$. $\square$

Proof of Proposition 2.7.11. The set $\overline{\mathscr{K}}_*$ is compact in the topology of $\mathscr{L}$. The functions $\mathscr{S}_\star^{s_{\max}} : \overline{\mathscr{K}}_* \to \mathbb{R}$ are continuous by Lemma 2.7.12 and compactness of $[0, s_{\max}]$. On any sequence of $s_{\max}$ tending to ∞, the sequence of $\mathscr{S}_\star^{s_{\max}}$ is decreasing with pointwise limit $\mathscr{S}_\star$. Since Lemma 2.7.15 implies $\mathscr{S}_\star$ is not $-\infty$ everywhere, the result follows from Lemma 2.7.13. $\square$

2.7.4 No boundary maximizers and conclusion

The results proved so far imply that the exponential order of $\mathbb{E}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{Pl}} Z_N(\boldsymbol{G})$ is bounded up to vanishing error by $\sup_{\boldsymbol{\Lambda}\in\overline{\mathscr{K}}_*} \mathscr{S}_\star(\boldsymbol{\Lambda})$. Condition 2.1.3 provides a bound on $\sup_{\boldsymbol{\Lambda}\in\mathscr{K}_*} \mathscr{S}_\star(\boldsymbol{\Lambda})$. Since $\mathscr{S}_\star$ (unlike $\mathscr{S}_\star^{s_{\max}}$) is not a priori continuous, to complete the proof we verify in the following proposition that it is not maximized on the boundary.

Proposition 2.7.16. *The maximum of $\mathscr{S}_\star(\boldsymbol{\Lambda})$ on $\overline{\mathscr{K}}_*$ (which exists by Proposition 2.7.11) is not attained on $\overline{\mathscr{K}}_* \setminus \mathscr{K}_*$.*

Lemma 2.7.17. *Let $d_0 = \alpha_\star \mathbb{E}[F'_{1-q_0}(q_0^{1/2}Z)]$, and*

$$\mathscr{O} = \left\{\boldsymbol{\Lambda} \in \mathscr{K} : d_0\, \mathbb{E}[\boldsymbol{M}\boldsymbol{\Lambda}(\dot{\boldsymbol{H}})] + \mathbb{E}[\dot{\boldsymbol{H}}\boldsymbol{\Lambda}(\dot{\boldsymbol{H}})] > \alpha_\star\kappa\right\}.$$

Then, for $\boldsymbol{\Lambda} \in \mathscr{K}$,

$$\lim_{s\to+\infty} \mathscr{S}_\star(\boldsymbol{\Lambda}, s) = \begin{cases} +\infty & \boldsymbol{\Lambda} \in \mathscr{O}, \\ -\infty & \boldsymbol{\Lambda} \not\in \mathscr{O}. \end{cases}$$

Proof. A well-known gaussian tail bound gives $\frac{\varphi(x)}{x} < \Psi(x) < \frac{x\varphi(x)}{1+x^2}$ for all $x > 0$. Thus, for large x,

$$\log \Psi(x) = -\frac{1}{2}x^2 - \log x + O(1). \tag{2.84}$$

Let s be large and define

$$\xi(x) = -\frac{1}{2}x^2 - \mathbf{1}\{s^{1/2} \le x \le s^2\} \log x, \qquad \boldsymbol{U} = \frac{\kappa - \frac{\mathbb{E}[\boldsymbol{M}\boldsymbol{\Lambda}(\dot{\boldsymbol{H}})]}{q_0}\hat{\boldsymbol{H}} - \frac{\mathbb{E}[\dot{\boldsymbol{H}}\boldsymbol{\Lambda}(\dot{\boldsymbol{H}})]}{\psi_0}\boldsymbol{N}}{\sqrt{1 - \frac{\mathbb{E}[\boldsymbol{M}\boldsymbol{\Lambda}(\dot{\boldsymbol{H}})]^2}{q_0}}}, \qquad \boldsymbol{V} = \boldsymbol{U} + s\boldsymbol{N}.$$

Note that

$$\begin{aligned} |\mathbb{E}\log\Psi(\boldsymbol{V}) - \mathbb{E}\,\xi(\boldsymbol{V})| &\le |\mathbb{E}\,\mathbf{1}\{\boldsymbol{V} \le \log\log s\}(\log\Psi(\boldsymbol{V}) - \xi(\boldsymbol{V}))| \\ &+ \left|\mathbb{E}\,\mathbf{1}\{\log\log s \le \boldsymbol{V} \le s^{1/2}\}(\log\Psi(\boldsymbol{V}) - \xi(\boldsymbol{V}))\right| \\ &+ \left|\mathbb{E}\,\mathbf{1}\{s^{1/2} \le \boldsymbol{V} \le s^2\}(\log\Psi(\boldsymbol{V}) - \xi(\boldsymbol{V}))\right| \\ &+ \left|\mathbb{E}\,\mathbf{1}\{\boldsymbol{V} \ge s^2\}(\log\Psi(\boldsymbol{V}) - \xi(\boldsymbol{V}))\right|. \end{aligned}$$

We will show each of these terms is $o(\log s)$. Let $\boldsymbol{V}_+ = \max(\boldsymbol{V}, 0)$, $\boldsymbol{V}_- = -\min(\boldsymbol{V}, 0)$, and let $C > 0$ be a constant varying from line to line. Then,

$$\begin{aligned} &|\mathbb{E}\,\mathbf{1}\{\boldsymbol{V} \le \log\log s\}(\log\Psi(\boldsymbol{V}) - \xi(\boldsymbol{V}))| \\ &\le \mathbb{E}\,\mathbf{1}\{\boldsymbol{V} \le \log\log s\}|\log\Psi(\boldsymbol{V})| + \mathbb{E}\,\mathbf{1}\{\boldsymbol{V} \le \log\log s\}\boldsymbol{V}_+^2 + \mathbb{E}\,\boldsymbol{V}_-^2 \\ &\le C(\log\log s)^2 + \mathbb{E}\,\boldsymbol{U}_-^2 \le C(\log\log s)^2. \end{aligned}$$

In the last line we used that $\boldsymbol{N} > 0$ almost surely, and thus $\boldsymbol{U}_- \ge \boldsymbol{V}_-$. By the estimate (2.84), if $\log\log s \le \boldsymbol{V} < s^{1/2}$, then $|\log\Psi(\boldsymbol{V}) - \xi(\boldsymbol{V})| \le C\log s$. Thus

$$\begin{aligned} \left|\,\mathbb{E}\,\mathbf{1}\{\log\log s \le \boldsymbol{V} < s^{1/2}\}(\log\Psi(\boldsymbol{V}) - \xi(\boldsymbol{V}))\right| &\le (C\log s)\,\mathbb{P}(\boldsymbol{V} \le s^{1/2}) \\ &\le (C\log s)\left(\mathbb{P}(\boldsymbol{U} \le -s^{1/2}) + \mathbb{P}(s\boldsymbol{N} \le 2s^{1/2})\right) = o(\log s). \end{aligned}$$

The estimate (2.84) directly implies

$$\left|\mathbb{E}\,\mathbf{1}\{s^{1/2} \le \boldsymbol{V} \le s^2\}(\log\Psi(\boldsymbol{V}) - \xi(\boldsymbol{V}))\right| = O(1).$$

Finally, Lemma 2.4.21(a) gives $0 \le \mathcal{E}(x) \le |x| + 1$. Thus

$$|\boldsymbol{V}| \le |\boldsymbol{U}| + \frac{s}{\sqrt{1-q_0}} \mathcal{E}\left(\frac{\kappa - \hat{\boldsymbol{H}}}{\sqrt{1-q_0}}\right) \le Cs(|\hat{\boldsymbol{H}}| + 1).$$

It follows that for $t \ge s^2$, we have $\mathbb{P}(|\boldsymbol{V}| \ge t) \le \exp(-t^2/Cs^2)$. So, crudely

$$\begin{aligned}
\left|\mathbb{E}\,\mathbf{1}\{\boldsymbol{V} \ge s^2\}(\log \Psi(\boldsymbol{V}) - \xi(\boldsymbol{V}))\right| &\le C'\,\mathbb{E}\,\mathbf{1}\{\boldsymbol{V} \ge s^2\}\boldsymbol{V}^2 \\
&\le C'\left(s^2 \exp(-s^2/C) + \int_{s^2}^{\infty} 2t \exp(-t^2/Cs^2)\,\mathsf{d}t\right) \\
&\le C's^2 \exp(-s^2/C).
\end{aligned}$$

Thus $|\mathbb{E} \log \Psi(\boldsymbol{V}) - \mathbb{E}\,\xi(\boldsymbol{V})| = o(\log s)$. So,

$$\mathscr{S}_\star(\boldsymbol{\Lambda}, s) = \frac{1}{2}s^2\psi_0 + \alpha_\star\,\mathbb{E}\,\xi(\boldsymbol{V}) + o(\log s).$$

We now evaluate $\alpha_\star\,\mathbb{E}\,\xi(\boldsymbol{V})$. First,

$$\begin{aligned}
\frac{1}{2}\alpha_\star\,\mathbb{E}\,\boldsymbol{V}^2 &= \frac{1}{2}\alpha_\star s^2\,\mathbb{E}[\boldsymbol{N}^2] + \alpha_\star s\,\mathbb{E}[\boldsymbol{U}\boldsymbol{N}] + O(1) \\
&= \frac{1}{2}s^2\psi_0 + \frac{s\left(\alpha_\star\kappa - d_0\,\mathbb{E}[\boldsymbol{M}\boldsymbol{\Lambda}(\dot{\boldsymbol{H}})] - \mathbb{E}[\dot{\boldsymbol{H}}\boldsymbol{\Lambda}(\dot{\boldsymbol{H}})]\right)}{\sqrt{1 - \frac{\mathbb{E}[\boldsymbol{M}\boldsymbol{\Lambda}(\dot{\boldsymbol{H}})]^2}{q_0}}} + O(1).
\end{aligned}$$

Thus

$$\mathscr{S}_\star(\boldsymbol{\Lambda}, s) = \frac{s\left(d_0\,\mathbb{E}[\boldsymbol{M}\boldsymbol{\Lambda}(\dot{\boldsymbol{H}})] + \mathbb{E}[\dot{\boldsymbol{H}}\boldsymbol{\Lambda}(\dot{\boldsymbol{H}})] - \alpha_\star\kappa\right)}{\sqrt{1 - \frac{\mathbb{E}[\boldsymbol{M}\boldsymbol{\Lambda}(\dot{\boldsymbol{H}})]^2}{q_0}}} - \mathbb{E}\,\mathbf{1}\{s^{1/2} \le \boldsymbol{V} \le s^2\}\log \boldsymbol{V} + o(\log s).$$

The logarithmic term clearly has magnitude $O(\log s)$. So, $\lim_{s\to+\infty} \mathscr{S}_\star(\boldsymbol{\Lambda}, s) = +\infty$ if $\boldsymbol{\Lambda} \in \mathscr{O}$, and $-\infty$ if $\boldsymbol{\Lambda}$ is in the interior of $\mathscr{K} \setminus \mathscr{O}$. Finally, we have shown above that $\mathbb{P}(\boldsymbol{V} < s^{1/2}), \mathbb{P}(\boldsymbol{V} > s^2) = o_s(1)$, so

$$\mathbb{E}\,\mathbf{1}\{s^{1/2} \le \boldsymbol{V} \le s^2\}\log \boldsymbol{V} \ge \frac{1}{2}(1 - o_s(1))\log s.$$

Thus $\lim_{s\to+\infty} \mathscr{S}_\star(\boldsymbol{\Lambda}, s) = -\infty$ for $\boldsymbol{\Lambda}$ on the boundary of $\mathscr{K} \setminus \mathscr{O}$. □

Proof of Proposition 2.7.16. Suppose for contradiction that $\boldsymbol{\Lambda} \in \overline{\mathscr{K}}_* \setminus \mathscr{K}_*$ maximizes $\mathscr{S}_\star(\boldsymbol{\Lambda})$ in $\overline{\mathscr{K}}_*$. By Proposition 2.7.9, $\boldsymbol{\Lambda}$ is also a maximizer of $\mathscr{S}_\star(\boldsymbol{\Lambda})$ in $\mathscr{K}$.

By Lemma 2.7.17, if $\boldsymbol{\Lambda} \notin \mathscr{O}$, then $\mathscr{S}_\star(\boldsymbol{\Lambda}) = -\infty$ is not a maximizer (recall Lemma 2.7.15). Thus $\boldsymbol{\Lambda} \in \mathscr{O}$. Let $\boldsymbol{\Lambda}^t = (1-t)\boldsymbol{\Lambda}$. Since $\mathscr{O}$ is open, $\boldsymbol{\Lambda}^t \in \mathscr{O}$ for $t \in [0, t_+)$, for sufficiently small t_+.

By Lemma 2.7.17, for $t \in [0, t_+)$, the infimum of $\mathscr{S}(\boldsymbol{\Lambda}^t, s)$ is attained at some $s(\boldsymbol{\Lambda}^t) \in [0, +\infty)$. Note that

$$\left.\frac{\partial}{\partial s}\mathscr{S}_\star(\boldsymbol{\Lambda}^t, s)\right|_{s=0} = -\alpha_\star\,\mathbb{E}\left\{\mathcal{E}\left(\frac{\kappa - \frac{\mathbb{E}[\boldsymbol{M}\boldsymbol{\Lambda}^t(\dot{\boldsymbol{H}})]}{q_0}\hat{\boldsymbol{H}} - \frac{\mathbb{E}[\dot{\boldsymbol{H}}\boldsymbol{\Lambda}^t(\dot{\boldsymbol{H}})]}{\psi_0}\boldsymbol{N}}{\sqrt{1 - \frac{\mathbb{E}[\boldsymbol{M}\boldsymbol{\Lambda}^t(\dot{\boldsymbol{H}})]^2}{q_0}}} + s\boldsymbol{N}\right)\boldsymbol{N}\right\} < 0$$

because $\boldsymbol{N} > 0$ almost surely and the image of $\mathcal{E}$ is positive. Combined with Lemma 2.7.14, this implies $s(\boldsymbol{\Lambda}^t)$ is the unique solution to $\frac{\partial}{\partial s}\mathscr{S}_\star(\boldsymbol{\Lambda}, s) = 0$, and $s(\boldsymbol{\Lambda}^t) > 0$.

Note that $\frac{\partial}{\partial s}\mathscr{S}_\star(\boldsymbol{\Lambda}^t, s)$ is differentiable in t, as the denominator $\sqrt{1 - \frac{\mathbb{E}[\boldsymbol{M}\boldsymbol{\Lambda}^t(\dot{\boldsymbol{H}})]^2}{q_0}}$ is bounded away from 0 by Lemma 2.7.2. By Lemma 2.7.14 and the implicit function theorem, $s(\boldsymbol{\Lambda}^t)$ is differentiable in t for all $t \in [0, t_+)$. It follows that

$$\frac{\mathsf{d}}{\mathsf{d}t}\left.\left\{\frac{1}{2}s(\boldsymbol{\Lambda}^t)^2\psi_0 + \alpha_\star\,\mathbb{E}\log\Psi\left(\frac{\kappa - \frac{\mathbb{E}[\boldsymbol{M}\boldsymbol{\Lambda}^t(\dot{\boldsymbol{H}})]}{q_0}\hat{\boldsymbol{H}} - \frac{\mathbb{E}[\dot{\boldsymbol{H}}\boldsymbol{\Lambda}^t(\dot{\boldsymbol{H}})]}{\psi_0}\boldsymbol{N}}{\sqrt{1 - \frac{\mathbb{E}[\boldsymbol{M}\boldsymbol{\Lambda}^t(\dot{\boldsymbol{H}})]^2}{q_0}}} + s(\boldsymbol{\Lambda}^t)\boldsymbol{N}\right)\right\}\right|_{t=0}$$

exists and is finite. However, since $\mathbf{\Lambda} \in \overline{\mathscr{K}}_* \setminus \mathscr{K}_*$, we have $\mathbf{\Lambda}(\dot{\boldsymbol{H}}) \in \{-1, 1\}$ $\dot{\boldsymbol{H}}$-almost surely. Thus

$$\frac{\mathsf{d}}{\mathsf{d}t} \mathsf{ent}(\mathbf{\Lambda}^t)\big|_{t=0} = \frac{\mathsf{d}}{\mathsf{d}t} \mathcal{H}(t/2)\big|_{t=0} = +\infty.$$

Hence $\frac{\mathsf{d}}{\mathsf{d}t} \mathscr{S}_\star(\mathbf{\Lambda}^t)\big|_{t=0} = +\infty$, and $\mathbf{\Lambda}$ is not a maximizer of $\mathscr{S}_\star(\mathbf{\Lambda})$ in $\mathscr{K}$. □

Proof of Proposition 2.3.9. By Propositions 2.7.1, 2.7.9, for any $s_{\max} > 0$,

$$\frac{1}{N} \log \mathbb{E}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{Pl}}[Z_N(\boldsymbol{G})] \le \sup_{\mathbf{\Lambda} \in \mathscr{K}} \mathscr{S}^{s_{\max}}_\star(\mathbf{\Lambda}) + o_{\varepsilon,\upsilon}(1) = \sup_{\mathbf{\Lambda} \in \overline{\mathscr{K}}_*} \mathscr{S}^{s_{\max}}_\star(\mathbf{\Lambda}) + o_{\varepsilon,\upsilon}(1). \tag{2.85}$$

By Propositions 2.7.11 and 2.7.16 and Condition 2.1.3,

$$\lim_{s_{\max} \to \infty} \sup_{\mathbf{\Lambda} \in \overline{\mathscr{K}}_*} \mathscr{S}^{s_{\max}}_\star(\mathbf{\Lambda}) = \sup_{\mathbf{\Lambda} \in \overline{\mathscr{K}}_*} \mathscr{S}_\star(\mathbf{\Lambda}) = \sup_{\mathbf{\Lambda} \in \mathscr{K}_*} \mathscr{S}_\star(\mathbf{\Lambda}) = \sup_{\lambda_1,\lambda_2 \in \mathbb{R}} \mathscr{S}_\star(\lambda_1, \lambda_2) \le 0.$$

Thus, taking the limit $\varepsilon, \upsilon \to 0$ followed by $s_{\max} \to \infty$ in (2.85) implies the result. □

2.7.5 Local analysis of first moment functional at $(1, 0)$

We now prove Lemma 2.2.5. Note that part (a) follows from Proposition 2.7.16, and part (b) was already proved in Lemma 2.7.15. We turn to the proofs of the remaining parts.

Proof of Lemma 2.2.5(c). Let $\mathscr{S}_\star(\lambda_1, \lambda_2, s) = \mathscr{S}_\star(\mathbf{\Lambda}_{\lambda_1,\lambda_2}, s)$, and let $s(\lambda_1, \lambda_2)$ minimize $\mathscr{S}_\star(\lambda_1, \lambda_2, s)$. Lemma 2.7.15 shows $s(1, 0) = \sqrt{1 - q_0}$, and the proof of Proposition 2.7.16 shows that for (λ_1, λ_2) in a neighborhood of $(1, 0)$, $s(\lambda_1, \lambda_2)$ is the unique solution to $\partial_s \mathscr{S}_\star(\lambda_1, \lambda_2, s) = 0$. By Lemma 2.7.14 and the implicit function theorem, $s(\lambda_1, \lambda_2)$ is differentiable in this neighborhood. So,

$$\begin{aligned}\nabla \mathscr{S}_\star(\lambda_1, \lambda_2) &= \nabla_{\lambda_1,\lambda_2} \mathscr{S}_\star(\lambda_1, \lambda_2, s(\lambda_1, \lambda_2)) + \partial_s \mathscr{S}_\star(\lambda_1, \lambda_2, s(\lambda_1, \lambda_2)) \nabla s(\lambda_1, \lambda_2) \\ &= \nabla_{\lambda_1,\lambda_2} \mathscr{S}_\star(\lambda_1, \lambda_2, s(\lambda_1, \lambda_2)),\end{aligned} \tag{2.86}$$

and in particular $\nabla \mathscr{S}_\star(1, 0) = \nabla \overline{\mathscr{S}}_\star(1, 0)$. To calculate the latter gradient, let $u_1, u_2 \in \mathbb{R}$ be arbitrary and

$$\mathbf{\Delta} \equiv (u_1 \partial_{\lambda_1} + u_2 \partial_{\lambda_2}) \mathbf{\Lambda} = (1 - \mathbf{\Lambda}^2)(u_1 \dot{\boldsymbol{H}} + u_2 \boldsymbol{M}).$$

Then

$$\begin{aligned}\langle \nabla \overline{\mathscr{S}}_\star(\lambda_1, \lambda_2), (u_1, u_2) \rangle = &- \mathbb{E}[\mathrm{th}^{-1}(\mathbf{\Lambda})\mathbf{\Delta}] - \alpha_\star \mathbb{E}\left\{ \mathcal{E}\left(\frac{\kappa - \frac{\mathbb{E}[\boldsymbol{M}\mathbf{\Lambda}]}{q_0}\hat{\boldsymbol{H}} - \frac{\mathbb{E}[\dot{\boldsymbol{H}}\mathbf{\Lambda}]}{\psi_0}\boldsymbol{N}}{\sqrt{1 - \frac{\mathbb{E}[\boldsymbol{M}\mathbf{\Lambda}]^2}{q_0}}} + \sqrt{1 - q_0}\boldsymbol{N} \right) \right. \\ &\left. \times \left(\frac{-\frac{\mathbb{E}[\boldsymbol{M}\mathbf{\Delta}]}{q_0}\hat{\boldsymbol{H}} - \frac{\mathbb{E}[\dot{\boldsymbol{H}}\mathbf{\Delta}]}{\psi_0}\boldsymbol{N}}{\sqrt{1 - \frac{\mathbb{E}[\boldsymbol{M}\mathbf{\Lambda}]^2}{q_0}}} + \frac{\kappa - \frac{\mathbb{E}[\boldsymbol{M}\mathbf{\Lambda}]}{q_0}\hat{\boldsymbol{H}} - \frac{\mathbb{E}[\dot{\boldsymbol{H}}\mathbf{\Lambda}]}{\psi_0}\boldsymbol{N}}{\left(1 - \frac{\mathbb{E}[\boldsymbol{M}\mathbf{\Lambda}]^2}{q_0}\right)^{3/2}} \cdot \frac{\mathbb{E}[\boldsymbol{M}\mathbf{\Lambda}]\,\mathbb{E}[\boldsymbol{M}\mathbf{\Delta}]}{q_0} \right) \right\}.\end{aligned} \tag{2.87}$$

Specializing to $(\lambda_1, \lambda_2) = (1, 0)$,

$$\begin{aligned}&\langle \nabla \overline{\mathscr{S}}_\star(1, 0), (u_1, u_2) \rangle \\ &= -\mathbb{E}[\mathrm{th}^{-1}(\boldsymbol{M})\mathbf{\Delta}] - \alpha_\star \mathbb{E}\left\{ \mathcal{E}\left(\frac{\kappa - \hat{\boldsymbol{H}}}{\sqrt{1 - q_0}} \right) \left(\frac{-\frac{\mathbb{E}[\boldsymbol{M}\mathbf{\Delta}]}{q_0}\hat{\boldsymbol{H}} - \frac{\mathbb{E}[\dot{\boldsymbol{H}}\mathbf{\Delta}]}{\psi_0}\boldsymbol{N}}{\sqrt{1 - q_0}} + \frac{\kappa - \hat{\boldsymbol{H}} - (1 - q_0)\boldsymbol{N}}{(1 - q_0)^{3/2}} \mathbb{E}[\boldsymbol{M}\mathbf{\Delta}] \right) \right\} \\ &= -\mathbb{E}[\dot{\boldsymbol{H}}\mathbf{\Delta}] - \alpha_\star \mathbb{E}\left\{ F_{1-q_0}(\hat{\boldsymbol{H}}) \left(-\frac{\mathbb{E}[\boldsymbol{M}\mathbf{\Delta}]}{q_0}\hat{\boldsymbol{H}} - \frac{\mathbb{E}[\dot{\boldsymbol{H}}\mathbf{\Delta}]}{\psi_0}\boldsymbol{N} + \frac{\kappa - \hat{\boldsymbol{H}} - (1 - q_0)\boldsymbol{N}}{1 - q_0} \mathbb{E}[\boldsymbol{M}\mathbf{\Delta}] \right) \right\} \\ &= -\mathbb{E}[\dot{\boldsymbol{H}}\mathbf{\Delta}] + \frac{\alpha_\star \mathbb{E}[\boldsymbol{N}^2]}{\psi_0} \mathbb{E}[\dot{\boldsymbol{H}}\mathbf{\Delta}] + \alpha_\star \left(\frac{\mathbb{E}[\boldsymbol{N}\hat{\boldsymbol{H}}]}{q_0} + \mathbb{E}\left[\boldsymbol{N}\left(\boldsymbol{N} - \frac{\kappa - \hat{\boldsymbol{H}}}{1 - q_0} \right) \right] \right) \mathbb{E}[\boldsymbol{M}\mathbf{\Delta}].\end{aligned}$$

The first two terms cancel because $\alpha_\star \mathbb{E}[\boldsymbol{N}^2] = \psi_0$. Finally, note the identity

$$F'_{1-q_0}(x) = -F_{1-q_0}(x)\left(F_{1-q_0}(x) - \frac{x}{1-q_0}\right).$$

By gaussian integration by parts,

$$\mathbb{E}[\boldsymbol{N}\hat{\boldsymbol{H}}] = \mathbb{E}[\hat{\boldsymbol{H}}F_{1-q_0}(\hat{\boldsymbol{H}})] = \mathbb{E}[\hat{\boldsymbol{H}}^2]\,\mathbb{E}[F'_{1-q_0}(\hat{\boldsymbol{H}})] = -q_0\,\mathbb{E}\left[\boldsymbol{N}\left(\boldsymbol{N} - \frac{\kappa - \hat{\boldsymbol{H}}}{1-q_0}\right)\right].$$

It follows that $\langle \nabla \mathscr{S}_\star(1,0), (u_1,u_2)\rangle = 0$. Since u_1, u_2 were arbitrary, $\nabla \mathscr{S}_\star(1,0) = 0$. ☐

Proof of Lemma 2.2.5(d). Differentiating (2.86) and applying the implicit function theorem yields

$$\begin{aligned}
\nabla^2 \mathscr{S}_\star(\lambda_1,\lambda_2) &= \nabla^2_{\lambda_1,\lambda_2}\mathscr{S}_\star(\lambda_1,\lambda_2,s(\lambda_1,\lambda_2)) + \nabla_{\lambda_1,\lambda_2}\partial_s \mathscr{S}_\star(\lambda_1,\lambda_2,s(\lambda_1,\lambda_2))(\nabla s(\lambda_1,\lambda_2))^\top \\
&= \nabla^2_{\lambda_1,\lambda_2}\mathscr{S}_\star(\lambda_1,\lambda_2,s(\lambda_1,\lambda_2)) - \frac{(\nabla_{\lambda_1,\lambda_2}\partial_s \mathscr{S}_\star(\lambda_1,\lambda_2,s(\lambda_1,\lambda_2)))^{\otimes 2}}{\partial_s^2 \mathscr{S}_\star(\lambda_1,\lambda_2,s(\lambda_1,\lambda_2))} \\
&\preceq \nabla^2_{\lambda_1,\lambda_2}\mathscr{S}_\star(\lambda_1,\lambda_2,s(\lambda_1,\lambda_2)).
\end{aligned}$$

Specializing to $(\lambda_1,\lambda_2) = (1,0)$ yields the result. ☐

Appendix

2.A Deferred proofs

In this appendix, we provide proofs of various results deferred from the paper.

2.A.1 Well definedness and $\varepsilon \downarrow 0$ limit of $(q_\varepsilon, \psi_\varepsilon, \varrho_\varepsilon)$

Proof of Proposition 2.4.1. Let ι_0 be small enough that $[q_0 - 3\iota_0, q_0 + 3\iota_0] \subseteq [0,1]$. Note that $\zeta_0(\psi) = (R_{\alpha_\star} \circ P)(\psi)$. By Condition 2.3.1, $\zeta_0(\psi_0) = \psi_0$ and

$$\zeta_0'(\psi_0) = R_{\alpha_\star}'(q_0)P'(\psi_0) = (P \circ R_{\alpha_\star})'(q_0) < 1.$$

By continuity of ζ_0 and ζ_0', we can find $\iota > 0$ such that for all $\psi \in [\psi_0 - \iota, \psi_0 + \iota]$, $P(\psi) \in [q_0 - \iota_0, q_0 + \iota_0]$ and $\zeta_0'(\psi) < 1$. Set ι_1 small enough that

$$\zeta_0(\psi_0 - \iota) \geq \psi_0 - \iota + 2\iota_1, \qquad \zeta_0(\psi_0 + \iota) \leq \psi_0 + \iota - 2\iota_1, \qquad \sup_{\psi \in [\psi_0 - \iota, \psi_0 + \iota]} \zeta_0'(\psi) \leq 1 - 2\iota_1.$$

We will show that for sufficiently small ε,

$$\sup_{\psi \in [\psi_0 - \iota, \psi_0 + \iota]} |\zeta_\varepsilon(\psi) - \zeta_0(\psi)|, \sup_{\psi \in [\psi_0 - \iota, \psi_0 + \iota]} |\zeta_\varepsilon'(\psi) - \zeta_0'(\psi)| = o_\varepsilon(1). \tag{2.88}$$

We first explain why this implies the result. First, (2.88) implies that for sufficiently small ε,

$$\zeta_\varepsilon(\psi_0 - \iota) \geq \psi_0 - \iota + \iota_1, \qquad \zeta_\varepsilon(\psi_0 + \iota) \leq \psi_0 + \iota - \iota_1, \qquad \sup_{\psi \in [\psi_0 - \iota, \psi_0 + \iota]} \zeta_\varepsilon'(\psi) \leq 1 - \iota_1.$$

This implies that ζ_ε has a unique fixed point ψ_ε in $[\psi_0 - \iota, \psi_0 + \iota]$. Furthermore, it implies $|\zeta_\varepsilon(\psi_0) - \psi_0| = o_\varepsilon(1)$, which combined with the above derivative estimate gives

$$|\psi_\varepsilon - \psi_0| \leq |\zeta_\varepsilon(\psi_0) - \psi_0|/\iota_1 = o_\varepsilon(1).$$

Continuity considerations then imply $(q_\varepsilon, \psi_\varepsilon, \varrho_\varepsilon) \to (q_0, \psi_0, 1 - q_0)$ as $\varepsilon \downarrow 0$. We now turn to the proof of (2.88). Let $\psi \in [\psi_0 - \iota, \psi_0 + \iota]$. Below, $o_\varepsilon(1)$ is an error uniform over ψ. Let $q = P^\varepsilon(\psi)$ and $\widetilde{q} = P(\psi)$. Note that

$$|q - \widetilde{q}| \leq \mathbb{E}\left[|(\mathrm{th}((\psi + \varepsilon)^{1/2} Z) + \varepsilon(\psi + \varepsilon)^{1/2} Z)^2 - \mathrm{th}^2(\psi^{1/2} Z)|\right] \leq o_\varepsilon(1).$$

Let $\varrho = \varrho_\varepsilon(q, \psi)$, and note that

$$|\varrho - (1 - q)| = o_\varepsilon(1).$$

Thus

$$\varrho \geq (1 - \widetilde{q}) - |\widetilde{q} - q| - |\varrho - (1 - q)| \geq 2\iota_0 - o_\varepsilon(1) \geq \iota_0,$$

so ϱ is bounded away from 0. By Cauchy-Schwarz,

$$\begin{aligned}
&|\zeta_\varepsilon(\psi) - \zeta_0(\psi)| = |R^\varepsilon(q, \psi) - R_{\alpha_\star}(\widetilde{q})| \\
&= \alpha_\star \mathbb{E}\left[|F_{\varepsilon,\varrho}((q + \varepsilon)^{1/2} Z) - F_{1-q_0}(q^{1/2} Z)||F_{\varepsilon,\varrho}((q + \varepsilon)^{1/2} Z) + F_{1-q_0}(q^{1/2} Z)|\right] \\
&\leq \alpha_\star \mathbb{E}\left[(F_{\varepsilon,\varrho}((q + \varepsilon)^{1/2} Z) - F_{1-\widetilde{q}}(\widetilde{q}^{1/2} Z))^2\right]^{1/2} \mathbb{E}\left[(F_{\varepsilon,\varrho}((q + \varepsilon)^{1/2} Z) + F_{1-\widetilde{q}}(\widetilde{q}^{1/2} Z))^2\right]^{1/2}.
\end{aligned}$$

Expanding $F_{\varepsilon,\varrho}$ using (2.19) shows the first expectation is $o_\varepsilon(1)$, while the second is bounded by Lemma 2.4.21(a). Thus $|\zeta_\varepsilon(\psi) - \zeta_0(\psi)| = o_\varepsilon(1)$ uniformly in $\psi \in [\psi_0 - \iota, \psi_0 + \iota]$. Furthermore,

$$\zeta_\varepsilon'(\psi) = \frac{\partial R^\varepsilon}{\partial q}(q,\psi)(P^\varepsilon)'(\psi) + \frac{\partial R^\varepsilon}{\partial \psi}(q,\psi), \qquad \zeta_0'(\psi) = R'_{\alpha_\star}(\widetilde{q})P'(\psi).$$

Similar computations to above show

$$\left|\frac{\partial R^\varepsilon}{\partial q}(q,\psi) - R'_{\alpha_\star}(\widetilde{q})\right|, |(P^\varepsilon)'(\psi) - P'(\psi)|, \left|\frac{\partial R^\varepsilon}{\partial \psi}(q,\psi)\right| = o_\varepsilon(1),$$

and thus $|\zeta_\varepsilon'(\psi) - \zeta_0'(\psi)| = o_\varepsilon(1)$ uniformly in ψ. This proves (2.88). □

2.A.2 Approximation for (pseudo)-Lipschitz functions

Proof of Fact 2.4.20. Let (x, y) be a sample from the optimal coupling of (μ, μ'). Then

$$\begin{aligned}
|\mathbb{E}_\mu[f] - \mathbb{E}_{\mu'}[f]| &\le \mathbb{E}\,|f(x) - f(y)| \le L\,\mathbb{E}\,[|x-y|(|x|+|y|+1)] \\
&\le L\,\mathbb{E}[|x-y|^2]^{1/2}\,\mathbb{E}[3(|x|^2+|y|^2+1)]^{1/2} \\
&\le L\,\mathbb{E}[|x-y|^2]^{1/2}\,\mathbb{E}[3(3|x|^2+2|x-y|^2+1)]^{1/2} \\
&\le 3L\mathbb{W}_2(\mu,\mu')(\mu_2 + \mathbb{W}_2(\mu,\mu') + 1),
\end{aligned}$$

where we have used the estimate $|y|^2 \le 2|x|^2 + 2|x-y|^2$. □

Proof of Fact 2.6.11. Couple $(x, y, z) \sim \mu$ and $(x', y', z') \sim \mu'$ in the $\mathbb{W}_2$-optimal way. Then, the left-hand side of (2.62) is bounded by the sum of:

$$\begin{aligned}
\mathbb{E}|f_1(x)||f_2(y)||f_3(z) - f_3(z')| &\le L(\mathbb{E}f_1(x)^4)^{1/4}(\mathbb{E}f_2(y)^4)^{1/4}(\mathbb{E}|z-z'|^2)^{1/2} \\
&\le L(\mathbb{E}f_1(x)^4)^{1/4}(\mathbb{E}f_2(y)^4)^{1/4}\mathbb{W}_2(\mu,\mu'), \\
\mathbb{E}|f_1(x)||f_3(z')||f_2(y) - f_2(y')| &\le L^2(\mathbb{E}f_1(x)^2)^{1/2}(\mathbb{E}|y-y'|^2)^{1/2} \le L^2(\mathbb{E}f_1(x)^2)^{1/2}\mathbb{W}_2(\mu,\mu') \\
\mathbb{E}|f_2(y')||f_3(z')||f_1(x) - f_1(x')| &\le L^2(\mathbb{E}f_2(y')^2)^{1/2}(\mathbb{E}|x-x'|^2)^{1/2} \le L^2(\mathbb{E}f_2(y')^2)^{1/2}\mathbb{W}_2(\mu,\mu').
\end{aligned}$$

Finally, by Fact 2.4.20,

$$\mathbb{E}f_2(y')^2 \le \mathbb{E}f_2(y)^2 + 3\mathbb{W}_2(\mu,\mu')(\mathbb{E}f_2(y)^2 + \mathbb{W}_2(\mu,\mu') + 1).$$

Combining gives the conclusion. □

2.A.3 Gradient and Hessian formulas for $\mathcal{F}^\varepsilon_{\mathsf{TAP}}$, and regularity estimates

Proof of Lemma 2.4.16. By standard properties of convex duals,

$$(V_\varepsilon^*)'(m) = -\arg\min_{\dot h}\left\{-m\dot h + V_\varepsilon(\dot h)\right\} = -\mathrm{th}_\varepsilon^{-1}(m).$$

We differentiate the interaction term in $\mathcal{F}^\varepsilon_{\mathsf{TAP}}$ by gaussian integration by parts. For each $i \in [N]$, $a \in [M]$,

$$\begin{aligned}
&\frac{\partial}{\partial m_i}\overline{F}_{\varepsilon,\rho_\varepsilon(q(\boldsymbol{m}))}\left(\frac{\langle \boldsymbol{g}^a, \boldsymbol{m}\rangle}{\sqrt{N}} + \varepsilon^{1/2}\widehat{g}_a - \rho_\varepsilon(q(\boldsymbol{m}))n_a\right) \\
&= \frac{\partial}{\partial m_i}\log\mathbb{E}\,\chi_\varepsilon\left(\frac{\langle \boldsymbol{g}^a, \boldsymbol{m}\rangle}{\sqrt{N}} + \varepsilon^{1/2}\widehat{g}_a - \rho_\varepsilon(q(\boldsymbol{m}))n_a + \rho_\varepsilon(q(\boldsymbol{m}))^{1/2}Z\right) \\
&= \frac{\mathbb{E}\,\chi_\varepsilon'\left(\frac{\langle \boldsymbol{g}^a, \boldsymbol{m}\rangle}{\sqrt{N}} + \varepsilon^{1/2}\widehat{g}_a - \rho_\varepsilon(q(\boldsymbol{m}))n_a + \rho_\varepsilon(q(\boldsymbol{m}))^{1/2}Z\right)\left(\frac{g_i^a}{\sqrt{N}} - \rho_\varepsilon'(q(\boldsymbol{m}))\frac{2m_in_a}{N} + \frac{\rho_\varepsilon'(q(\boldsymbol{m}))}{\rho_\varepsilon(q(\boldsymbol{m}))^{1/2}}\frac{m_i}{N}Z\right)}{\mathbb{E}\,\chi_\varepsilon'\left(\frac{\langle \boldsymbol{g}^a, \boldsymbol{m}\rangle}{\sqrt{N}} + \varepsilon^{1/2}\widehat{g}_a - \rho_\varepsilon(q(\boldsymbol{m}))n_a + \rho_\varepsilon(q(\boldsymbol{m}))^{1/2}Z\right)} \\
&= F_{\varepsilon,\rho_\varepsilon(q(\boldsymbol{m}))}(\acute{h}_a)\left(\frac{g_i^a}{\sqrt{N}} - \rho_\varepsilon'(q(\boldsymbol{m}))\frac{2m_in_a}{N}\right) + \frac{\mathbb{E}\,\chi_\varepsilon''(\acute{h}_a + \rho_\varepsilon(q(\boldsymbol{m}))^{1/2}Z)}{\mathbb{E}\,\chi_\varepsilon(\acute{h}_a + \rho_\varepsilon(q(\boldsymbol{m}))^{1/2}Z)}\cdot\frac{\rho_\varepsilon'(q(\boldsymbol{m}))m_i}{N}. \\
&= \frac{g_i^a}{\sqrt{N}}F_{\varepsilon,\rho_\varepsilon(q(\boldsymbol{m}))}(\acute{h}_a) + \frac{\rho_\varepsilon'(q(\boldsymbol{m}))m_i}{N}\left(-2F_{\varepsilon,\rho_\varepsilon(q(\boldsymbol{m}))}(\acute{h}_a)n_a + F_{\varepsilon,\rho_\varepsilon(q(\boldsymbol{m}))}(\acute{h}_a)^2 + F'_{\varepsilon,\rho_\varepsilon(q(\boldsymbol{m}))}(\acute{h}_a)\right).
\end{aligned}$$

Thus

$$\frac{\partial}{\partial m_i}\mathcal{F}^{\varepsilon}_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n}) = -\mathrm{th}_{\varepsilon}^{-1}(m_i) + \varepsilon^{1/2}\dot{g}_i + \frac{(\boldsymbol{G}^{\top}F_{\varepsilon,\rho_\varepsilon(q(\boldsymbol{m}))}(\acute{h}_a))_i}{\sqrt{N}}$$
$$+ \frac{\rho_\varepsilon'(q(\boldsymbol{m}))m_i}{N}\sum_{a=1}^{M}\Big((n_a - F_{\varepsilon,\rho_\varepsilon(q(\boldsymbol{m}))}(\acute{h}_a))^2 + F'_{\varepsilon,\rho_\varepsilon(q(\boldsymbol{m}))}(\acute{h}_a)\Big),$$

which implies (2.34). The formula (2.35) follows by directly differentiating $\mathcal{F}^{\varepsilon}_{\mathsf{TAP}}$. Setting (2.35) to zero shows that $\nabla_{\boldsymbol{n}}\mathcal{F}^{\varepsilon}_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n}) = 0$ if and only if $\acute{\boldsymbol{h}} = \widehat{\boldsymbol{h}}$, which rearranges to (2.36). This implies $F_{\varepsilon,\rho_\varepsilon(q(\boldsymbol{m}))}(\acute{\boldsymbol{h}}) = \boldsymbol{n}$, so setting (2.34) to zero yields (2.37). □

Proof of Fact 2.6.5. Note that

$$\frac{\partial}{\partial m_i}\mathrm{th}_{\varepsilon}^{-1}(m_i) = \frac{1}{\mathrm{th}_\varepsilon'(\dot{h}_i)} = \frac{1}{1+\varepsilon-\mathrm{th}^2(\dot{h}_i)} = \frac{\mathrm{ch}^2(\dot{h}_i)}{1+\varepsilon\mathrm{ch}^2(\dot{h}_i)}.$$

The functions $F_{\varepsilon,\varrho}, F'_{\varepsilon,\varrho}$ can be differentated in ϱ as follows. By gaussian integration by parts (or Itô's formula),

$$\frac{\mathsf{d}}{\mathsf{d}\varrho}\,\mathbb{E}\,\chi_\varepsilon(x+\varrho^{1/2}Z) = \frac{1}{2}\,\mathbb{E}\,\chi_\varepsilon''(x+\varrho^{1/2}Z),$$

and similarly for χ_ε'. Thus, abbreviating $\chi_{\varepsilon,\varrho}(x) = \mathbb{E}\,\chi_\varepsilon(x+\varrho^{1/2}Z)$,

$$\frac{\mathsf{d}}{\mathsf{d}\varrho}F_{\varepsilon,\varrho}(x) = \frac{\mathsf{d}}{\mathsf{d}\varrho}\frac{\chi_{\varepsilon,\varrho}(x)}{\chi'_{\varepsilon,\varrho}(x)} = \frac{1}{2}\left(\frac{\chi^{(3)}_{\varepsilon,\varrho}(x)}{\chi_{\varepsilon,\varrho}(x)} - \frac{\chi'_{\varepsilon,\varrho}(x)\chi''_{\varepsilon,\varrho}(x)}{\chi_{\varepsilon,\varrho}(x)^2}\right).$$

We also have

$$F'_{\varepsilon,\varrho}(x) = \frac{\chi''_{\varepsilon,\varrho}(x)}{\chi_{\varepsilon,\varrho}(x)} - \frac{(\chi'_{\varepsilon,\varrho}(x))^2}{\chi_{\varepsilon,\varrho}(x)^2}, \qquad F''_{\varepsilon,\varrho}(x) = \frac{\chi'''_{\varepsilon,\varrho}(x)}{\chi_{\varepsilon,\varrho}(x)} - \frac{3(\chi'_{\varepsilon,\varrho}(x))(\chi''_{\varepsilon,\varrho}(x))}{\chi_{\varepsilon,\varrho}(x)^2} + \frac{2(\chi'_{\varepsilon,\varrho}(x))^3}{\chi_{\varepsilon,\varrho}(x)^3}.$$

Thus

$$\frac{\mathsf{d}}{\mathsf{d}\varrho}F_{\varepsilon,\varrho}(x) = \frac{1}{2}\left(2F_{\varepsilon,\varrho}(x)F'_{\varepsilon,\varrho}(x) + F''_{\varepsilon,\varrho}(x)\right).$$

A similar calculation shows

$$\frac{\mathsf{d}}{\mathsf{d}\varrho}F'_{\varepsilon,\varrho}(x) = \frac{1}{2}\left(2F_{\varepsilon,\varrho}(x)F''_{\varepsilon,\varrho}(x) + 2F'_{\varepsilon,\varrho}(x)^2 + F^{(3)}_{\varepsilon,\varrho}(x)\right).$$

The result follows by directly differentiating (2.34) and (2.35) using the above formulas. □

Proof of Lemma 2.6.6. As $(\boldsymbol{m},\boldsymbol{n}) \in \mathcal{S}_{\varepsilon,r_0}$, approximation arguments identical to the proof of Corollary 2.4.18 show the estimates for $q(\boldsymbol{m}), \psi(\boldsymbol{n}), d_\varepsilon(\boldsymbol{m},\boldsymbol{n})$ in part (a). The regularity estimate (2.23) of ρ_ε and its derivatives proves the rest of part (a). Differentiating (2.19) yields

$$F'_{\varepsilon,\varrho}(x) = -\frac{\varepsilon}{1+\varepsilon\varrho} - \frac{1}{(\varrho+\varepsilon(1+\varepsilon\varrho))(1+\varepsilon\varrho)}\mathcal{E}'\left(\frac{\kappa(1+\varepsilon\varrho)-x}{\sqrt{(\varrho+\varepsilon(1+\varepsilon\varrho))(1+\varepsilon\varrho)}},\right).$$

By Lemma 2.4.21, we see that for ϱ in a neighborhood of ϱ_ε, $\sup_{x\in\mathbb{R}}\left|\frac{\mathsf{d}}{\mathsf{d}\varrho}F'_{\varepsilon,\varrho}(x)\right|$ is bounded by an absolute constant. Note that

$$\sup_{x\in\mathbb{R}}\left|\frac{\mathsf{d}}{\mathsf{d}\varrho}\frac{F'_{\varepsilon,\varrho}(x)}{1+\varrho F'_{\varepsilon,\varrho}(x)}\right| \le \sup_{x\in\mathbb{R}}\left|\frac{F'_{\varepsilon,\varrho}(x)}{(1+\varrho F'_{\varepsilon,\varrho}(x))^2}\right| + \sup_{x\in\mathbb{R}}\left|\frac{1}{(1+\varrho F'_{\varepsilon,\varrho}(x))^2}\right|\cdot\sup_{x\in\mathbb{R}}\left|\frac{\mathsf{d}}{\mathsf{d}\varrho}F'_{\varepsilon,\varrho}(x)\right|. \tag{2.89}$$

By (2.42),

$$\frac{1}{1+\varrho F'_{\varepsilon,\varrho}(x)} \ge \frac{\varrho+\varepsilon(1+\varepsilon\varrho)}{\varepsilon},$$

which for ϱ in a neighborhood of ϱ_ε is bounded depending only on ε. It follows that (2.89) is is bounded depending only on ε. So,

$$\|\boldsymbol{D}_2 - \widetilde{\boldsymbol{D}}_2\|_{\mathsf{op}} \le \left| \frac{F'_{\varepsilon,\varrho_\varepsilon}(x)}{1+\varrho_\varepsilon F'_{\varepsilon,\varrho_\varepsilon}(x)} - \frac{F'_{\varepsilon,\rho_\varepsilon(q(\boldsymbol{m}))}(x)}{1+\rho_\varepsilon(q(\boldsymbol{m}))F'_{\varepsilon,\rho_\varepsilon(q(\boldsymbol{m}))}(x)} \right| = o_{r_0}(1).$$

This proves part (b). Part (c) follows from Fact 2.4.22, as (for $\rho_\varepsilon(q(\boldsymbol{m}))$ in a neighborhood of $\varrho_\varepsilon > 0$) the images of $F'_{\varepsilon,\rho_\varepsilon(q(\boldsymbol{m}))}$ and $F^{(3)}_{\varepsilon,\rho_\varepsilon(q(\boldsymbol{m}))}$ are bounded. Similarly,

$$\frac{1}{\sqrt{N}}\|\boldsymbol{D}_4^{-1}F''(\acute{\boldsymbol{h}})\| \le \|\boldsymbol{D}_4^{-1}\|_{\mathsf{op}}\|F''(\acute{\boldsymbol{h}})\|_\infty \overset{(2.42)}{\le} \frac{\rho_\varepsilon(q(\boldsymbol{m})) + \varepsilon(1+\varepsilon\rho_\varepsilon(q(\boldsymbol{m})))}{\varepsilon}\|F''(\acute{\boldsymbol{h}})\|_\infty.$$

Since the image of $F''_{\varepsilon,\rho_\varepsilon(q(\boldsymbol{m}))}$ is bounded by Fact 2.4.22, this proves part (d). □

Proof of Proposition 2.4.7. We will show that the matrices $\nabla^2_{\boldsymbol{m},\boldsymbol{m}}\mathcal{F}^\varepsilon_{\mathsf{TAP}}$, $\nabla^2_{\boldsymbol{m},\boldsymbol{n}}\mathcal{F}^\varepsilon_{\mathsf{TAP}}$, $\nabla^2_{\boldsymbol{n},\boldsymbol{n}}\mathcal{F}^\varepsilon_{\mathsf{TAP}}$ in Fact 2.6.5 have bounded operator norm (with bound depending on $\varepsilon, C_{\mathsf{cvx}}, C_{\mathsf{bd}}, D$). Throughout this proof, C is a constant depending on $\varepsilon, C_{\mathsf{cvx}}, C_{\mathsf{bd}}, D$, which may change from line to line.

Under $\mathbb{P}$, we have $\|\boldsymbol{G}\|_{\mathsf{op}}, \|\widehat{\boldsymbol{g}}\| \le C\sqrt{N}$ with high probability. Under $\mathbb{P}^{\boldsymbol{m}',\boldsymbol{n}'}_{\varepsilon,\mathsf{Pl}}$, we may write $\boldsymbol{G} = \mathbb{E}^{\boldsymbol{m}',\boldsymbol{n}'}_{\varepsilon,\mathsf{Pl}}\boldsymbol{G} + \widetilde{\boldsymbol{G}}$ for $\widetilde{\boldsymbol{G}}$ as in Lemma 2.4.17. Then $\|\widetilde{\boldsymbol{G}}\|_{\mathsf{op}} \le C\sqrt{N}$ with high probability, and by Lemma 2.4.17, $\|\mathbb{E}^{\boldsymbol{m}',\boldsymbol{n}'}_{\varepsilon,\mathsf{Pl}}\boldsymbol{G}\| \le C\sqrt{N}$. On this event, $\|\boldsymbol{G}\|_{\mathsf{op}} \le C\sqrt{N}$. Since $\rho_\varepsilon(q(\boldsymbol{m}')) \in [C_{\mathsf{bd}}^{-1}, C_{\mathsf{bd}}]$, $\widehat{\boldsymbol{h}}' = F^{-1}_{\varepsilon,\rho_\varepsilon(q(\boldsymbol{m}'))}(\boldsymbol{n})$ satisfies $\|\widehat{\boldsymbol{h}}'\| \le C\sqrt{N}$. Then, (2.37) implies $\|\widehat{\boldsymbol{g}}\| \le C\sqrt{N}$. So, under both $\mathbb{P}$ and $\mathbb{P}^{\boldsymbol{m}',\boldsymbol{n}'}_{\varepsilon,\mathsf{Pl}}$, we have $\|\boldsymbol{G}\|_{\mathsf{op}}, \|\widehat{\boldsymbol{g}}\| \le C\sqrt{N}$ with high probability. For the remainder of this proof, we assume this event holds.

Consider any $\|\boldsymbol{m}\|^2, \|\boldsymbol{n}\|^2 \le DN$. The above bounds on $\|\boldsymbol{G}\|_{\mathsf{op}}, \|\widehat{\boldsymbol{g}}\|$ imply $\|\acute{\boldsymbol{h}}\| \le C\sqrt{N}$. By (2.23), $C_{\mathsf{bd}}^{-1} \le \rho_\varepsilon(q(\boldsymbol{m})) \le C_{\mathsf{bd}}$ and $|\rho'_\varepsilon(q(\boldsymbol{m}))|, |\rho''_\varepsilon(q(\boldsymbol{m}))| \le C_{\mathsf{bd}}$. Abbreviate $F = F_{\varepsilon,\rho_\varepsilon(q(\boldsymbol{m}))}$ as above. By Fact 2.4.22,

$$\sup_{x\in\mathbb{R}}|F'(x)|, \sup_{x\in\mathbb{R}}|F''(x)|, \sup_{x\in\mathbb{R}}|F^{(3)}(x)| \le C. \tag{2.90}$$

Thus F is C-Lipschitz. By (2.19),

$$F(0) = \frac{1}{\sqrt{(\rho_\varepsilon(q(\boldsymbol{m})) + \varepsilon(1+\varepsilon\rho_\varepsilon(q(\boldsymbol{m}))))(1+\varepsilon\rho_\varepsilon(q(\boldsymbol{m})))}}\mathcal{E}\left(\frac{\kappa\sqrt{1+\varepsilon\rho_\varepsilon(q(\boldsymbol{m}))}}{\sqrt{\rho_\varepsilon(q(\boldsymbol{m})) + \varepsilon(1+\varepsilon\rho_\varepsilon(q(\boldsymbol{m})))}}\right)$$

is bounded, and thus

$$\|F(\acute{\boldsymbol{h}})\| \le \|F(\boldsymbol{0})\| + C\|\acute{\boldsymbol{h}}\| \le C\sqrt{N}.$$

By (2.90) we also have $\|F'(\acute{\boldsymbol{h}})\|, \|F''(\acute{\boldsymbol{h}})\|, \|F^{(3)}(\acute{\boldsymbol{h}})\| \le C\sqrt{N}$. This also implies $d_\varepsilon(\boldsymbol{m},\boldsymbol{n}) \le C$.

Since $\dot{f}_\varepsilon$ is bounded, $\|\boldsymbol{D}_1\|_{\mathsf{op}} \le C$. Since F' is bounded, $\|\boldsymbol{D}_3\|_{\mathsf{op}}, \|\boldsymbol{D}_4\|_{\mathsf{op}} \le C$. The estimate (2.42) also implies $\|\widetilde{\boldsymbol{D}}_2\|_{\mathsf{op}}, \|\boldsymbol{D}_4^{-1}\|_{\mathsf{op}} \le C$. Combining these estimates shows $\|\nabla^2_{\boldsymbol{m},\boldsymbol{m}}\mathcal{F}^\varepsilon_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n})\|_{\mathsf{op}}$, $\|\nabla^2_{\boldsymbol{m},\boldsymbol{n}}\mathcal{F}^\varepsilon_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n})\|_{\mathsf{op}}$, $\|\nabla^2_{\boldsymbol{n},\boldsymbol{n}}\mathcal{F}^\varepsilon_{\mathsf{TAP}}(\boldsymbol{m},\boldsymbol{n})\|_{\mathsf{op}} \le C$. □

2.A.4 Analysis of AMP iteration in planted model

Proof of Proposition 2.5.4. The state evolution [BMN20, Theorem 1] implies that

$$\frac{1}{N}\sum_{i=1}^N \delta(\dot{h}_i, \dot{\xi}_i, \dot{h}_i^{(1),1}, \dots, \dot{h}_i^{(1),k}) \overset{\mathbb{W}_2}{\to} \mathcal{N}(0, \dot{\Sigma}^{(1)}_{\le k}), \qquad \frac{1}{M}\sum_{a=1}^M \delta(\widehat{h}_a, \widehat{\xi}_a, \widehat{h}_a^{(1),0}, \dots, \widehat{h}_a^{(1),k}) \overset{\mathbb{W}_2}{\to} \mathcal{N}(0, \widehat{\Sigma}^{(1)}_{\le k}),$$

for the following arrays $\dot{\Sigma}^{(1)}, \widehat{\Sigma}^{(1)}$. First, $\widehat{\Sigma}^{(1)}$ agrees with $\widehat{\Sigma}^+$ on indices (i,j) where $\{(i,j)\} \cap \{\diamond, \bowtie\} \neq \emptyset$, and $\dot{\Sigma}^{(1)}$ agrees with $\dot{\Sigma}^+$ on (i,j) where $\{(i,j)\} \cap \{\diamond, \bowtie, 0\} \neq \emptyset$. The remaining entries are defined by the following recursion. For $(\dot{H}, \dot{\Xi}, \dot{H}_1, \dots, \dot{H}_k) \sim \mathcal{N}(0, \dot{\Sigma}^{(1)}_{\le k})$ and $0 \le i \le k$,

$$\widehat{\Sigma}^{(1)}_{i,k} = \mathbb{E}\left[\left(\mathrm{th}_\varepsilon(\dot{H}_i) - \frac{\overline{q}_i}{q_\varepsilon}\mathrm{th}_\varepsilon(\dot{H})\right)\left(\mathrm{th}_\varepsilon(\dot{H}_k) - \frac{\overline{q}_k}{q_\varepsilon}\mathrm{th}_\varepsilon(\dot{H})\right)\right] + \frac{\varepsilon(q_\varepsilon - \overline{q}_i)(q_\varepsilon - \overline{q}_k)}{q_\varepsilon(q_\varepsilon+\varepsilon)} + \frac{(\overline{q}_i+\varepsilon)(\overline{q}_k+\varepsilon)}{q_\varepsilon+\varepsilon}. \tag{2.91}$$

For $(\widehat{H}, \widehat{\Xi}, \widehat{H}_0, \ldots, \widehat{H}_k) \sim \mathcal{N}(0, \widehat{\Sigma}^{(1)}_{\le k})$ and $0 \le i \le k$, we have

$$\dot{\Sigma}^{(1)}_{i+1,k+1} = \alpha_\star \, \mathbb{E}\left[\left(F_{\varepsilon,\varrho_\varepsilon}(\widehat{H}_i) - \frac{\overline{\psi}_{i+1}}{\psi_\varepsilon} F_{\varepsilon,\varrho_\varepsilon}(\widehat{H})\right)\left(F_{\varepsilon,\varrho_\varepsilon}(\widehat{H}_k) - \frac{\overline{\psi}_{k+1}}{\psi_\varepsilon} F_{\varepsilon,\varrho_\varepsilon}(\widehat{H})\right)\right] + \frac{\varepsilon(\psi_\varepsilon - \overline{\psi}_{i+1})(\psi_\varepsilon - \overline{\psi}_{k+1})}{\psi_\varepsilon(\psi_\varepsilon + \varepsilon)} + \frac{(\overline{\psi}_{i+1} + \varepsilon)(\overline{\psi}_{k+1} + \varepsilon)}{\psi_\varepsilon + \varepsilon}. \tag{2.92}$$

We now verify by induction that $\widehat{\Sigma}^{(1)}$ and $\dot{\Sigma}^{(1)}$ coincide with $\widehat{\Sigma}^+$ and $\dot{\Sigma}^+$. Suppose $\dot{\Sigma}^{(1)}_{\le k} = \dot{\Sigma}^+_{\le k}$. Then,

$$\mathbb{E}[\mathrm{th}_\varepsilon(\dot{H}_i)\mathrm{th}_\varepsilon(\dot{H}_k)] = \dot{\Sigma}_{i,k}, \qquad \mathbb{E}[\mathrm{th}_\varepsilon(\dot{H}_i)\mathrm{th}_\varepsilon(\dot{H})] = \overline{q}_i, \qquad \mathbb{E}[\mathrm{th}_\varepsilon(\dot{H})^2] = q_\varepsilon,$$

so the right-hand side of (2.91) simplifies as

$$\dot{\Sigma}_{i,k} - \frac{\overline{q}_i\overline{q}_k}{q_\varepsilon} + \frac{\varepsilon(q_\varepsilon - \overline{q}_i)(q_\varepsilon - \overline{q}_k)}{q_\varepsilon(q_\varepsilon + \varepsilon)} + \frac{(\overline{q}_i + \varepsilon)(\overline{q}_k + \varepsilon)}{q_\varepsilon + \varepsilon} = \dot{\Sigma}_{i,k} + \varepsilon = \dot{\Sigma}^+_{i,k}.$$

Now, suppose $\widehat{\Sigma}^{(1)}_{\le k} = \widehat{\Sigma}^+_{\le k}$. Then,

$$\alpha_\star \, \mathbb{E}[F_{\varepsilon,\varrho_\varepsilon}(\widehat{H}_i)F_{\varepsilon,\varrho_\varepsilon}(\widehat{H}_k)] = \widehat{\Sigma}_{i+1,k+1}, \qquad \alpha_\star \, \mathbb{E}[F_{\varepsilon,\varrho_\varepsilon}(\widehat{H}_i)F_{\varepsilon,\varrho_\varepsilon}(\widehat{H})] = \overline{\psi}_{i+1}, \qquad \alpha_\star \, \mathbb{E}[F_{\varepsilon,\varrho_\varepsilon}(\widehat{H})^2] = \psi_\varepsilon,$$

so the right-hand side of (2.92) simplifies as

$$\widehat{\Sigma}_{i+1,k+1} - \frac{\overline{\psi}_{i+1}\overline{\psi}_{k+1}}{\psi_\varepsilon} + \frac{\varepsilon(\psi_\varepsilon - \overline{\psi}_{i+1})(\psi_\varepsilon - \overline{\psi}_{k+1})}{\psi_\varepsilon(\psi_\varepsilon + \varepsilon)} + \frac{(\overline{\psi}_{i+1} + \varepsilon)(\overline{\psi}_{k+1} + \varepsilon)}{\psi_\varepsilon + \varepsilon} = \widehat{\Sigma}_{i+1,k+1} + \varepsilon = \widehat{\Sigma}^+_{i+1,k+1}.$$

This completes the induction. □

To prove Proposition 2.5.5, we introduce two additional auxiliary AMP iterations. They are initialized at $\boldsymbol{n}^{(2),-1} = \boldsymbol{n}^{(3),-1} = \boldsymbol{0}$, $\boldsymbol{m}^{(2),0} = \boldsymbol{m}^{(3),0} = q_\varepsilon^{1/2}\boldsymbol{1}$, with iteration

$$\boldsymbol{m}^{(i),k} = \mathrm{th}_\varepsilon(\dot{\boldsymbol{h}}^{(i),k}), \qquad \boldsymbol{n}^{(i),k} = F_{\varepsilon,\varrho_\varepsilon}(\widehat{\boldsymbol{h}}^{(i),k}),$$

for $i \in \{2,3\}$ and $\dot{\boldsymbol{h}}^{(i),k}, \widehat{\boldsymbol{h}}^{(i),k}$ as follows. Recall that $\overline{\boldsymbol{G}}$ is the matrix (2.44), and $\overline{\psi}_0 = 0$. Then,

$$\widehat{\boldsymbol{h}}^{(2),k} = \frac{1}{\sqrt{N}}\overline{\boldsymbol{G}}\left(\boldsymbol{m}^{(2),k} - \frac{\overline{q}_k}{q_\varepsilon}\boldsymbol{m}\right) + \frac{\sqrt{\varepsilon}(q_\varepsilon - \overline{q}_k)}{\sqrt{q_\varepsilon(q_\varepsilon + \varepsilon)}}\widehat{\boldsymbol{\xi}} + \frac{\overline{q}_k + \varepsilon}{q_\varepsilon + \varepsilon}\widehat{\boldsymbol{h}} - \varrho_\varepsilon\left(\boldsymbol{n}^{(2),k-1} - \frac{\overline{\psi}_k}{\psi_\varepsilon}\boldsymbol{n}\right) \tag{2.93}$$

$$\dot{\boldsymbol{h}}^{(2),k+1} = \frac{1}{\sqrt{N}}\overline{\boldsymbol{G}}^\top\left(\boldsymbol{n}^{(2),k} - \frac{\overline{\psi}_{k+1}}{\psi_\varepsilon}\boldsymbol{n}\right) + \frac{\sqrt{\varepsilon}(\psi_\varepsilon - \psi_{k+1})}{\sqrt{\psi_\varepsilon(\psi_\varepsilon + \varepsilon)}}\dot{\boldsymbol{\xi}} + \frac{\overline{\psi}_{k+1} + \varepsilon}{\psi_\varepsilon + \varepsilon}\dot{\boldsymbol{h}} - d_\varepsilon\left(\boldsymbol{m}^{(2),k} - \frac{\overline{q}_k}{q_\varepsilon}\boldsymbol{m}\right)$$

$$\widehat{\boldsymbol{h}}^{(3),k} = \frac{1}{\sqrt{N}}\widetilde{\boldsymbol{G}}\left(\boldsymbol{m}^{(3),k} - \boldsymbol{m}\right) + \frac{\overline{q}_k + \varepsilon}{q_\varepsilon + \varepsilon}\widehat{\boldsymbol{h}} - \varrho_\varepsilon\left(\boldsymbol{n}^{(3),k-1} - \frac{\overline{\psi}_k + \boldsymbol{1}\{k \ge 1\}\varepsilon}{\psi_\varepsilon + \varepsilon}\boldsymbol{n}\right) \tag{2.94}$$

$$\dot{\boldsymbol{h}}^{(3),k+1} = \frac{1}{\sqrt{N}}\widetilde{\boldsymbol{G}}^\top\left(\boldsymbol{n}^{(3),k} - \boldsymbol{n}\right) + \frac{\overline{\psi}_{k+1} + \varepsilon}{\psi_\varepsilon + \varepsilon}\dot{\boldsymbol{h}} - d_\varepsilon\left(\boldsymbol{m}^{(3),k} - \frac{\overline{q}_k + \varepsilon}{q_\varepsilon + \varepsilon}\boldsymbol{m}\right).$$

The following proposition shows that all these AMP iterations approximate each other.

Proposition 2.A.1. *For any $k \ge 0$, as $N \to \infty$ we have the following convergences in probability under $\mathbb{P}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{Pl}}$.*

(a) $\|\widehat{\boldsymbol{h}}^{(1),k} - \widehat{\boldsymbol{h}}^{(2),k}\|/\sqrt{N} \to 0$, *and if* $k \ge 1$, $\|\dot{\boldsymbol{h}}^{(1),k} - \dot{\boldsymbol{h}}^{(2),k}\|/\sqrt{N} \to 0$.

(b) $\|\widehat{\boldsymbol{h}}^{(2),k} - \widehat{\boldsymbol{h}}^{(3),k}\|/\sqrt{N} \to 0$, *and if* $k \ge 1$, $\|\dot{\boldsymbol{h}}^{(2),k} - \dot{\boldsymbol{h}}^{(3),k}\|/\sqrt{N} \to 0$.

(c) $\|\widehat{\boldsymbol{h}}^{(3),k} - \widehat{\boldsymbol{h}}^{k}\|/\sqrt{N} \to 0$, *and if* $k \ge 1$, $\|\dot{\boldsymbol{h}}^{(3),k} - \dot{\boldsymbol{h}}^{k}\|/\sqrt{N} \to 0$.

Proof of Proposition 2.A.1(a). Similarly to (2.45), we can sample $Z' \sim \mathcal{N}(0,1)$, $\dot{\boldsymbol{\xi}}' \sim \mathcal{N}(0, \boldsymbol{I}_N)$, $\widehat{\boldsymbol{\xi}}' \sim \mathcal{N}(0, \boldsymbol{I}_M)$ coupled to $\widehat{\boldsymbol{G}}$ such that

$$\widehat{\boldsymbol{G}} + \boldsymbol{\Delta}' = \overline{\boldsymbol{G}} - \frac{\widehat{\boldsymbol{\xi}}' \boldsymbol{m}^\top}{\|\boldsymbol{m}\|} - \frac{\boldsymbol{n}(\dot{\boldsymbol{\xi}}')^\top}{\|\boldsymbol{n}\|}, \qquad \boldsymbol{\Delta}' = \frac{\boldsymbol{n}\boldsymbol{m}^\top}{\|\boldsymbol{n}\|\|\boldsymbol{m}\|} Z' \tag{2.95}$$

Note that $\|\boldsymbol{\Delta}'\|_{\mathsf{op}} = o(\sqrt{N})$ with high probability. Let $\simeq$ denote equality up to additive $o_N(1)$. By Proposition 2.5.4, for $(\dot{H}, \dot{\Xi}, \dot{H}_1, \ldots, \dot{H}_k) \sim \mathcal{N}(0, \dot{\Sigma}^{(1)}_{\le k})$ and $(\widehat{H}, \widehat{\Xi}, \widehat{H}_0, \ldots, \widehat{H}_k) \sim \mathcal{N}(0, \widehat{\Sigma}^{(1)}_{\le k})$,

$$\frac{1}{N}\langle \boldsymbol{m}, \dot{\boldsymbol{h}}^{(1),k}\rangle \simeq \mathbb{E}[\mathrm{th}_\varepsilon(\dot{H})\dot{H}_k] = \varrho_\varepsilon(\overline{\psi}_k + \varepsilon), \qquad \frac{1}{N}\langle \boldsymbol{n}, \widehat{\boldsymbol{h}}^{(1),k}\rangle \simeq \alpha_\star \mathbb{E}[F_{\varepsilon,\varrho_\varepsilon}(\widehat{H})\widehat{H}_k] = d_\varepsilon(\overline{q}_k + \varepsilon),$$
$$\frac{1}{N}\langle \boldsymbol{m}, \dot{\boldsymbol{h}}\rangle \simeq \mathbb{E}[\mathrm{th}_\varepsilon(\widehat{H})\widehat{H}] = \varrho_\varepsilon(\psi_\varepsilon + \varepsilon), \qquad \frac{1}{N}\langle \boldsymbol{n}, \widehat{\boldsymbol{h}}\rangle \simeq \alpha_\star \mathbb{E}[F_{\varepsilon,\varrho_\varepsilon}(\widehat{H})\widehat{H}] = d_\varepsilon(q_\varepsilon + \varepsilon).$$

Also,

$$\frac{1}{N}\left\langle \boldsymbol{m}, \boldsymbol{m}^{(1),k} - \frac{\overline{q}_k}{q_\varepsilon}\boldsymbol{m}\right\rangle \simeq \overline{q}_k - \frac{\overline{q}_k}{q_\varepsilon}\cdot q_\varepsilon = 0, \qquad \frac{1}{N}\left\langle \boldsymbol{n}, \boldsymbol{n}^{(1),k-1} - \frac{\overline{\psi}_k}{\psi_\varepsilon}\boldsymbol{n}\right\rangle \simeq \overline{\psi}_k - \frac{\overline{\psi}_k}{\psi_\varepsilon}\cdot \psi_\varepsilon = 0. \tag{2.96}$$

Finally $\frac{1}{N}\langle \dot{\boldsymbol{\xi}}, \boldsymbol{m}\rangle \simeq \frac{1}{N}\langle \widehat{\boldsymbol{\xi}}, \boldsymbol{n}\rangle \simeq 0$. Considering the inner product of (2.47) with $\boldsymbol{n}$ shows

$$0 \simeq \frac{1}{N}\left\langle \boldsymbol{n}, \frac{1}{\sqrt{N}}\widehat{\boldsymbol{G}}\left(\boldsymbol{m}^{(1),k} - \frac{\overline{q}_k}{q_\varepsilon}\boldsymbol{m}\right)\right\rangle.$$

We can expand $\widehat{\boldsymbol{G}}$ using (2.95). Since $\boldsymbol{n}^\top \overline{\boldsymbol{G}} = \boldsymbol{0}$, $\frac{1}{N}\langle \boldsymbol{n}, \widehat{\boldsymbol{\xi}}'\rangle \simeq 0$ in probability, and $\|\boldsymbol{\Delta}'\|_{\mathsf{op}} = o(\sqrt{N})$,

$$0 \simeq \frac{1}{N}\left\langle \boldsymbol{n}, \frac{1}{\sqrt{N}}\left(\overline{\boldsymbol{G}} - \frac{\widehat{\boldsymbol{\xi}}'\boldsymbol{m}^\top}{\|\boldsymbol{m}\|} - \frac{\boldsymbol{n}(\dot{\boldsymbol{\xi}}')^\top}{\|\boldsymbol{n}\|} - \boldsymbol{\Delta}'\right)\left(\boldsymbol{m}^{(1),k} - \frac{\overline{q}_k}{q_\varepsilon}\boldsymbol{m}\right)\right\rangle \simeq \frac{\|\boldsymbol{n}\|}{N^{3/2}}\left\langle \dot{\boldsymbol{\xi}}', \boldsymbol{m}^{(1),k} - \frac{\overline{q}_k}{q_\varepsilon}\boldsymbol{m}\right\rangle.$$

Thus,

$$\frac{1}{N}\left\langle \dot{\boldsymbol{\xi}}', \boldsymbol{m}^{(1),k} - \frac{\overline{q}_k}{q_\varepsilon}\boldsymbol{m}\right\rangle \simeq 0 \tag{2.97}$$

in probability for all k. An analogous computation shows

$$\frac{1}{N}\left\langle \widehat{\boldsymbol{\xi}}', \boldsymbol{n}^{(1),k-1} - \frac{\overline{\psi}_k}{\psi_\varepsilon}\boldsymbol{n}\right\rangle \simeq 0.$$

By (2.95),

$$\begin{aligned}\frac{1}{\sqrt{N}}(\widehat{\boldsymbol{G}} - \overline{\boldsymbol{G}})\left(\boldsymbol{m}^{(1),k} - \frac{\overline{q}_k}{q_\varepsilon}\boldsymbol{m}\right) &= \frac{\widehat{\boldsymbol{\xi}}'}{\sqrt{N}\|\boldsymbol{m}\|}\left\langle \boldsymbol{m}^\top, \boldsymbol{m}^{(1),k} - \frac{\overline{q}_k}{q_\varepsilon}\boldsymbol{m}\right\rangle + \frac{\boldsymbol{n}}{\sqrt{N}\|\boldsymbol{n}\|}\left\langle \dot{\boldsymbol{\xi}}', \boldsymbol{m}^{(1),k} - \frac{\overline{q}_k}{q_\varepsilon}\boldsymbol{m}\right\rangle \\ &\quad - \frac{1}{\sqrt{N}}\boldsymbol{\Delta}'\left(\boldsymbol{m}^{(1),k} - \frac{\overline{q}_k}{q_\varepsilon}\boldsymbol{m}\right),\end{aligned}$$

and this has norm $o(\sqrt{N})$ by (2.96), (2.97). Subtracting (2.47) and (2.93) yields

$$\begin{aligned}\widehat{\boldsymbol{h}}^{(1),k} - \widehat{\boldsymbol{h}}^{(2),k} &= \frac{1}{\sqrt{N}}(\widehat{\boldsymbol{G}} - \overline{\boldsymbol{G}})\left(\boldsymbol{m}^{(1),k} - \frac{\overline{q}_k}{q_\varepsilon}\boldsymbol{m}\right) + \frac{1}{\sqrt{N}}\overline{\boldsymbol{G}}(\boldsymbol{m}^{(1),k} - \boldsymbol{m}^{(2),k}) - \varrho_\varepsilon(\boldsymbol{n}^{(1),k-1} - \boldsymbol{n}^{(2),k-1}) \\ &= \frac{1}{\sqrt{N}}\overline{\boldsymbol{G}}(\boldsymbol{m}^{(1),k} - \boldsymbol{m}^{(2),k}) - \varrho_\varepsilon(\boldsymbol{n}^{(1),k-1} - \boldsymbol{n}^{(2),k-1}) + o(\sqrt{N}),\end{aligned}$$

where $o(\sqrt{N})$ denotes a vector with this norm. Analogously,

$$\dot{\boldsymbol{h}}^{(1),k+1} - \dot{\boldsymbol{h}}^{(2),k+1} = \frac{1}{\sqrt{N}}\overline{\boldsymbol{G}}^\top(\boldsymbol{n}^{(1),k} - \boldsymbol{n}^{(2),k}) - d_\varepsilon(\boldsymbol{m}^{(1),k} - \boldsymbol{m}^{(2),k}) + o(\sqrt{N}).$$

On the high probability event that $\|\overline{\boldsymbol{G}}\|_{\mathsf{op}} = O(\sqrt{N})$, we have

$$\begin{aligned}
\|\widehat{\boldsymbol{h}}^{(1),k} - \widehat{\boldsymbol{h}}^{(2),k}\| &\le O(1)\|\boldsymbol{m}^{(1),k} - \boldsymbol{m}^{(2),k}\| + \varrho_\varepsilon\|\boldsymbol{n}^{(1),k-1} - \boldsymbol{n}^{(2),k-1}\| + o(\sqrt{N}), \\
\|\dot{\boldsymbol{h}}^{(1),k+1} - \dot{\boldsymbol{h}}^{(2),k+1}\| &\le O(1)\|\boldsymbol{n}^{(1),k} - \boldsymbol{n}^{(2),k}\| + |d_\varepsilon|\|\boldsymbol{m}^{(1),k} - \boldsymbol{m}^{(2),k}\| + o(\sqrt{N}).
\end{aligned}$$

The claim now follows by induction on k: $\|\boldsymbol{m}^{(1),0} - \boldsymbol{m}^{(2),0}\| = \|\boldsymbol{n}^{(1),-1} - \boldsymbol{n}^{(2),-1}\| = 0$ by initialization, and because th_ε and $F_{\varepsilon,\varrho_\varepsilon}$ are $O(1)$-Lipschitz,

$$\|\boldsymbol{m}^{(1),k} - \boldsymbol{m}^{(2),k}\| \le O(1)\|\dot{\boldsymbol{h}}^{(1),k} - \dot{\boldsymbol{h}}^{(2),k}\|, \qquad \|\boldsymbol{n}^{(1),k} - \boldsymbol{n}^{(2),k}\| \le O(1)\|\widehat{\boldsymbol{h}}^{(1),k} - \widehat{\boldsymbol{h}}^{(2),k}\|,$$

for all $k \ge 1$, $k \ge 0$ respectively. □

Proof of Proposition 2.A.1(b). Note that $\boldsymbol{\Delta}$ defined in (2.46) w.h.p. satisfies $\|\boldsymbol{\Delta}\|_{\mathsf{op}} = o(\sqrt{N})$. We write (2.94) as

$$\begin{aligned}
\widehat{\boldsymbol{h}}^{(3),k} &= \frac{1}{\sqrt{N}}\widetilde{\boldsymbol{G}}(\boldsymbol{m}^{(2),k} - \boldsymbol{m}) + \frac{\overline{q}_k + \varepsilon}{q_\varepsilon + \varepsilon}\widehat{\boldsymbol{h}} - \varrho_\varepsilon\left(\boldsymbol{n}^{(2),k-1} - \frac{\overline{\psi}_k + \mathbf{1}\{k \ge 1\}\varepsilon}{\psi_\varepsilon + \varepsilon}\boldsymbol{n}\right) \\
&\quad + \frac{1}{\sqrt{N}}\widetilde{\boldsymbol{G}}(\boldsymbol{m}^{(3),k} - \boldsymbol{m}^{(2),k}) - \varrho_\varepsilon(\boldsymbol{n}^{(3),k-1} - \boldsymbol{n}^{(2),k-1}).
\end{aligned}$$

By Proposition 2.A.1(a), $\mathbb{W}_2(\mu_{\dot{\boldsymbol{h}}^{(2),k}}, \mu_{\dot{\boldsymbol{h}}^{(1),k}}) = o_N(1)$. So, Fact 2.4.20 and Proposition 2.5.4 imply

$$\begin{aligned}
\frac{1}{N}\langle \boldsymbol{m}, \boldsymbol{m}^{(2),k}\rangle &\simeq \frac{1}{N}\langle \boldsymbol{m}, \boldsymbol{m}^{(1),k}\rangle \simeq \overline{q}_k, \\
\frac{1}{N}\langle \dot{\boldsymbol{\xi}}, \boldsymbol{m}^{(2),k}\rangle &\simeq \frac{1}{N}\langle \dot{\boldsymbol{\xi}}, \boldsymbol{m}^{(1),k}\rangle \simeq \mathbf{1}\{k \ge 1\}\varrho_\varepsilon\frac{\sqrt{\varepsilon}(\psi_\varepsilon - \overline{\psi}_k)}{\sqrt{\psi_\varepsilon(\psi_\varepsilon + \varepsilon)}}.
\end{aligned} \tag{2.98}$$

By (2.45),

$$\begin{aligned}
\frac{1}{\sqrt{N}}\widetilde{\boldsymbol{G}}(\boldsymbol{m}^{(2),k} - \boldsymbol{m}) &= \frac{1}{\sqrt{N}}\left(\overline{\boldsymbol{G}} - \sqrt{\frac{\varepsilon}{q(\boldsymbol{m}) + \varepsilon}} \cdot \frac{\widehat{\boldsymbol{\xi}}\boldsymbol{m}^\top}{\|\boldsymbol{m}\|} - \sqrt{\frac{\varepsilon}{\psi(\boldsymbol{n}) + \varepsilon}} \cdot \frac{\boldsymbol{n}\dot{\boldsymbol{\xi}}^\top}{\|\boldsymbol{n}\|} - \boldsymbol{\Delta}\right)(\boldsymbol{m}^{(2),k} - \boldsymbol{m}) \\
&= \frac{1}{\sqrt{N}}\overline{\boldsymbol{G}}(\boldsymbol{m}^{(2),k} - \boldsymbol{m}) + \frac{\sqrt{\varepsilon}(q_\varepsilon - \overline{q}_k)}{\sqrt{q_\varepsilon(q_\varepsilon + \varepsilon)}}\widehat{\boldsymbol{\xi}} - \mathbf{1}\{k \ge 1\}\varrho_\varepsilon\frac{\varepsilon(\psi_\varepsilon - \overline{\psi}_k)}{\psi_\varepsilon(\psi_\varepsilon + \varepsilon)}\boldsymbol{n} + o(\sqrt{N}).
\end{aligned}$$

Since $\overline{\boldsymbol{G}}\boldsymbol{m} = \mathbf{0}$, we have $\overline{\boldsymbol{G}}(\boldsymbol{m}^{(2),k} - \boldsymbol{m}) = \overline{\boldsymbol{G}}(\boldsymbol{m}^{(2),k} - \frac{\overline{q}_k}{q_\varepsilon}\boldsymbol{m})$. Moreover,

$$\frac{\overline{\psi}_k + \mathbf{1}\{k \ge 1\}\varepsilon}{\psi_\varepsilon + \varepsilon} - \mathbf{1}\{k \ge 1\}\frac{\varepsilon(\psi_\varepsilon - \overline{\psi}_k)}{\psi_\varepsilon(\psi_\varepsilon + \varepsilon)} = \frac{\overline{\psi}_k}{\psi_\varepsilon}.$$

Combining the above and comparing with (2.93) shows

$$\widehat{\boldsymbol{h}}^{(3),k} = \widehat{\boldsymbol{h}}^{(2),k} + \frac{1}{\sqrt{N}}\widetilde{\boldsymbol{G}}(\boldsymbol{m}^{(3),k} - \boldsymbol{m}^{(2),k}) - \varrho_\varepsilon(\boldsymbol{n}^{(3),k-1} - \boldsymbol{n}^{(2),k-1}) + o(\sqrt{N}).$$

Similarly,

$$\dot{\boldsymbol{h}}^{(3),k+1} = \widehat{\boldsymbol{h}}^{(2),k+1} + \frac{1}{\sqrt{N}}\widetilde{\boldsymbol{G}}^\top(\boldsymbol{n}^{(3),k} - \boldsymbol{n}^{(2),k}) - d_\varepsilon(\boldsymbol{m}^{(3),k} - \boldsymbol{m}^{(2),k}) + o(\sqrt{N}).$$

On the high-probability event that $\|\widetilde{\boldsymbol{G}}\|_{\mathsf{op}} = O(\sqrt{N})$, this implies

$$\begin{aligned}
\|\widehat{\boldsymbol{h}}^{(3),k} - \widehat{\boldsymbol{h}}^{(2),k}\| &\le O(1)\|\boldsymbol{m}^{(3),k} - \boldsymbol{m}^{(2),k}\| + \varrho_\varepsilon\|\boldsymbol{n}^{(3),k-1} - \boldsymbol{n}^{(2),k-1}\| + o(\sqrt{N}), \\
\|\dot{\boldsymbol{h}}^{(3),k+1} - \widehat{\boldsymbol{h}}^{(2),k+1}\| &\le O(1)\|\boldsymbol{n}^{(3),k} - \boldsymbol{n}^{(2),k}\| + |d_\varepsilon|\|\boldsymbol{m}^{(3),k} - \boldsymbol{m}^{(2),k}\| + o(\sqrt{N}).
\end{aligned}$$

The result follows by induction on k, like above. □

Proof of Proposition 2.A.1(c). By Corollary 2.4.18, we have

$$\frac{\boldsymbol{G}}{\sqrt{N}} \stackrel{d}{=} \frac{(1+o_N(1))\widehat{\boldsymbol{h}}\boldsymbol{m}^\top}{N(q_\varepsilon+\varepsilon)} + \frac{(1+o_N(1))\boldsymbol{n}\dot{\boldsymbol{h}}^\top}{N(\psi_\varepsilon+\varepsilon)} + \frac{o_N(1)\boldsymbol{n}\boldsymbol{m}^\top}{N} + \frac{\widetilde{\boldsymbol{G}}}{\sqrt{N}}$$
$$= \frac{\widehat{\boldsymbol{h}}\boldsymbol{m}^\top}{N(q_\varepsilon+\varepsilon)} + \frac{\boldsymbol{n}\dot{\boldsymbol{h}}^\top}{N(\psi_\varepsilon+\varepsilon)} + \frac{\widetilde{\boldsymbol{G}}}{\sqrt{N}} + o_N(1), \tag{2.99}$$

for $\widetilde{\boldsymbol{G}}$ as above and $o_N(1)$ a matrix with this operator norm. Since $q(\boldsymbol{m}) \simeq q_\varepsilon$, $\psi(\boldsymbol{n}) \simeq \psi_\varepsilon$, and under $\mathbb{P}^{\boldsymbol{m},\boldsymbol{n}}_{\varepsilon,\mathsf{Pl}}$ we have a.s. $\acute{\boldsymbol{h}} = F^{-1}_{\varepsilon,\rho_\varepsilon(q(\boldsymbol{m}))}(\boldsymbol{n})$, the following terms appearing in (2.36), (2.37) satisfy

$$\rho_\varepsilon(q(\boldsymbol{m})) \simeq \varrho_\varepsilon, \qquad \rho'_\varepsilon(q(\boldsymbol{m})) \simeq -1, \qquad d_\varepsilon(\boldsymbol{m},\boldsymbol{n}) \simeq d_\varepsilon.$$

Combining the AMP iteration (2.20) with (2.37) yields

$$\widehat{\boldsymbol{h}}^k = \frac{1}{\sqrt{N}}\boldsymbol{G}(\boldsymbol{m}^k - \boldsymbol{m}) + \widehat{\boldsymbol{h}} + \varrho_\varepsilon(\boldsymbol{n} - \boldsymbol{n}^{k-1})$$
$$= \frac{1}{\sqrt{N}}\boldsymbol{G}(\boldsymbol{m}^{(3),k} - \boldsymbol{m}) + \widehat{\boldsymbol{h}} - \varrho_\varepsilon(\boldsymbol{n}^{(3),k-1} - \boldsymbol{n}) + \frac{1}{\sqrt{N}}\boldsymbol{G}(\boldsymbol{m}^k - \boldsymbol{m}^{(3),k}) - \varrho_\varepsilon(\boldsymbol{n}^{k-1} - \boldsymbol{n}^{(3),k-1}).$$

By Proposition 2.A.1(a)(b), $\mathbb{W}_2(\mu_{\dot{\boldsymbol{h}}^{(3),k}}, \mu_{\dot{\boldsymbol{h}}^{(1),k}}) = o_N(1)$. So, Fact 2.4.20 and Proposition 2.5.4 imply $\frac{1}{N}\langle \boldsymbol{m}, \boldsymbol{m}^{(3),k}\rangle \simeq \overline{q}_k$ (similarly to (2.98)) and

$$\frac{1}{N}\langle \dot{\boldsymbol{h}}, \boldsymbol{m}^{(3),k}\rangle \simeq \frac{1}{N}\langle \dot{\boldsymbol{h}}, \boldsymbol{m}^{(1),k}\rangle \simeq (\overline{\psi}_k + \mathbf{1}\{k \ge 1\}\varepsilon)\varrho_\varepsilon.$$

Expanding $\boldsymbol{G}$ using (2.99) then yields

$$\widehat{\boldsymbol{h}}^k = \frac{1}{\sqrt{N}}\widetilde{\boldsymbol{G}}(\boldsymbol{m}^{(3),k} - \boldsymbol{m}) + \frac{\overline{q}_k + \varepsilon}{q_\varepsilon + \varepsilon}\widehat{\boldsymbol{h}} - \varrho_\varepsilon\left(\boldsymbol{n}^{(3),k-1} - \frac{\overline{\psi}_k + \varepsilon}{\psi_\varepsilon + \varepsilon}\boldsymbol{n}\right)$$
$$+ \frac{1}{\sqrt{N}}\boldsymbol{G}(\boldsymbol{m}^k - \boldsymbol{m}^{(3),k}) - \varrho_\varepsilon(\boldsymbol{n}^{k-1} - \boldsymbol{n}^{(3),k-1}) + o(\sqrt{N})$$
$$= \widehat{\boldsymbol{h}}^{(3),k} + \frac{1}{\sqrt{N}}\boldsymbol{G}(\boldsymbol{m}^k - \boldsymbol{m}^{(3),k}) - \varrho_\varepsilon(\boldsymbol{n}^{k-1} - \boldsymbol{n}^{(3),k-1}) + o(\sqrt{N}).$$

Analogously,

$$\dot{\boldsymbol{h}}^{k+1} = \dot{\boldsymbol{h}}^{(3),k+1} + \frac{1}{\sqrt{N}}\boldsymbol{G}^\top(\boldsymbol{n}^k - \boldsymbol{n}^{(3),k}) - d_\varepsilon(\boldsymbol{m}^{k-1} - \boldsymbol{m}^{(3),k-1}) + o(\sqrt{N}).$$

So, on the high probability event that $\|\boldsymbol{G}\|_{\mathsf{op}} = O(\sqrt{N})$,

$$\|\widehat{\boldsymbol{h}}^k - \widehat{\boldsymbol{h}}^{(3),k}\| = O(1)\|\boldsymbol{m}^k - \boldsymbol{m}^{(3),k}\| + \varrho_\varepsilon\|\boldsymbol{n}^{k-1} - \boldsymbol{n}^{(3),k-1}\| + o(\sqrt{N}),$$
$$\|\dot{\boldsymbol{h}}^{k+1} - \dot{\boldsymbol{h}}^{(3),k+1}\| = O(1)\|\boldsymbol{n}^k - \boldsymbol{n}^{(3),k}\| + |d_\varepsilon|\|\boldsymbol{m}^k - \boldsymbol{m}^{(3),k}\| + o(\sqrt{N}).$$

The result follows by induction on k, like above. □

Proof of Proposition 2.5.5. Immediate from Proposition 2.A.1. □

2.A.5 Continuity of first moment functional term

Proof of Lemma 2.7.4. Let C denote an absolute constant, which may change from line by line. By Lemma 2.7.3, $\log \Psi$ is $(2,1)$-pseudo-Lipschitz. By Cauchy–Schwarz (similarly to the proof of Fact 2.4.20),

$$\left|\mathbb{E}\log\Psi\left\{\frac{\kappa - a_1\hat{\boldsymbol{H}} - b_1\boldsymbol{N}}{c_1}\right\} - \log\Psi\left\{\frac{\kappa - a_2\hat{\boldsymbol{H}} - b_2\boldsymbol{N}}{c_2}\right\}\right| \le C\sqrt{T_1T_2},$$

where

$$T_1 = \mathbb{E}\left[\left(\frac{\kappa - a_1\hat{\boldsymbol{H}} - b_1\boldsymbol{N}}{c_1} - \frac{\kappa - a_2\hat{\boldsymbol{H}} - b_2\boldsymbol{N}}{c_2}\right)^2\right]$$
$$\le C\left(\frac{\max(a_1,a_2,b_1,b_2,c_1,c_2,1)(|a_1-a_2|+|b_1-b_2|+|c_1-c_2|)}{\min(c_1,c_2)^2}\right)^2$$

and

$$T_2 = \mathbb{E}\left[\left(\frac{\kappa - a_1\hat{\boldsymbol{H}} - b_1\boldsymbol{N}}{c_1}\right)^2 + \left(\frac{\kappa - a_2\hat{\boldsymbol{H}} - b_2\boldsymbol{N}}{c_2}\right)^2 + 1\right] \le C\left(\frac{\max(a_1,a_2,b_1,b_2,c_1,c_2,1)}{\min(c_1,c_2)}\right)^4.$$

□

2.B Verification of numerical conditions for $\kappa = 0$

In this appendix, we use rigorous interval arithmetic (implemented in the accompanying Python 3 file using `python-flint`) to verify the conditions in Theorem 2.3.6, other than Condition 2.1.3, at $\kappa = 0$. This proves Theorem 2.1.2. We also verify Claim 2.2.6 using interval arithmetic.

Throughout this section we take $\kappa = 0$, $\alpha_\star = \alpha_\star(0)$, $q_0 = q_\star(\alpha_\star, 0)$, and $\psi_0 = \psi_\star(\alpha_\star, 0)$. We will use Claims to denote statements whose proofs require interval arithmetic.

2.B.1 Numerical estimates of parameters and special functions

By [DS18, §7], the following are lower and upper bounds for $\alpha_\star$, q_0, ψ_0:

$$\alpha_{\rm lb} = 0.833078599, \qquad q_{\rm lb} = 0.56394907949, \qquad \psi_{\rm lb} = 2.5763513100,$$
$$\alpha_{\rm ub} = 0.833078600, \qquad q_{\rm ub} = 0.56394908030, \qquad \psi_{\rm ub} = 2.5763513224.$$

Let $\gamma_0 = \frac{q_0}{1-q_0}$, $\gamma_{\rm lb} = \frac{q_{\rm lb}}{1-q_{\rm lb}}$ and $\gamma_{\rm ub} = \frac{q_{\rm ub}}{1-q_{\rm ub}}$. Note that Condition 2.3.4 only requires us to exhibit a value of $z > -1$ such that $\lambda(z) < 0$. In the verification below we will use the value

$$\widehat{z} = -0.669316.$$

For $k \in \{2, 4\}$, define

$$p_k(\psi) = \mathbb{E}[\mathrm{th}(\psi^{1/2}Z)^k], \qquad r_k(\gamma) = \mathbb{E}[\mathcal{E}(\gamma^{1/2}Z)^k].$$

Note that the fixed-point condition in Condition 2.3.1 defining (q_0, ψ_0) implies (for $\kappa = 0$)

$$p_2(\psi_0) = q_0, \qquad r_2(\gamma_0) = \frac{(1-q_0)\psi_0}{\alpha_\star}. \tag{2.100}$$

Let

$$m(z,\psi) = \mathbb{E}[(z + \mathrm{ch}^2(\psi^{1/2}Z))^{-1}]. \tag{2.101}$$

Finally, define

$$g(m,q,\gamma) = \mathbb{E}\left\{\frac{\mathcal{E}'(\gamma^{1/2}Z)}{(1-q)(1-\mathcal{E}'(\gamma^{1/2}Z)) + m\mathcal{E}'(\gamma^{1/2}Z)}\right\}. \tag{2.102}$$

We now collect the main estimates in the verification whose proofs require computer assistance. The proofs of these claims are deferred to Appendix 2.B.4, with computer-assisted parts carried out in the accompanying Python file.

Claim 2.B.1. *We have* $p_4(\psi_0) \in [p_{4,\rm lb}, p_{4,\rm ub}] \equiv [0.4405902310, 0.4405902320]$.

Claim 2.B.2. *We have* $r_4(\gamma_0) \in [r_{4,\mathrm{lb}}, r_{4,\mathrm{ub}}] \equiv [5.297, 5.317]$.

Claim 2.B.3. *We have* $m(\widehat{z}) \le m_{\mathrm{ub}} \equiv 0.9309695$, *where* $m(z) = m(z, \psi_0)$ *is defined in Condition 2.3.4.*

Claim 2.B.4. *We have* $g(m(\widehat{z}), q_0, \gamma_0) \ge g_{\mathrm{lb}} \equiv 0.7739$.

We conclude this preparatory subsection with a few useful lemmas. First, we reduce several integrals that will appear below to the functions p_2, p_4, r_2, r_4.

Lemma 2.B.5. *The following identities hold.*

$$\mathfrak{t}(\psi) \equiv \mathbb{E}[\mathrm{th}'(\psi_0^{1/2} Z)^2] = 1 - 2p_2(\psi) + p_4(\psi), \tag{2.103}$$

$$\mathfrak{s}_1(\gamma) \equiv \mathbb{E}\left\{\mathcal{E}'(\gamma^{1/2} Z)\right\} = \frac{r_2(\gamma)}{1+\gamma}, \tag{2.104}$$

$$\mathfrak{s}_2(\gamma) \equiv \mathbb{E}\left\{\mathcal{E}(\gamma^{1/2} Z)^2 \mathcal{E}'(\gamma^{1/2} Z)\right\} = \frac{r_4(\gamma)}{1+3\gamma}, \tag{2.105}$$

$$\mathfrak{s}_3(\gamma) \equiv \mathbb{E}\left\{\gamma^{1/2} Z \mathcal{E}(\gamma^{1/2} Z) \mathcal{E}'(\gamma^{1/2} Z)\right\} = -\frac{\gamma}{1+2\gamma} r_2(\gamma) + \frac{3\gamma}{(1+2\gamma)(1+3\gamma)} r_4(\gamma), \tag{2.106}$$

$$\mathfrak{s}_4(\gamma) \equiv \mathbb{E}\left\{(\gamma^{1/2} Z)^2 \mathcal{E}'(\gamma^{1/2} Z)\right\} = -\frac{\gamma(4\gamma^2 + \gamma - 1)}{(1+\gamma)^2(1+2\gamma)} r_2(\gamma) + \frac{6\gamma^2}{(1+\gamma)(1+2\gamma)(1+3\gamma)} r_4(\gamma), \tag{2.107}$$

$$\mathfrak{s}_5(\gamma) \equiv \mathbb{E}\left\{\mathcal{E}'(\gamma^{1/2} Z)^2\right\} = \frac{\gamma}{1+2\gamma} r_2(\gamma) + \frac{1-\gamma}{(1+2\gamma)(1+3\gamma)} r_4(\gamma). \tag{2.108}$$

Proof. Equation (2.103) follows directly from the identity

$$\mathrm{th}'(x)^2 = (1 - \mathrm{th}^2(x))^2 = 1 - 2\mathrm{th}^2(x) + \mathrm{th}^4(x).$$

For the remaining parts, we apply the identity $\mathcal{E}'(x) = \mathcal{E}(x)(\mathcal{E}(x) - x)$ (Lemma 2.4.21(b)) and integrate by parts. First,

$$\begin{aligned}\mathfrak{s}_1(\gamma) &= \mathbb{E}\left\{\mathcal{E}(\gamma^{1/2} Z)^2\right\} - \mathbb{E}\left\{\mathcal{E}(\gamma^{1/2} Z)\gamma^{1/2} Z\right\} \\ &= \mathbb{E}\left\{\mathcal{E}(\gamma^{1/2} Z)^2\right\} - \gamma\,\mathbb{E}\left\{\mathcal{E}'(\gamma^{1/2} Z)\right\} = r_2(\gamma) - \gamma\mathfrak{s}_1(\gamma),\end{aligned}$$

which proves (2.104). Similarly,

$$\mathfrak{s}_2(\gamma) = \mathbb{E}\left\{\mathcal{E}(\gamma^{1/2} Z)^4\right\} - \mathbb{E}\left\{\mathcal{E}(\gamma^{1/2} Z)^3 \gamma^{1/2} Z\right\} = r_4(\gamma) - 3\gamma\mathfrak{s}_2(\gamma),$$

which proves (2.105). Then,

$$\begin{aligned}\mathfrak{s}_3(\gamma) &= \mathbb{E}\left\{\gamma^{1/2} Z \mathcal{E}(\gamma^{1/2} Z)^3\right\} - \mathbb{E}\left\{(\gamma^{1/2} Z)^2 \mathcal{E}(\gamma^{1/2} Z)^2\right\} \\ &= 3\gamma\,\mathbb{E}\left\{\mathcal{E}(\gamma^{1/2} Z)^2 \mathcal{E}'(\gamma^{1/2} Z)\right\} - \gamma\,\mathbb{E}\left\{\mathcal{E}(\gamma^{1/2} Z)^2\right\} - 2\gamma\,\mathbb{E}\left\{(\gamma^{1/2} Z)\mathcal{E}(\gamma^{1/2} Z)\mathcal{E}'(\gamma^{1/2} Z)\right\} \\ &= 3\gamma\mathfrak{s}_2(\gamma) - \gamma r_2(\gamma) - 2\gamma\mathfrak{s}_3(\gamma).\end{aligned}$$

Rearranging proves (2.106). Further,

$$\begin{aligned}\mathfrak{s}_4(\gamma) &= \mathbb{E}\left\{(\gamma^{1/2} Z)^2 \mathcal{E}(\gamma^{1/2} Z)^2\right\} - \mathbb{E}\left\{(\gamma^{1/2} Z)^3 \mathcal{E}(\gamma^{1/2} Z)\right\} \\ &= \gamma\,\mathbb{E}\left\{\mathcal{E}(\gamma^{1/2} Z)^2\right\} + 2\gamma\,\mathbb{E}\left\{(\gamma^{1/2} Z)\mathcal{E}(\gamma^{1/2} Z)\mathcal{E}'(\gamma^{1/2} Z)\right\} \\ &\quad - 2\gamma\,\mathbb{E}\left\{(\gamma^{1/2} Z)\mathcal{E}(\gamma^{1/2} Z)\right\} - \gamma\,\mathbb{E}\left\{(\gamma^{1/2} Z)^2 \mathcal{E}'(\gamma^{1/2} Z)\right\}.\end{aligned}$$

Integrating by parts again yields

$$\mathbb{E}\left\{(\gamma^{1/2} Z)\mathcal{E}(\gamma^{1/2} Z)\right\} = \gamma\,\mathbb{E}\left\{\mathcal{E}'(\gamma^{1/2} Z)\right\} = \gamma\mathfrak{s}_1(\gamma).$$

So

$$\mathfrak{s}_4(\gamma) = \gamma r_2(\gamma) + 2\gamma\mathfrak{s}_3(\gamma) - 2\gamma^2\mathfrak{s}_1(\gamma) - \gamma\mathfrak{s}_4(\gamma).$$

Rearranging proves (2.107). Finally,

$$\begin{aligned}
\mathfrak{s}_5(\gamma) &= \mathbb{E}\left\{\mathcal{E}(\gamma^{1/2}Z)^4\right\} - 2\,\mathbb{E}\left\{(\gamma^{1/2}Z)\mathcal{E}(\gamma^{1/2}Z)^3\right\} + \mathbb{E}\left\{(\gamma^{1/2}Z)^2\mathcal{E}(\gamma^{1/2}Z)^2\right\} \\
&= \mathbb{E}\left\{\mathcal{E}(\gamma^{1/2}Z)^4\right\} - 6\gamma\,\mathbb{E}\left\{\mathcal{E}(\gamma^{1/2}Z)^2\mathcal{E}'(\gamma^{1/2}Z)\right\} + \gamma\,\mathbb{E}\left\{\mathcal{E}(\gamma^{1/2}Z)^2\right\} \\
&\quad + 2\gamma\,\mathbb{E}\left\{(\gamma^{1/2}Z)\mathcal{E}(\gamma^{1/2}Z)\mathcal{E}'(\gamma^{1/2}Z)\right\} \\
&= r_4(\gamma) - 6\gamma\mathfrak{s}_2(\gamma) + \gamma r_2(\gamma) + 2\gamma\mathfrak{s}_3(\gamma).
\end{aligned}$$

Rearranging proves (2.108). □

Recall from Condition 2.3.4 that $d_0 = \alpha_\star\,\mathbb{E}[F'_{1-q_0}(q_0^{1/2}Z)]$. As a consequence of (2.100) and (2.104), we have

$$d_0 = -\frac{\alpha_\star}{1-q_0}\mathfrak{s}_1(\gamma_0) = -\frac{\alpha_\star}{1-q_0}\cdot\frac{r_2(\gamma_0)}{1+\gamma_0} = -(1-q_0)\psi_0, \tag{2.109}$$

where we have used that $(1-q_0)(1+\gamma_0) = 1$.

Lemma 2.B.6. *The functions p_4 and r_4 are increasing. Moreover, for any $z > -1$, and m defined in (2.101), the function $\psi \mapsto m(z,\psi)$ is decreasing.*

Proof. The function p_4 is increasing simply because the maps $\psi \mapsto \mathrm{th}(\psi^{1/2}x)^4$ are pointwise increasing for all $x \in \mathbb{R}$. Similarly, since the maps $\psi \mapsto (z + \mathrm{ch}^2(\psi^{1/2}x))^{-1}$ are pointwise increasing for all $x \in \mathbb{R}$, $z > -1$, the function $\psi \mapsto m(z,\psi)$ is decreasing. Finally,

$$r_4'(\gamma) = \mathbb{E}\left\{6\mathcal{E}(\gamma^{1/2}Z)^2\mathcal{E}'(\gamma^{1/2}Z)^2 + 2\mathcal{E}(\gamma^{1/2}Z)^3\mathcal{E}''(\gamma^{1/2}Z)\right\} \geq 0,$$

as Lemma 2.4.21(c) implies $\mathcal{E}'' > 0$. Thus r_4 is increasing. □

2.B.2 Verification of numerical conditions in Theorem 2.3.6

Condition 2.3.1 was proved in [DS18, Proposition 1.3] (recorded as Proposition 2.3.2). We now verify Conditions 2.3.3 and 2.3.4 by proving the following.

Claim 2.B.7. *Condition 2.3.3 holds for $\kappa = 0$, with $\alpha_\star\,\mathbb{E}[\mathrm{th}'(\psi_0^{1/2}Z)^2]\,\mathbb{E}[F'_{1-q_0}(q_0^{1/2}Z)^2] \leq a_{\mathrm{ub}} \equiv 0.5446$.*

Proof. We calculate:

$$\begin{aligned}
\alpha_\star\,\mathbb{E}[\mathrm{th}'(\psi_0^{1/2}Z)^2]\,\mathbb{E}[F'_{1-q_0}(q_0^{1/2}Z)^2] &= \frac{\alpha_\star}{(1-q_0)^2}\mathfrak{t}(\psi_0)\mathfrak{s}_5(\gamma_0) \\
&\overset{Lem.\ 2.B.5}{=} \frac{\alpha_\star}{(1-q_0)^2}(1 - 2p_2(\psi) + p_4(\psi))\left(\frac{\gamma_0}{1+2\gamma_0}r_2(\gamma_0) + \frac{1-\gamma_0}{(1+2\gamma_0)(1+3\gamma_0)}r_4(\gamma_0)\right) \\
&\overset{(2.100)}{=} \frac{\alpha_\star}{(1-q_0)^2}(1 - 2q_0 + p_4(\psi))\left(\frac{\gamma_0}{1+2\gamma_0}\cdot\frac{(1-q_0)\psi_0}{\alpha_\star} + \frac{1-\gamma_0}{(1+2\gamma_0)(1+3\gamma_0)}r_4(\gamma_0)\right) \\
&= (1 - 2q_0 + p_4(\psi))\left(\frac{\gamma_0\psi_0}{1+q_0} + \frac{\alpha_\star(1-\gamma_0)}{(1+q_0)(1+2q_0)}r_4(\gamma_0)\right) \\
&\leq (1 - 2q_{\mathrm{lb}} + p_{4,\mathrm{ub}})\left(\frac{\gamma_{\mathrm{ub}}\psi_{\mathrm{ub}}}{1+q_{\mathrm{lb}}} + \frac{\alpha_{\mathrm{lb}}(1-\gamma_{\mathrm{lb}})}{(1+q_{\mathrm{ub}})(1+2q_{\mathrm{ub}})}r_{4,\mathrm{lb}}\right) \overset{(*)}{\leq} a_{\mathrm{ub}}.
\end{aligned}$$

The estimate $(*)$ is verified in the accompanying Python file. We note that this is a simple arithmetic comparison, as all terms are explicitly defined decimal numbers. □

Claim 2.B.8. *Condition 2.3.4 holds for $\kappa = 0$, with $\lambda(\widehat{z}) \leq \lambda_{\mathrm{ub}} \equiv -0.1906$.*

Proof. Note that for g defined in (2.102),

$$\begin{aligned}\lambda(\widehat{z}) &= \widehat{z} - \alpha_\star g(m(\widehat{z}), q_0, \gamma_0) - d_0 \overset{(2.109)}{=} \widehat{z} - \alpha_\star g(m(\widehat{z}), q_0, \gamma_0) + (1-q_0)\psi_0 \\ &\le \widehat{z} - \alpha_{\mathrm{lb}} g_{\mathrm{lb}} + (1-q_{\mathrm{lb}})\psi_{\mathrm{ub}} \overset{(*)}{\le} \lambda_{\mathrm{ub}}.\end{aligned}$$

The step $(*)$ is verified in the accompanying Python file, and is a simple arithmetic comparison of explicitly defined decimal numbers. ☐

Proof of Theorem 2.1.2. Follows from Theorem 2.3.6, Proposition 2.3.2, and Claims 2.B.7 and 2.B.8. ☐

2.B.3 Local maximality of first moment functional at $(1,0)$

We next verify Claim 2.2.6.

Lemma 2.B.9. *For $\kappa = 0$, we have*

$$\begin{aligned}\langle \nabla^2 \overline{\mathscr{S}}_\star(1,0), (u_1,u_2)^{\otimes 2}\rangle = &-\mathbb{E}[(1-\boldsymbol{M}^2)(u_1\dot{\boldsymbol{H}} + u_2\boldsymbol{M})^2] + C_1\,\mathbb{E}[(1-\boldsymbol{M}^2)(u_1\dot{\boldsymbol{H}} + u_2\boldsymbol{M})\dot{\boldsymbol{H}}]^2 \\ &+ C_2\,\mathbb{E}[(1-\boldsymbol{M}^2)(u_1\dot{\boldsymbol{H}} + u_2\boldsymbol{M})\boldsymbol{M}]\,\mathbb{E}[(1-\boldsymbol{M}^2)(u_1\dot{\boldsymbol{H}} + u_2\boldsymbol{M})\dot{\boldsymbol{H}}] \\ &+ C_3\,\mathbb{E}[(1-\boldsymbol{M}^2)(u_1\dot{\boldsymbol{H}} + u_2\boldsymbol{M})\boldsymbol{M}]^2,\end{aligned}$$

where

$$C_1 = \frac{\alpha_\star}{\psi_0^2}\,\mathbb{E}\left\{F'_{1-q_0}(\hat{\boldsymbol{H}})\boldsymbol{N}^2\right\}, \qquad C_2 = \frac{2\alpha_\star}{\psi_0}\,\mathbb{E}\left\{F'_{1-q_0}(\hat{\boldsymbol{H}})\left(\frac{1}{q_0(1-q_0)}\hat{\boldsymbol{H}} + \boldsymbol{N}\right)\boldsymbol{N}\right\} + \frac{2}{1-q_0},$$

$$C_3 = \alpha_\star\,\mathbb{E}\left\{F'_{1-q_0}(\hat{\boldsymbol{H}})\left(\frac{1}{q_0(1-q_0)}\hat{\boldsymbol{H}} + \boldsymbol{N}\right)^2\right\} + \frac{\psi_0}{q_0}.$$

Proof. Analogously to the proof of Lemma 2.2.5(c), define $\boldsymbol{\Delta}_2 = (u_1\partial_{\lambda_1} + u_2\partial_{\lambda_2})^2\boldsymbol{\Lambda}$. Also abbreviate

$$\boldsymbol{V} = \frac{\kappa - \frac{\mathbb{E}[\boldsymbol{M}\boldsymbol{\Lambda}]}{q_0}\hat{\boldsymbol{H}} - \frac{\mathbb{E}[\dot{\boldsymbol{H}}\boldsymbol{\Lambda}]}{\psi_0}\boldsymbol{N}}{\sqrt{1-\frac{\mathbb{E}[\boldsymbol{M}\boldsymbol{\Lambda}]^2}{q_0}}} + \sqrt{1-q_0}\boldsymbol{N}.$$

We differentiate (2.87) to obtain

$$\begin{aligned}&\langle \overline{\mathscr{S}}_\star(\lambda_1,\lambda_2), (u_1,u_2)^{\otimes 2}\rangle = -\mathbb{E}[(u_1\dot{\boldsymbol{H}} + u_2\boldsymbol{M})\boldsymbol{\Delta}] \\ &- \alpha_\star\,\mathbb{E}\left\{\mathcal{E}'(\boldsymbol{V})\left(\frac{-\frac{\mathbb{E}[\boldsymbol{M}\boldsymbol{\Delta}]}{q_0}\hat{\boldsymbol{H}} - \frac{\mathbb{E}[\dot{\boldsymbol{H}}\boldsymbol{\Delta}]}{\psi_0}\boldsymbol{N}}{\sqrt{1-\frac{\mathbb{E}[\boldsymbol{M}\boldsymbol{\Lambda}]^2}{q_0}}} + \frac{\kappa - \frac{\mathbb{E}[\boldsymbol{M}\boldsymbol{\Lambda}]}{q_0}\hat{\boldsymbol{H}} - \frac{\mathbb{E}[\dot{\boldsymbol{H}}\boldsymbol{\Lambda}]}{\psi_0}\boldsymbol{N}}{\left(1-\frac{\mathbb{E}[\boldsymbol{M}\boldsymbol{\Lambda}]^2}{q_0}\right)^{3/2}} \cdot \frac{\mathbb{E}[\boldsymbol{M}\boldsymbol{\Lambda}]\,\mathbb{E}[\boldsymbol{M}\boldsymbol{\Delta}]}{q_0}\right)^2\right\} \\ &- \alpha_\star\,\mathbb{E}\left\{\mathcal{E}(\boldsymbol{V})\left(\frac{-\frac{2\,\mathbb{E}[\boldsymbol{M}\boldsymbol{\Delta}]}{q_0}\hat{\boldsymbol{H}} - \frac{2\,\mathbb{E}[\dot{\boldsymbol{H}}\boldsymbol{\Delta}]}{\psi_0}\boldsymbol{N}}{\left(1-\frac{\mathbb{E}[\boldsymbol{M}\boldsymbol{\Lambda}]^2}{q_0}\right)^{3/2}} \cdot \frac{\mathbb{E}[\boldsymbol{M}\boldsymbol{\Lambda}]\,\mathbb{E}[\boldsymbol{M}\boldsymbol{\Delta}]}{q_0}\right.\right. \\ &+ \left.\left.\frac{\kappa - \frac{\mathbb{E}[\boldsymbol{M}\boldsymbol{\Lambda}]}{q_0}\hat{\boldsymbol{H}} - \frac{\mathbb{E}[\dot{\boldsymbol{H}}\boldsymbol{\Lambda}]}{\psi_0}\boldsymbol{N}}{\left(1-\frac{\mathbb{E}[\boldsymbol{M}\boldsymbol{\Lambda}]^2}{q_0}\right)^{5/2}} \cdot \frac{3\,\mathbb{E}[\boldsymbol{M}\boldsymbol{\Lambda}]^2\,\mathbb{E}[\boldsymbol{M}\boldsymbol{\Delta}]^2}{q_0^2} + \frac{\kappa - \frac{\mathbb{E}[\boldsymbol{M}\boldsymbol{\Lambda}]}{q_0}\hat{\boldsymbol{H}} - \frac{\mathbb{E}[\dot{\boldsymbol{H}}\boldsymbol{\Lambda}]}{\psi_0}\boldsymbol{N}}{\left(1-\frac{\mathbb{E}[\boldsymbol{M}\boldsymbol{\Lambda}]^2}{q_0}\right)^{3/2}} \cdot \frac{\mathbb{E}[\boldsymbol{M}\boldsymbol{\Delta}]^2}{q_0}\right)\right\} + f(\boldsymbol{\Delta}_2),\end{aligned}$$

where $f(\boldsymbol{\Delta}_2)$ is (2.87) with $\boldsymbol{\Delta}$ replaced by $\boldsymbol{\Delta}_2$. We now specialize to $(\lambda_1,\lambda_2) = (1,0)$. As argued in the proof of Lemma 2.2.5(c), at $(\lambda_1,\lambda_2)=(1,0)$ we have $f(\boldsymbol{\Delta}_2)=0$. So,

$$\begin{aligned}
&\langle \overline{\mathscr{S}}_\star(1,0),(u_1,u_2)^{\otimes 2}\rangle = -\mathbb{E}[(u_1\dot{\boldsymbol{H}}+u_2\boldsymbol{M})\boldsymbol{\Delta}] \\
&+\alpha_\star\,\mathbb{E}\left\{F'_{1-q_0}(\hat{\boldsymbol{H}})\left(-\frac{\mathbb{E}[\boldsymbol{M}\boldsymbol{\Delta}]}{q_0}\hat{\boldsymbol{H}}-\frac{\mathbb{E}[\dot{\boldsymbol{H}}\boldsymbol{\Delta}]}{\psi_0}\boldsymbol{N}+\frac{\kappa-\hat{\boldsymbol{H}}-(1-q_0)\boldsymbol{N}}{1-q_0}\cdot\mathbb{E}[\boldsymbol{M}\boldsymbol{\Delta}]\right)^2\right\} \\
&-\alpha_\star\,\mathbb{E}\left\{F_{1-q_0}(\hat{\boldsymbol{H}})\left(\frac{-\frac{2\,\mathbb{E}[\boldsymbol{M}\boldsymbol{\Delta}]}{q_0}\hat{\boldsymbol{H}}-\frac{2\,\mathbb{E}[\dot{\boldsymbol{H}}\boldsymbol{\Delta}]}{\psi_0}\boldsymbol{N}}{1-q_0}\cdot\mathbb{E}[\boldsymbol{M}\boldsymbol{\Delta}]\right.\right. \\
&\left.\left.+\frac{\kappa-\hat{\boldsymbol{H}}-(1-q_0)\boldsymbol{N}}{(1-q_0)^2}\cdot 3\,\mathbb{E}[\boldsymbol{M}\boldsymbol{\Delta}]^2+\frac{\kappa-\hat{\boldsymbol{H}}-(1-q_0)\boldsymbol{N}}{1-q_0}\cdot\frac{\mathbb{E}[\boldsymbol{M}\boldsymbol{\Delta}]^2}{q_0}\right)\right\}.
\end{aligned}$$

Specializing further to $\kappa=0$ (which was not used up to here),

$$\begin{aligned}
&\langle \overline{\mathscr{S}}_\star(1,0),(u_1,u_2)^{\otimes 2}\rangle \\
&=-\mathbb{E}[(u_1\dot{\boldsymbol{H}}+u_2\boldsymbol{M})\boldsymbol{\Delta}]+\alpha_\star\,\mathbb{E}\left\{F'_{1-q_0}(\hat{\boldsymbol{H}})\left(\left(\frac{1}{q_0(1-q_0)}\hat{\boldsymbol{H}}+\boldsymbol{N}\right)\mathbb{E}[\boldsymbol{M}\boldsymbol{\Delta}]+\frac{\boldsymbol{N}}{\psi_0}\,\mathbb{E}[\dot{\boldsymbol{H}}\boldsymbol{\Delta}]\right)^2\right\} \\
&+\alpha_\star\,\mathbb{E}\left\{\boldsymbol{N}\left(\left(\frac{3}{q_0(1-q_0)^2}\hat{\boldsymbol{H}}+\frac{1+2q_0}{q_0(1-q_0)}\boldsymbol{N}\right)\mathbb{E}[\boldsymbol{M}\boldsymbol{\Delta}]^2+\frac{2}{\psi_0(1-q_0)}\boldsymbol{N}\,\mathbb{E}[\boldsymbol{M}\boldsymbol{\Delta}]\,\mathbb{E}[\dot{\boldsymbol{H}}\boldsymbol{\Delta}]\right)\right\}
\end{aligned}$$

Finally, as $\alpha_\star\,\mathbb{E}[\boldsymbol{N}\hat{\boldsymbol{H}}]=q_0d_0=-q_0(1-q_0)\psi_0$ (by (2.109)) and $\alpha_\star\,\mathbb{E}[\boldsymbol{N}^2]=\psi_0$, the last term simplifies to

$$\frac{\psi_0}{q_0}\,\mathbb{E}[\boldsymbol{M}\boldsymbol{\Delta}]^2+\frac{2}{1-q_0}\,\mathbb{E}[\boldsymbol{M}\boldsymbol{\Delta}]\,\mathbb{E}[\dot{\boldsymbol{H}}\boldsymbol{\Delta}].$$

Expanding $\boldsymbol{\Delta}=(1-\boldsymbol{M}^2)(u_1\dot{\boldsymbol{H}}+u_2\boldsymbol{M})$ concludes the proof. □

Claim 2.B.10. *The following estimates hold.*

(a) $C_1\in[C_{1,\mathrm{lb}},C_{1,\mathrm{ub}}]\equiv[-0.7193,-0.7165]$.

(b) $C_2\in[C_{2,\mathrm{lb}},C_{2,\mathrm{ub}}]\equiv[5.0439,5.0568]$.

(c) $C_3\in[C_{3,\mathrm{lb}},C_{3,\mathrm{ub}}]\equiv[1.1345,1.1526]$.

Proof. We compute using Lemma 2.B.5 and (2.100):

$$\begin{aligned}
C_1&=\frac{\alpha_\star}{\psi_0^2}\cdot\frac{-\mathfrak{s}_2(\gamma_0)}{(1-q_0)^2}=-\frac{\alpha_\star r_4(\gamma_0)}{\psi_0^2(1-q_0)^2(1+3\gamma_0)}=-\frac{\alpha_\star r_4(\gamma_0)}{\psi_0^2(1-q_0)(1+2q_0)},\\
C_2&=\frac{2\alpha_\star}{\psi_0(1-q_0)^2}\left(-\mathfrak{s}_2(\gamma_0)+\frac{\mathfrak{s}_3(\gamma_0)}{q_0}\right)+\frac{2}{1-q_0}\\
&=\frac{2\alpha_\star}{\psi_0(1-q_0)^2}\left(\frac{(2-q_0)(1-q_0)r_4(\gamma_0)}{(1+q_0)(1+2q_0)}-\frac{r_2(\gamma_0)}{1+q_0}\right)+\frac{2}{1-q_0}=\frac{2(2-q_0)\alpha_\star r_4(\gamma_0)}{\psi_0(1-q_0^2)(1+2q_0)}+\frac{2q_0}{1-q_0^2},\\
C_3&=-\frac{\alpha_\star}{(1-q_0)^2}\left(\mathfrak{s}_2(\gamma_0)-\frac{2\mathfrak{s}_3(\gamma_0)}{q_0}+\frac{\mathfrak{s}_4(\gamma_0)}{q_0^2}\right)+\frac{\psi_0}{q_0}\\
&=-\frac{\alpha_\star}{(1-q_0)^2}\left(\frac{1-q_0}{1+2q_0}r_4(\gamma_0)-\frac{2q_0-1}{q_0}r_2(\gamma_0)\right)+\frac{\psi_0}{q_0}=-\frac{\alpha_\star r_4(\gamma_0)}{(1-q_0)(1+2q_0)}+\frac{\psi_0}{1-q_0}.
\end{aligned}$$

So

$$\begin{aligned}
C_{1,\mathrm{lb}}&\overset{(*)}{\le}-\frac{\alpha_{\mathrm{ub}}r_{4,\mathrm{ub}}}{\psi_{\mathrm{lb}}^2(1-q_{\mathrm{ub}})(1+2q_{\mathrm{lb}})}\le C_1\le-\frac{\alpha_{\mathrm{lb}}r_{4,\mathrm{lb}}}{\psi_{\mathrm{ub}}^2(1-q_{\mathrm{lb}})(1+2q_{\mathrm{ub}})}\overset{(*)}{\le}C_{1,\mathrm{ub}},\\
C_{2,\mathrm{lb}}&\overset{(*)}{\le}\frac{2(2-q_{\mathrm{ub}})\alpha_{\mathrm{lb}}r_{4,\mathrm{lb}}}{\psi_{\mathrm{ub}}(1-q_{\mathrm{lb}}^2)(1+2q_{\mathrm{ub}})}+\frac{2q_{\mathrm{lb}}}{1-q_{\mathrm{lb}}^2}\le C_2\le\frac{2(2-q_{\mathrm{lb}})\alpha_{\mathrm{ub}}r_{4,\mathrm{ub}}}{\psi_{\mathrm{lb}}(1-q_{\mathrm{ub}}^2)(1+2q_{\mathrm{lb}})}+\frac{2q_{\mathrm{ub}}}{1-q_{\mathrm{ub}}^2}\overset{(*)}{\le}C_{2,\mathrm{ub}},\\
C_{3,\mathrm{lb}}&\overset{(*)}{\le}-\frac{\alpha_{\mathrm{ub}}r_{4,\mathrm{ub}}}{(1-q_{\mathrm{ub}})(1+2q_{\mathrm{lb}})}+\frac{\psi_{\mathrm{lb}}}{1-q_{\mathrm{lb}}}\le C_3\le-\frac{\alpha_{\mathrm{lb}}r_{4,\mathrm{lb}}}{(1-q_{\mathrm{lb}})(1+2q_{\mathrm{ub}})}+\frac{\psi_{\mathrm{ub}}}{1-q_{\mathrm{ub}}}\overset{(*)}{\le}C_{3,\mathrm{ub}}.
\end{aligned}$$

The steps marked $(*)$ are verified in the accompanying Python file, and are simple arithmetic comparisons of explicitly defined decimal numbers. $\square$

Claim 2.B.11. *Define* $I_1 = \mathbb{E}[(1-\boldsymbol{M}^2)\dot{\boldsymbol{H}}^2]$, $I_2 = \mathbb{E}[(1-\boldsymbol{M}^2)\dot{\boldsymbol{H}}\boldsymbol{M}]$, $I_3 = \mathbb{E}[(1-\boldsymbol{M}^2)\boldsymbol{M}^2]$. *Then,*

(a) $I_1 \in [I_{1,\mathrm{lb}}, I_{1,\mathrm{ub}}] \equiv [0.24759912, 0.24759923]$.

(b) $I_2 \in [I_{2,\mathrm{lb}}, I_{2,\mathrm{ub}}] \equiv [0.16997315, 0.16997318]$.

(c) $I_3 \in [I_{3,\mathrm{lb}}, I_{3,\mathrm{ub}}] \equiv [0.12335884, 0.12335885]$.

Proof. By repeated integration by parts and (2.100):

$$I_1 = \psi_0(1-q_0) - 2\psi_0^2(1-4q_0+3p_4(\psi_0)), \qquad I_2 = \psi_0(1-4q_0+3p_4(\psi_0)), \qquad I_3 = q_0 - p_4(\psi_0).$$

Thus

$$\begin{gathered}
I_{1,\mathrm{lb}} \overset{(*)}{\le} \psi_{\mathrm{lb}}(1-q_{\mathrm{ub}}) - 2\psi_{\mathrm{ub}}^2(1-4q_{\mathrm{lb}}+3p_{4,\mathrm{ub}}) \le I_1 \le \psi_{\mathrm{ub}}(1-q_{\mathrm{lb}}) - 2\psi_{\mathrm{lb}}^2(1-4q_{\mathrm{ub}}+3p_{4,\mathrm{lb}}) \overset{(*)}{\le} I_{1,\mathrm{ub}}, \\
I_{2,\mathrm{lb}} \overset{(*)}{\le} \psi_{\mathrm{lb}}(1-4q_{\mathrm{ub}}+3p_{4,\mathrm{lb}}) \le I_2 \le \psi_{\mathrm{ub}}(1-4q_{\mathrm{lb}}+3p_{4,\mathrm{ub}}) \overset{(*)}{\le} I_{2,\mathrm{ub}}, \\
I_{3,\mathrm{lb}} \overset{(*)}{\le} q_{\mathrm{lb}} - p_{4,\mathrm{ub}} \le I_3 \le q_{\mathrm{ub}} - p_{4,\mathrm{lb}} \overset{(*)}{\le} I_{3,\mathrm{ub}}.
\end{gathered}$$

The steps marked $(*)$ are verified in the accompanying Python file, and are simple arithmetic comparisons of explicitly defined decimal numbers. $\square$

Claim 2.B.12. *Let* $M = \nabla^2\overline{\mathscr{S}}_\star(1,0)$. *The following estimates hold.*

(a) $M_{1,1} \le M_{1,1,\mathrm{ub}} \equiv -0.045408$.

(b) $M_{2,2} \le M_{2,2,\mathrm{ub}} \equiv -0.020490$.

(c) $M_{1,2} \in [M_{1,2,\mathrm{lb}}, M_{1,2,\mathrm{ub}}] \equiv [-0.025685, -0.026567]$.

(d) $\det(M) \ge M_{\det,\mathrm{lb}} \equiv 0.0002246$.

Proof. By Lemma 2.B.9,

$$\begin{aligned}
M_{1,1} &= -I_1 + C_1 I_1^2 + C_2 I_1 I_2 + C_3 I_2^2, \\
M_{1,2} &= -I_2 + C_1 I_1 I_2 + \frac{1}{2} C_2 (I_2^2 + I_1 I_3) + C_3 I_2 I_3, \\
M_{2,2} &= -I_3 + C_1 I_2^2 + C_2 I_2 I_3 + C_3 I_3^2.
\end{aligned}$$

Estimating with Claims 2.B.10 and 2.B.11, we find

$$\begin{aligned}
M_{1,1} &\le -I_{1,\mathrm{lb}} + C_{1,\mathrm{ub}} I_{1,\mathrm{lb}}^2 + C_{2,\mathrm{ub}} I_{1,\mathrm{ub}} I_{2,\mathrm{ub}} + C_{3,\mathrm{ub}} I_{2,\mathrm{ub}}^2 \overset{(*)}{\le} M_{1,1,\mathrm{ub}}, \\
M_{2,2} &\le -I_{3,\mathrm{lb}} + C_{1,\mathrm{ub}} I_{2,\mathrm{lb}}^2 + C_{2,\mathrm{ub}} I_{2,\mathrm{ub}} I_{3,\mathrm{ub}} + C_{3,\mathrm{ub}} I_{3,\mathrm{ub}}^2 \overset{(*)}{\le} M_{2,2,\mathrm{ub}}, \\
M_{1,2} &\le -I_{2,\mathrm{lb}} + C_{1,\mathrm{ub}} I_{1,\mathrm{lb}} I_{2,\mathrm{lb}} + \frac{1}{2} C_{2,\mathrm{ub}} (I_{2,\mathrm{ub}}^2 + I_{1,\mathrm{ub}} I_{3,\mathrm{ub}}) + C_{3,\mathrm{ub}} I_{2,\mathrm{ub}} I_{3,\mathrm{ub}} \overset{(*)}{\le} M_{1,2,\mathrm{ub}}, \\
M_{1,2} &\ge -I_{2,\mathrm{ub}} + C_{1,\mathrm{lb}} I_{1,\mathrm{ub}} I_{2,\mathrm{ub}} + \frac{1}{2} C_{2,\mathrm{lb}} (I_{2,\mathrm{lb}}^2 + I_{1,\mathrm{lb}} I_{3,\mathrm{lb}}) + C_{3,\mathrm{lb}} I_{2,\mathrm{lb}} I_{3,\mathrm{lb}} \overset{(*)}{\ge} M_{1,2,\mathrm{lb}}.
\end{aligned}$$

The steps marked $(*)$ are verified in the accompanying Python file, and are simple arithmetic comparisons of explicitly defined decimal numbers. This proves parts (a), (b), and (c). Finally,

$$\det(M) = M_{1,1} M_{2,2} - M_{1,2}^2 \ge M_{1,1,\mathrm{ub}} M_{2,2,\mathrm{ub}} - M_{1,2,\mathrm{lb}}^2 \overset{(*)}{\ge} M_{\det,\mathrm{lb}},$$

where the step $(*)$ is verified in the accompanying Python file. This proves part (d). $\square$

Proof of Claim 2.2.6. Follows from Claim 2.B.12, which implies $M_{1,1}, M_{2,2} < 0$ and $\det(M) > 0$. $\square$

2.B.4 Interval arithmetic estimates

We now describe the computer-assisted proofs of Claims 2.B.1, 2.B.2, 2.B.3, and 2.B.4. We begin with the more straightforward Claims 2.B.1 and 2.B.3.

Proof of Claim 2.B.1. We first show the upper bound. Set $L = 10$. Since th^4 takes values in $[0, 1]$,

$$\begin{aligned} p_4(\psi_0) &\overset{Lem.\ 2.B.6}{\leq} p_4(\psi_{\mathrm{ub}}) \leq \mathbb{E}[\mathrm{th}^4(\psi_{\mathrm{ub}}^{1/2} Z)\mathbf{1}\{|Z| \leq L\}] + \mathbb{P}[|Z| \geq L] \\ &\leq \int_{-L}^{L} \mathrm{th}^4(\psi_{\mathrm{ub}}^{1/2} x)\varphi(x)\ \mathsf{d}x + 2e^{-L^2/2} \overset{(*)}{\leq} p_{4,\mathrm{ub}}, \end{aligned}$$

where the step $(*)$ is verified in the accompanying Python file. Similarly,

$$p_4(\psi_0) \overset{Lem.\ 2.B.6}{\geq} p_4(\psi_{\mathrm{lb}}) \geq \mathbb{E}[\mathrm{th}^4(\psi_{\mathrm{lb}}^{1/2} Z)\mathbf{1}\{|Z| \leq L\}] = \int_{-L}^{L} \mathrm{th}^4(\psi_{\mathrm{lb}}^{1/2} x)\varphi(x)\ \mathsf{d}x \overset{(*)}{\geq} p_{4,\mathrm{lb}},$$

where the step $(*)$ is verified in the accompanying Python file. □

Proof of Claim 2.B.3. Let $L = 10$. Note that for any $x \in \mathbb{R}$, $(\widehat{z} + \mathrm{ch}^2(x))^{-1} \leq (1 + \widehat{z})^{-1}$. Then,

$$\begin{aligned} m(\widehat{z}) = m(\widehat{z}, \psi_0) &\overset{Lem.\ 2.B.6}{\leq} m(\widehat{z}, \psi_{\mathrm{lb}}) \leq \mathbb{E}[(\widehat{z} + \mathrm{ch}^2(\psi_{\mathrm{lb}}^{1/2} Z))^{-1}\mathbf{1}\{|Z| \leq L\}] + (1 + \widehat{z})^{-1}\,\mathbb{P}[|Z| \geq L] \\ &\leq \int_{-L}^{L} (\widehat{z} + \mathrm{ch}^2(\psi_{\mathrm{lb}}^{1/2} x))^{-1}\varphi(x)\ \mathsf{d}x + 2(1 + \widehat{z})^{-1} e^{-L^2/2} \overset{(*)}{\leq} m_{\mathrm{ub}}, \end{aligned}$$

where the step $(*)$ is verified in the accompanying Python file. □

Claims 2.B.2 and 2.B.4 will involve integrating functions that involve $\mathcal{E}$ against the gaussian measure. This is more challenging because $\mathcal{E}$ is itself defined in terms of an integral, which makes these claims less amenable to numerical integration. We take a cruder approach of discretizing these integrals into small intervals, and bounding the integral on each small interval using monotonicity properties of $\mathcal{E}$ and $\mathcal{E}'$.

Proof of Claim 2.B.2. Let $L = 8$, $\delta = 10^{-3}$, and $J = L/\delta$. For integer $j \in [-J, J]$, let $x_j = j\delta$. Then,

$$r_4(\gamma_0) \overset{Lem.\ 2.B.6}{\leq} r_4(\gamma_{\mathrm{ub}}) = \sum_{j=-J}^{J-1} \mathbb{E}\left\{\mathcal{E}(\gamma_{\mathrm{ub}}^{1/2} Z)^4 \mathbf{1}\{Z \in [x_j, x_{j+1}]\}\right\} + \mathbb{E}\left\{\mathcal{E}(\gamma_{\mathrm{ub}}^{1/2} Z)^4 \mathbf{1}\{|Z| \geq L\}\right\}.$$

These terms can be bounded as follows. Since $\mathcal{E}$ is nonnegative and increasing (by Lemma 2.4.21(a)(b)),

$$\mathbb{E}\left\{\mathcal{E}(\gamma_{\mathrm{ub}}^{1/2} Z)^4 \mathbf{1}\{Z \in [x_j, x_{j+1}]\}\right\} \leq \mathcal{E}(\gamma_{\mathrm{ub}}^{1/2} x_{j+1})^4\, \mathbb{P}[Z \in [x_j, x_{j+1}]],$$

and this probability is bounded above by $\delta\varphi(x_{j+1})$ if $j \leq -1$ and $\delta\varphi(x_j)$ if $j \geq 0$. We estimate the tail term using Cauchy–Schwarz:

$$\mathbb{E}\left\{\mathcal{E}(\gamma_{\mathrm{ub}}^{1/2} Z)^4 \mathbf{1}\{|Z| \geq L\}\right\} \leq \mathbb{E}\left\{\mathcal{E}(\gamma_{\mathrm{ub}}^{1/2} Z)^8\right\}^{1/2} \mathbb{P}[|Z| \geq L]^{1/2}.$$

The probability is bounded by $2e^{-L^2/2}$. For the remaining expectation, recall from Lemma 2.4.21(a) that $0 \leq \mathcal{E}(x) \leq |x| + 1$. So,

$$\mathbb{E}\left\{\mathcal{E}(\gamma_{\mathrm{ub}}^{1/2} Z)^8\right\} \leq \mathbb{E}\left\{(1 + \gamma_{\mathrm{ub}}^{1/2} |Z|)^8\right\} \leq 2^7\,\mathbb{E}\left\{1 + \gamma_{\mathrm{ub}}^4 Z^8\right\} = 2^7(1 + 105\gamma_{\mathrm{ub}}^2).$$

Combining these estimates yields

$$\mathbb{E}\left\{\mathcal{E}(\gamma_{\mathrm{ub}}^{1/2} Z)^4 \mathbf{1}\{|Z| \geq L\}\right\} \leq 2^4(1 + 105\gamma_{\mathrm{ub}}^2)^{1/2} e^{-L^2/4} \leq 2^4(1 + 11\gamma_{\mathrm{ub}}) e^{-L^2/4}.$$

All in all,

$$r_4(\gamma_0) \le \delta \sum_{j=-J}^{-1} \mathcal{E}(\gamma_{\text{ub}}^{1/2} x_{j+1})^4 \varphi(x_{j+1}) + \delta \sum_{j=0}^{J-1} \mathcal{E}(\gamma_{\text{ub}}^{1/2} x_{j+1})^4 \varphi(x_j) + 2^4(1+11\gamma_{\text{ub}})e^{-L^2/4} \overset{(*)}{\le} r_{4,\text{ub}},$$

where the step $(*)$ is verified in the accompanying Python file. (See Remark 2.B.13 below for how the function $\mathcal{E}$ is evaluated numerically). For the lower bound, we similarly have

$$\begin{aligned} r_4(\gamma_0) \overset{\text{Lem. 2.B.6}}{\ge} r_4(\gamma_{\text{lb}}) &= \sum_{j=-J}^{J-1} \mathbb{E}\left\{\mathcal{E}(\gamma_{\text{lb}}^{1/2} Z)^4 \mathbf{1}\{Z \in [x_j, x_{j+1}]\}\right\} \\ &\ge \delta \sum_{j=-J}^{-1} \mathcal{E}(\gamma_{\text{lb}}^{1/2} x_j)^4 \varphi(x_j) + \delta \sum_{j=0}^{J-1} \mathcal{E}(\gamma_{\text{lb}}^{1/2} x_j)^4 \varphi(x_{j+1}) \overset{(*)}{\ge} r_{4,\text{lb}}, \end{aligned}$$

where the step $(*)$ is verified in the accompanying Python file. $\square$

Remark 2.B.13. The above computer-assisted proof requires evaluating the function $\mathcal{E}(x) = \varphi(x)/\Psi(x)$, where $\Psi(x) = \mathbb{P}[Z \ge x]$ is itself an integral. We evaluate this as follows. Note that the inputs x on which we numerically evaluate $\mathcal{E}$ are bounded above by $\gamma_{\text{ub}}^{1/2} L \le 10$. Define $L_+ = 12$. We estimate

$$\mathcal{E}(x)^{-1} = \int_x^{L_+} \frac{\varphi(y)}{\varphi(x)} \, \mathsf{d}y + \frac{\mathbb{P}[Z \ge L_+]}{\varphi(x)} \le \int_x^{L_+} e^{-(y^2-x^2)/2} \, \mathsf{d}y + \frac{e^{-(L_+^2-x^2)/2}}{\sqrt{2\pi}}$$

and

$$\mathcal{E}(x)^{-1} \ge \int_x^{L_+} \frac{\varphi(y)}{\varphi(x)} \, \mathsf{d}y = \int_x^{L_+} e^{-(y^2-x^2)/2} \, \mathsf{d}y.$$

The remaining integral can be rigorously bounded by numerical integration, and for $x \le 10$ the term $e^{-(L_+^2-x^2)/2}/\sqrt{2\pi}$ will contribute an error that is multiplicatively small.

Finally, we turn to Claim 2.B.4. By Lemma 2.4.21(b), $\mathcal{E}'$ takes values in $(0,1)$. Thus the function g defined in (2.102) is decreasing in m and increasing in q, and

$$g(m(\widehat{z}), q_0, \gamma_0) \ge g(m_{\text{ub}}, q_{\text{lb}}, \gamma_0). \tag{2.110}$$

However, g is not clearly monotone in γ, so we instead control the derivative of g in γ.

Lemma 2.B.14. *Let $\widetilde{g}(\gamma) = g(m_{\text{ub}}, q_{\text{lb}}, \gamma)$. Then, for all $\gamma \ge 0$, $|\widetilde{g}'(\gamma)| \le 20$.*

Proof. We write $\widetilde{g}(\gamma) = \mathbb{E}[\widehat{g}(\gamma^{1/2} Z)]$, where

$$\widehat{g}(x) = \frac{\mathcal{E}'(x)}{(1-q_{\text{lb}})(1-\mathcal{E}'(x)) + m_{\text{ub}}\mathcal{E}'(x)}. \tag{2.111}$$

A straightforward calculation shows that

$$\widehat{g}''(x) = \frac{(1-q_{\text{lb}})\mathcal{E}^{(3)}(x)}{((1-q_{\text{lb}})(1-\mathcal{E}'(x)) + m_{\text{ub}}\mathcal{E}'(x))^2} - \frac{2(1-q_{\text{lb}})(m_{\text{ub}}+q_{\text{lb}}-1)\mathcal{E}''(x)^2}{(((1-q_{\text{lb}})(1-\mathcal{E}'(x)) + m_{\text{ub}}\mathcal{E}'(x))^2)^3}.$$

Since $\mathcal{E}'(x) \in (0,1)$ by Lemma 2.4.21(b),

$$(1-q_{\text{lb}})(1-\mathcal{E}'(x)) + m_{\text{ub}}\mathcal{E}'(x) \ge \min(1-q_{\text{lb}}, m_{\text{ub}}) = 1-q_{\text{lb}}.$$

Lemma 2.4.21(c)(d) yields $|\mathcal{E}''(x)| \le 1$, $|\mathcal{E}^{(3)}(x)| \le 13$. Thus

$$|\widehat{g}''(x)| \le \frac{13}{1-q_{\text{lb}}} + \frac{2(m_{\text{ub}}+q_{\text{lb}}-1)}{(1-q_{\text{lb}})^2} \le 40,$$

where the final estimate follows from the simple bounds $q_{\text{lb}} \le 3/5$, $m_{\text{ub}} \le 1$. Finally, a gaussian integration by parts calculation yields

$$\widetilde{g}'(\gamma) = \frac{1}{2}\mathbb{E}[\widehat{g}''(\gamma^{1/2} Z)],$$

which implies the result. $\square$

Proof of Claim 2.B.4. In light of (2.110) and Lemma 2.B.14, we will estimate

$$g(m(\widehat{z}), q_0, \gamma_0) \geq g(m_{\mathrm{ub}}, q_{\mathrm{lb}}, \gamma_{\mathrm{lb}}) - 20|\gamma_{\mathrm{ub}} - \gamma_{\mathrm{lb}}|.$$

We will estimate $g(m_{\mathrm{ub}}, q_{\mathrm{lb}}, \gamma_{\mathrm{lb}})$ by discretization, like in the proof of CLaim 2.B.2. Let $L = 8$, $\delta = 10^{-3}$, and $J = L/\delta$. For integer $j \in [-J, J]$, let $x_j = j\delta$.

Note that $\widehat{g}(x)$ defined in (2.111) takes positive values, and is an increasing function of $\mathcal{E}'(x)$. Moreover, by Lemma 2.4.21(c), $\mathcal{E}'(x)$ is an increasing function of x. Thus $\widehat{g}(x)$ is an increasing function of x. Hence,

$$\begin{aligned} g(m_{\mathrm{ub}}, q_{\mathrm{lb}}, \gamma_{\mathrm{lb}}) = \mathbb{E}[\widehat{g}(\gamma_{\mathrm{lb}}^{1/2} Z)] &\geq \sum_{j=-J}^{J-1} \mathbb{E}[\widehat{g}(\gamma_{\mathrm{lb}}^{1/2} Z)\mathbf{1}\{Z \in [x_j, x_{j+1}]\}] \\ &\geq \delta \sum_{j=-J}^{-1} \widehat{g}(\gamma_{\mathrm{lb}}^{1/2} x_j)\varphi(x_j) + \delta \sum_{j=0}^{J-1} \widehat{g}(\gamma_{\mathrm{lb}}^{1/2} x_j)\varphi(x_{j+1}). \end{aligned}$$

Combining the above,

$$g(m(\widehat{z}), q_0, \gamma_0) \geq \delta \sum_{j=-J}^{-1} \widehat{g}(\gamma_{\mathrm{lb}}^{1/2} x_j)\varphi(x_j) + \delta \sum_{j=0}^{J-1} \widehat{g}(\gamma_{\mathrm{lb}}^{1/2} x_j)\varphi(x_{j+1}) - 20|\gamma_{\mathrm{ub}} - \gamma_{\mathrm{lb}}| \stackrel{(*)}{\geq} g_{\mathrm{lb}},$$

where the step $(*)$ is verified in the accompanying Python file. We numerically evaluate $\widehat{g}$ using the identity $\mathcal{E}'(x) = \mathcal{E}(x)(\mathcal{E}(x) - x)$ (Lemma 2.4.21(b)), evaluating $\mathcal{E}$ as in Remark 2.B.13. $\square$

Chapter 3

A constructive proof of the spherical Parisi formula

Abstract – The Parisi formula for the free energy is among the crown jewels in the theory of spin glasses. We present a simpler proof of the lower bound in the case of the spherical mean-field model. Our method follows the TAP approach developed recently in e.g. [Sub24]: we obtain an ultrametric tree of pure states, each with approximately the same free energy as the entire model, which are hierarchically arranged in accordance with the Parisi ansatz. We construct this tree "layer by layer" given the minimizer to Parisi's variational problem. On overlap intervals with full RSB, the tree is built by an optimization algorithm due to Subag. On overlap intervals with finite RSB, the tree is constructed by a new truncated second moment argument; a similar argument also characterizes the free energy of the resulting pure states. Notably we do not use the Aizenman–Sims–Starr scheme, and require interpolation bounds only up to the 1RSB level. Our methods also yield results for large deviations of the ground state, including the entire upper tail rate function for all 1RSB models without external field.

3.1 Introduction

We consider the mixed p-spin Hamiltonian

$$H_N(\boldsymbol{\sigma}) = \sum_{p\geq 1} \frac{\gamma_p}{N^{(p-1)/2}} \sum_{i_1,\ldots,i_p=1}^{N} g_{i_1,\ldots,i_p} \sigma_{i_1} \cdots \sigma_{i_p}. \tag{3.1}$$

Here $g_{i_1,\ldots,i_p}$ are i.i.d. standard Gaussians, and the coefficients γ_p satisfy $\sum_{p\geq 1} 2^p \gamma_p^2 < \infty$.[1] This model is described by the **mixture function** $\xi(t) = \sum_{p=1}^{P} \gamma_p^2 t^p$. For $\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2 \in \mathbb{R}^N$, define the **overlap** $R(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2) = \langle \boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2 \rangle / N$. Then H_N is the Gaussian process with covariance

$$\mathbb{E} H_N(\boldsymbol{\sigma}^1) H_N(\boldsymbol{\sigma}^2) = N\xi(R(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2)).$$

Since the introduction of this model by Sherrington and Kirkpatrick [SK75], a central question has been to understand the free energy, defined as follows. In this paper we consider the case of spherical spins. Let

$$\mathcal{S}_N = \left\{ \boldsymbol{\sigma} \in \mathbb{R}^N : \|\boldsymbol{\sigma}\|_2 = \sqrt{N} \right\}$$

be the sphere of radius $\sqrt{N}$. The partition function and free energy density are defined by

$$Z_N = \int_{\mathcal{S}_N} \exp H_N(\boldsymbol{\sigma}) \, \mathsf{d}\boldsymbol{\sigma}, \tag{3.2}$$

$$F_N = \frac{1}{N} \log Z_N, \tag{3.3}$$

[1] For notational convenience, we have written the model's external field as the degree-1 term of H_N, rather than the traditional $h \sum_{i=1}^{N} \sigma_i$. In the spherical models we consider, these are of course equivalent by rotational invariance.

where in (3.2) the integration is with respect to the uniform probability measure on $\mathcal{S}_N$.

The in-probability limit of the free energy was first predicted by Parisi in [Par79], and proved in the breakthrough works of Talagrand [Tal06b, Tal06a] and Panchenko [Pan13a] following decades of progress in the probability and statistical physics communities. In the equivalent formulation due to Crisanti and Sommers [CS92], the limiting free energy is described as follows. Let $x : [0,1] \to [0,1]$ be a right-continuous non-decreasing function such that $x(\widehat{q}) = 1$ for some $\widehat{q} < 1$ (which may depend on x). Let

$$\widehat{x}(q) = \int_q^1 x(s)\ \mathsf{d}s.$$

Define the Crisanti-Sommers functional

$$\mathsf{P}(x;\xi) = \frac{1}{2}\left\{\xi'(0)\widehat{x}(0) + \int_0^1 \xi''(q)\widehat{x}(q)\ \mathsf{d}q + \int_0^{\widehat{q}} \frac{\mathsf{d}q}{\widehat{x}(q)} + \log(1-\widehat{q})\right\}. \tag{3.4}$$

Note that $\widehat{x}(q) = 1 - q$ for $q > \widehat{q}$, so this functional is independent of $\widehat{q}$. Finally define

$$\mathsf{P}(\xi) = \inf_x \mathsf{P}(x;\xi). \tag{3.5}$$

Theorem 3.1.1 ([Tal06a, Che13]). *The limiting free energy exists and equals*

$$\operatorname*{p-lim}_{N\to\infty} F_N = \mathsf{P}(\xi).$$

3.1.1 Main result

The purpose of this paper is to give a new constructive proof of the (more difficult) lower bound for $\text{p-lim}_{N\to\infty} F_N$ in the Parisi formula. In fact, we will construct an ultrametric tree of pure states, each with the same free energy as the entire model, taking all overlaps in the model's overlap distribution (in fact a slight extension thereof, see (3.6) below) as predicted by Parisi's ultrametric ansatz [Par79, Par83].

We will use the following characterization of the unique minimizer of (3.5). We emphasize that this description (including existence and uniqueness) is a comparatively elementary fact about the variational problem, and as yet says nothing about the free energy F_N. For given x, define

$$F(q) = \xi'(q) - \int_0^q \frac{\mathsf{d}s}{\widehat{x}(s)^2}, \qquad\qquad f(s) = \int_0^s F(q)\ \mathsf{d}q,$$

and

$$S = \{s \le 1 : f(s) = f_{\max}\}, \qquad\qquad f_{\max} = \sup\{f(q) : q \in [0,1)\}. \tag{3.6}$$

Lemma 3.1.2 ([Tal06a, Proposition 2.1]). *There is a unique x attaining the infimum (3.5), which is characterized as follows. Let ν be the probability measure on $[0,1]$ such that $x(q) = \nu([0,q])$. Then $\nu(S) = 1$.*

Remark 3.1.3. The measure ν is the **overlap distribution** of the model ξ. Namely in the *generic* models where $\sum_{p\ \text{even}:\gamma_p>0} 1/p = \sum_{p\ \text{odd}:\gamma_p>0} 1/p = \infty$, one has $\lim_{N\to\infty} \mathbb{E}G^{\otimes 2}(f(R(\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2))) = \int f(x)\mathsf{d}\nu(x)$ for all continuous $f : [-1,1] \to \mathbb{R}$, where G is the Gibbs measure of the model and $\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2$ are independent samples from G. The same holds for arbitrary ξ modulo small "generic perturbations" that do not affect the free energy; see [Pan13b, Chapter 3].

Remark 3.1.4. It is possible for $\mathsf{supp}(\nu) \subseteq S$ to be a strict inclusion, and one may think of overlaps $q \in S \setminus \mathsf{supp}(\nu)$ as "atoms of mass zero" in the overlap distribution. Indeed, [Sub24, Theorem 10] showed that (for generic models) all overlaps in S are *multi-samplable*, meaning that the Gibbs probability of sampling several points with this pairwise overlap is not exponentially small.

The following two definitions describe the geometry of the pure states that our main result will construct.

Definition 3.1.5. For $k, D \in \mathbb{N}$, let $\mathbb{T} = \mathbb{T}(k, D)$ be the tree with vertices $\{\emptyset\} \cup [k] \cup [k]^2 \cup \cdots \cup [k]^D$ rooted at $\emptyset$, where $u \in [k]^d$ is the parent of $v \in [k]^{d+1}$ if u is the length-d prefix of v. For $u, v \in \mathbb{T}$, write $|u|$ for the length of u and $u \wedge v$ for the length of the least common ancestor of u, v. Let $\mathbb{L} = \mathbb{L}(k, D) = [k]^D$ be the leaf set of $\mathbb{T}$.

Definition 3.1.6. Let $k, D \in \mathbb{N}$, $0 \le q_0 < \cdots < q_D \le 1$, $\vec{q} = (q_0, \ldots, q_D)$, and $\delta > 0$. A $(k, D, \vec{q}, \delta)$-ultrametric tree is a collection of points $(\boldsymbol{\sigma}^u)_{u \in \mathbb{T}}$ such that

$$|R(\boldsymbol{\sigma}^u, \boldsymbol{\sigma}^v) - q_{u \wedge v}| \le \delta, \qquad u, v \in \mathbb{T}. \tag{3.7}$$

For $q \in [0, 1)$, define

$$E(q) = \frac{1}{2}\left\{\xi'(0)\widehat{x}(0) + \int_0^q \xi''(s)\widehat{x}(s)\ \mathsf{d}s + \int_0^q \frac{\mathsf{d}s}{\widehat{x}(s)}\right\}.$$

For $k \in \mathbb{N}$, $\delta > 0$, and $\|\boldsymbol{\sigma}\|_2 \le \sqrt{qN}$, let

$$\begin{aligned}\mathsf{Band}_{k,q,\delta}(\boldsymbol{\sigma}) = \Big\{\vec{\boldsymbol{\rho}} = (\boldsymbol{\rho}^1, \ldots, \boldsymbol{\rho}^k) : \|\boldsymbol{\rho}^i\|_2 = \sqrt{qN},\\ |R(\boldsymbol{\rho}^i - \boldsymbol{\sigma}, \boldsymbol{\sigma})| \le \delta, |R(\boldsymbol{\rho}^i - \boldsymbol{\sigma}, \boldsymbol{\rho}^j - \boldsymbol{\sigma})| \le \delta, \quad \forall 1 \le i < j \le k\Big\}.\end{aligned}$$

Theorem 3.1.7. *For any $\delta, \varepsilon > 0$, $D \in \mathbb{N}$ and increasing $q_0, \ldots, q_D \in S$ with $q_D = \sup(S)$, there exists $c > 0$ such that the following holds for any $k \le e^{cN}$. With probability $1 - e^{-cN}$ there is a $(k, D, \vec{q}, \delta)$-ultrametric tree $(\boldsymbol{\sigma}^u)_{u \in \mathbb{T}}$ with the following properties.*

(i) Energy of tree nodes: for each $u \in \mathbb{T}$, $\frac{1}{N} H_N(\boldsymbol{\sigma}^u) \ge E(q_{|u|}) - \varepsilon$.

(ii) Free energy of pure states: for each $u \in \mathbb{L}$,

$$\frac{1}{kN} \log \int_{\mathsf{Band}_{k,1,\delta}(\boldsymbol{\sigma}^u)} \exp\left(\sum_{i=1}^k H_N(\boldsymbol{\rho}^i)\right)\ \mathsf{d}\vec{\boldsymbol{\rho}} \ge \mathsf{P}(\xi) - \varepsilon. \tag{3.8}$$

The free energy lower bound (3.8) holds even in a "k-replicated" sense, where we average over k replicas $\boldsymbol{\rho}^i$ constrained to be nearly orthogonal. This of course lower bounds the free energy of a single replica, as

$$\int_{\mathsf{Band}_{k,1,\delta}(\boldsymbol{\sigma}^u)} \exp\left(\sum_{i=1}^k H_N(\boldsymbol{\rho}^i)\right)\ \mathsf{d}\vec{\boldsymbol{\rho}} \le \left(\int_{\mathsf{Band}_{1,1,\delta}(\boldsymbol{\sigma}^u)} \exp H_N(\boldsymbol{\rho})\ \mathsf{d}\boldsymbol{\rho}\right)^k, \tag{3.9}$$

and this shows there is no free energy cost to taking k approximately orthogonal replicas. In our proof of Theorem 3.1.7, we derive an analogous k-replicated lower bound on the energy increment from any $\boldsymbol{\sigma}^u$, where $u \in \mathbb{T} \setminus \mathbb{L}$, to its children $\boldsymbol{\sigma}^{u1}, \ldots, \boldsymbol{\sigma}^{uk}$, see Theorem 3.4.4(ii); this allows us to construct the ultrametric tree $(\boldsymbol{\sigma}_u)_{u \in \mathbb{T}}$.

As a consequence of Theorem 3.1.7, we obtain the lower bound in the Parisi formula.

Corollary 3.1.8. *We have* $\text{p-liminf}_{N \to \infty} F_N \ge \mathsf{P}(\xi)$.

Proof. Equations (3.8) and (3.9) imply

$$\begin{aligned}F_N = \frac{1}{N} \log \int_{S_N} \exp H_N(\boldsymbol{\rho})\ \mathsf{d}\boldsymbol{\rho} &\ge \frac{1}{N} \log \int_{\mathsf{Band}_{1,1,\delta}(\boldsymbol{\sigma}^u)} \exp H_N(\boldsymbol{\rho})\ \mathsf{d}\boldsymbol{\rho}\\ &\ge \frac{1}{kN} \log \int_{\mathsf{Band}_{k,1,\delta}(\boldsymbol{\sigma}^u)} \exp\left(\sum_{i=1}^k H_N(\boldsymbol{\rho}^i)\right)\ \mathsf{d}\vec{\boldsymbol{\rho}} \ge \mathsf{P}(\xi) - \varepsilon.\end{aligned}$$

Since this holds for any $\varepsilon > 0$, the result follows. $\square$

Taking the temperature to zero, we also obtain the following consequence on near-ground states.

Corollary 3.1.9. *Let ν_∞ be the zero-temperature overlap measure defined in (3.14), and let $q_1 < q_2 < \cdots < q_D = 1$ lie in $\mathsf{supp}(\nu_\infty)$. Then for any $\delta, \varepsilon > 0$, there exists $c > 0$ such that for all $k \le e^{cN}$, with probability $1 - e^{-cN}$ there exists a $(k, D, \vec{q}, \delta)$-ultrametric tree $\mathbb{T} \subseteq \mathcal{S}_N$ such that:*

$$\min_{\boldsymbol{\sigma} \in \mathbb{L}} H_N(\boldsymbol{\sigma})/N \ge \max_{\boldsymbol{\sigma} \in \mathcal{S}_N} H_N(\boldsymbol{\sigma})/N - \varepsilon. \tag{3.10}$$

In fact the same holds with $\mathsf{supp}(\nu_\infty)$ replaced by T from (3.15), which is the zero-temperature analog of S.

Finally in Section 3.5, we study the large deviations of the ground state energy $GS_N = \max_{\boldsymbol{\sigma} \in \mathcal{S}_N} H_N(\boldsymbol{\sigma})/N$. Confirming predictions of [LACTFLD24], we determine the upper tail rate function for all 1RSB ξ with $\gamma_1 = 0$, and identify a sharp phase transition in the speed from $O(N)$ to $\Omega(N^2)$ for general ξ. The former follows by the methods of Section 3.3; the latter uses Corollary 3.1.9, and in particular the fact that k can be taken exponentially large in N.

Remark 3.1.10. Related ultrametric decompositions for spin glasses have appeared in several previous works, including [Jag17, Sub24, CS21]. Our work follows the approach of [Sub24], and in particular uses a uniform concentration idea introduced therein. In the aforementioned works, this idea is used with previously established properties of Gibbs measures and free energies to construct ultrametric decompositions.

Our work proceeds in the opposite direction, using this idea to prove the lower bound in the Parisi formula. Uniform concentration reduces the proof of the general lower bound to four special cases, which we term fundamental types (see Section 3.2). For two of these cases, elementary proofs of the lower bound are known that do not depend on the full Parisi formula. Our main contribution is to provide such a proof for the two remaining cases, and thereby complete an independent proof of the lower bound. As a consequence, we are able to construct the tree in Theorem 3.1.7 one layer at a time "by hand."

A notable aspect of Theorem 3.1.7 is that it gives an ultrametric tree with exponentially large branching factor at each level. At zero temperature, with $\gamma_1 = 0$ so that $0 \in \mathsf{supp}(\nu_\infty)$, the existence of many approximately orthogonal near ground states is closely related to disorder chaos; see [Cha14, DEZ15, AC18, CHL18, Eld20a]. Corollary 3.1.9 is the first to show e^{cN} approximately orthogonal near ground states exist without additional assumptions on ξ.

3.1.2 Previous approaches to the Parisi formula

Mean-field spin glasses were introduced in [SK75, Der81] to model disordered magnetic materials. Soon after, [TAP77, dAT78] observed that the replica-symmetric ansatz made by Sherrington and Kirkpatrick could not be correct at low temperatures. This was resolved by Parisi's ground-breaking replica symmetry breaking solution, yielding a formula for the free energy at any temperature [Par79, Par80, Par83]. Several mysterious, fascinating features were present in this highly non-rigorous ansatz, including the prediction of ultrametricity [MPS+84a, MPS+84b, MPV87] which led to the introduction of Ruelle cascades [Rue87]. *Spherical* spin glasses were also introduced in [CS92], where it was observed that a similar replica ansatz should apply and lead to simpler formulas. Despite this, rigorous results were for some time mainly restricted to high-temperature settings with similar behavior to classical spin systems [ALR87, CN95].

A crucial breakthrough was made in [GT02], which proved the existence of a limiting free energy at all temperatures using the interpolation method. Then in [Gue03], Guerra gave an inspired interpolation upper bound for the free energy, which matched the conjectural Parisi ansatz. Finally, Talagrand used a difficult interpolation scheme (analyzing its intermediate-time behavior using another interpolation) to prove the Parisi formula at all temperatures for both Ising and spherical models, with the slight restriction that $\gamma_p = 0$ for all p odd [Tal06b, Tal06a].

While Talagrand's solution was rather complicated, it was realized in [ASS03] that Guerra's upper bound allows one to transparently deduce an extended variational formula for the free energy over a space of "random overlap structures", relaxing the ultrametricity condition in the Parisi ansatz. Combined with asymptotic ultrametricity of the Gibbs measures shown by [Pan13a], this led to an alternate proof of the Parisi formula with no parity restriction [Pan14, Che13]. More recently, the intrinsic behavior of the associated variational formula was clarified via connection to stochastic Hamilton-Jacobi equations [AC15, JT16], and a limiting zero temperature formula was obtained for the ground state energy [AC17, CS17].

Acknowledgements

We thank Wei-Kuo Chen and Peter Mottishaw for helpful comments.

3.2 Fundamental model types

In this subsection we define four types of models, which we term **topologically trivial**, **strictly RS**, **strictly 1RSB**, and **strictly FRSB**. We state lower bounds on the free energy of strictly RS models and the ground state energies of the other three model types. These models will serve as the basic building blocks for any overlap distribution. The proof of Theorem 3.1.7, carried out in Section 3.4, will decompose a model ξ into several sub-models of these types and apply these results. These lower bounds are then combined back together via a uniform concentration lemma of [Sub24].

Definition 3.2.1. The model ξ is **strictly RS** if $S = \{0\}$.

The remaining three types of models will be defined using a zero-temperature version of the Crisanti-Sommers formula introduced in [CS17], which is obtained as a limit of (3.4). For right-continuous non-decreasing integrable $\alpha : [0,1] \to [0,\infty)$ and $L > \int_0^1 \alpha(s)\, \mathsf{d}s$, let

$$\widehat{\alpha}(q) = L - \int_0^q \alpha(s)\, \mathsf{d}s.$$

Then define

$$\mathcal{Q}(L,\alpha;\xi) = \frac{1}{2}\left\{\xi'(0)L + \int_0^1 \xi''(q)\widehat{\alpha}(q)\, \mathsf{d}q + \int_0^1 \frac{\mathsf{d}q}{\widehat{\alpha}(q)}\right\} \tag{3.11}$$

and

$$\mathcal{Q}(\xi) = \inf_{L,\alpha} \mathcal{Q}(L,\alpha;\xi). \tag{3.12}$$

Theorem 3.2.2 ([CS17, Theorem 1])**.** *The limiting ground state energy of the model ξ is*

$$\operatorname*{p-lim}_{N\to\infty} \frac{1}{N} \max_{\boldsymbol{\sigma}\in\mathcal{S}_N} H_N(\boldsymbol{\sigma}) = \mathcal{Q}(\xi).$$

The minimizer of (3.12) has a similar characterization to Lemma 3.1.2 above. For given L, α, define

$$G(q) = \xi'(q) - \int_0^q \frac{\mathsf{d}s}{\widehat{\alpha}(s)^2}, \qquad\qquad g(s) = \int_s^1 G(q)\, \mathsf{d}q. \tag{3.13}$$

Similarly to before, we let ν_∞ be the finite Borel measure on $[0,1]$ defined by

$$\nu_\infty([0,q]) = \alpha(q) \qquad \forall q \in [0,1] \tag{3.14}$$

and define the set

$$T = \{q \in [0,1] : g(q) = 0\}. \tag{3.15}$$

Note that we always have $1 \in T$.

Lemma 3.2.3 ([CS17, Theorem 2])**.** *There is a unique (L,α) attaining the infimum (3.12), which is characterized by the following properties:*

$$G(1) = 0; \qquad \min_{q\in[0,1]} g(q) = 0; \qquad \nu_\infty(T^c) = 0.$$

Definition 3.2.4. The model ξ is **topologically trivial** if $T = \{1\}$, **strictly 1RSB** if $T = \{0,1\}$, and **strictly FRSB** if $T = [0,1]$.

Remark 3.2.5. Note that ξ can only be strictly RS, strictly 1RSB, or strictly FRSB if $\xi'(0) = \gamma_1^2 = 0$, i.e. there is no external field. Indeed if $\xi'(0) > 0$, then $F(q), G(q) > 0$ for q in a neighborhood of 0, so we cannot have $0 \in S, T$. Conversely, by Lemma 3.2.13(a), ξ can only be topologically trivial if $\xi'(1) \geq \xi''(1)$, which implies $\xi'(0) > 0$ except in the simple case that ξ is quadratic.

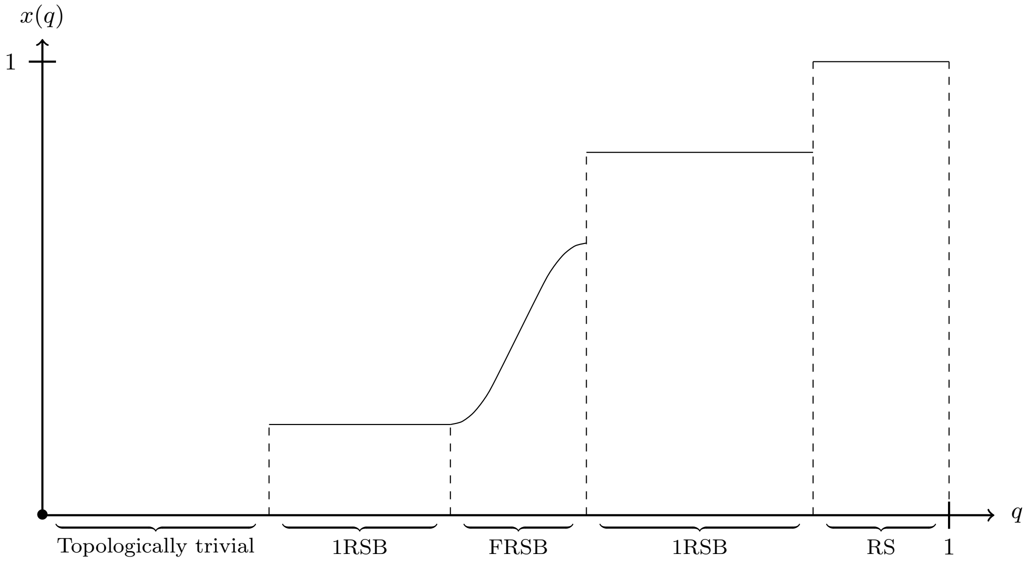


Figure 3.2.1: Decomposition of an overlap distribution into fundamental components. Our proof of Theorem 3.1.7 combines lower bounds on the free or ground state energy of each piece.

3.2.1 Proof outline for Theorem 3.1.7

Decomposition into fundamental types We will construct the ultrametric tree in Theorem 3.1.7 layer by layer, as follows. Let x attain the infimum in (3.5); it is known from [JT18] (see Lemma 3.4.1) that the associated S (3.6) is a finite union of intervals (including possibly atoms). We may assume without loss of generality that the sequence $q_0, \dots, q_D$ contains all endpoints of these intervals, so that $q_0 = \inf S$, $q_D = \sup S$, and each interval $[q_d, q_{d+1}]$ (where we take as convention $q_{-1} = 0$, $q_{D+1} = 1$) either is contained in S or intersects S at exactly its endpoints. Recall that (modulo Remark 3.1.4) S is the support of the overlap distribution $\nu([0, q]) = x(q)$, so these intervals comprise the overlap support and overlap gaps of the model ξ.

Following a construction of [Sub24], we define a sub-model of ξ for each interval $[q_d, q_{d+1}]$ (see (3.47)), which represents the landscape of H_N on an orthogonal band of radius $\sqrt{(q_{d+1} - q_d)N}$ around a point of radius $\sqrt{q_d N}$. Due to the choice of $q_0, \dots, q_D$, the sub-model for $[0, q_0]$ will be topologically trivial, that for $[q_D, 1]$ strictly RS, and the remaining sub-models either strictly 1RSB or strictly FRSB (see Figure 3.2.1). We then prove sharp lower bounds for the free energy of each strictly RS component and the ground state energy of the remaining components. Furthermore, for all but topologically trivial components, this lower bound will hold in a k-replicated sense. This allows us to combine the bounds and construct the tree described by Theorem 3.1.7 in Section 3.4.

The aforementioned lower bounds are stated in Subsection 3.2.2. For topologically trivial and strictly FRSB models, they are already known, using the Kac–Rice formula and an explicit optimization algorithm respectively. Namely [FLD14, BČNS22] showed that topologically trivial models have w.h.p. two critical points, the global maximum and minimum, and characterized their energies (the intuition is that the strong external field overpowers the remainder of the disorder). Meanwhile [Sub21a] showed how to construct an approximate ground state of any strictly FRSB model (which can easily be made k-replicated). It is worth pointing out that locating the ground state of the topologically trivial model in the first stage is analogous to the recentering step from the conditional second moment method approach used in [Bol19, DS18, BY22] for related problems. These works studied replica-symmetric models with external field, which for the purposes of this paper amount to a combination of topologically trivial and strictly RS models.

Given this, our primary remaining tasks are the lower bounds for strictly RS and 1RSB models. We prove these bounds using a new truncated second moment argument, explained below.

We note that the "decomposition" strategy we follow was introduced by [Sub24], and has been subsequently implemented to recover the Parisi formula in several restricted settings. [BSZ20] showed that under the subset of the 1RSB phase called "Condition M", the free energy can be understood at low enough temperature by decomposing the model into 1RSB and RS parts. A similar "full RSB + RS" decomposition was observed in [Sub21a]. In the shattered phase (a subset of the replica-symmetric phase), [BJ24] used a similar decomposition to understand geometric properties of the landscape. These results were highly suggestive and motivational for our work. However they did not apply in fully general models because the RS and 1RSB phases could not always be handled without depending on the Parisi formula. Because we solve these cases independently without this dependency, we can combine this with the above ingredients to arrive at a new proof of the Parisi formula.

Truncated second moment It is natural to study the free energy of strictly RS models using the second moment method. A direct calculation shows that if

$$\xi(q) + \frac{1}{2}\log(1-q^2) \le 0 \qquad \forall q \in [0,1), \tag{3.16}$$

then the second moment method succeeds and $\text{p-lim}_{N\to\infty} F_N = \frac{1}{2}\xi(1)$. However, this does not encompass the entire RS phase. Indeed, by Lemma 3.1.2 (with $x \equiv 1$), the model is strictly RS if (and only if)

$$\xi(q) + q + \log(1-q) \le 0 \qquad \forall q \in [0,1) \tag{3.17}$$

with equality only at $q = 0$. These conditions do not agree, so even in the strictly RS phase it is possible for the dominant contribution to the second moment to come from pairs with nonzero overlap. Similarly, [BSZ20, Condition M] gives a condition under which the second moment method, applied to a suitable critical point count, identifies the ground state energy. However, this condition does not encompass the entire zero-temperature 1RSB phase, given by (3.27) below.

We overcome these difficulties by truncating the moment calculation to *typical* points, see Definitions 3.3.1 and 3.3.8. Roughly, $\boldsymbol{\sigma} \in \mathcal{S}_N$ will be said to be **free energy typical** if for any q, the free energy on the band

$$\mathsf{Band}_q(\boldsymbol{\sigma}) = \{\boldsymbol{\rho} \in \mathcal{S}_N : R(\boldsymbol{\rho}, \boldsymbol{\sigma}) = q\}$$

is at most $\mathsf{P}(\xi)$, and **ground state typical** if the ground state energy on such bands is at most $\mathcal{Q}(\xi)$. We will show that typical points dominate the respective first moments for both of the model types described above, by applying Guerra's interpolation bound to appropriate conditional models. This implies that truncation has only a slight effect on the first moment. On the other hand, truncation immediately ensures that the second moment is dominated by pairs of orthogonal points, causing the second moment method to succeed throughout the RS and 1RSB regimes.

To prove these highly non-obvious typicality properties, we rely on the following upper bounds which follow from the interpolation method of [Gue03]. In fact we require only replica-symmetric bounds at positive temperature, and 1RSB bounds at zero temperature.

Proposition 3.2.6. *For any x as in (3.4) of the form $x(q) = \mathbf{1}\{q \ge q_*\}$,*

$$\operatorname*{p-limsup}_{N\to\infty} F_N \le \mathsf{P}(x;\xi).$$

Proposition 3.2.7. *For any (L,α) as in (3.11) of the form $\alpha(q) = u\mathbf{1}\{q \ge q_*\}$,*

$$\operatorname*{p-limsup}_{N\to\infty} \frac{1}{N} \max_{\boldsymbol{\sigma}\in\mathcal{S}_N} H_N(\boldsymbol{\sigma}) \le \mathcal{Q}(L,\alpha;\xi).$$

Remark 3.2.8. Proposition 3.2.6 follows directly from [Tal06a], while Proposition 3.2.7 follows from the zero temperature limits taken in [CS17] or [JT17]. We note that proving them for general ξ requires Talagrand's positivity principle [Pan07, Tal11], which follows from the Ghirlanda–Guerra identities. If one wishes to avoid these to keep things elementary, one may assume throughout this paper that ξ is convex on $[-1,1]$.

3.2.2 Lower bounds for free and ground state energy

The following propositions lower bound the free or ground state energies in the four fundamental model types. They are special cases of Theorem 3.1.7 where ξ is of these types.

Proposition 3.2.9. *Suppose ξ is strictly RS. For all $\delta, \varepsilon > 0$, there exists $c = c(\xi, \delta, \varepsilon)$ such that if $k \le e^{cN}$, with probability $1 - e^{-cN}$,*

$$\frac{1}{kN} \log \int_{\mathsf{Band}_{k,1,\delta}(\mathbf{0})} \exp\left(\sum_{i=1}^{k} H_N(\boldsymbol{\sigma}^i)\right) \, \mathsf{d}\vec{\boldsymbol{\sigma}} \ge \mathsf{P}(\xi) - \varepsilon.$$

Proposition 3.2.10. *Suppose ξ is strictly 1RSB. For all $\delta, \varepsilon > 0$, there exists $c = c(\xi, \delta, \varepsilon)$ such that if $k \le e^{cN}$, with probability $1 - e^{-cN}$ there exists $\vec{\boldsymbol{\sigma}} \in \mathsf{Band}_{k,1,\delta}(\mathbf{0})$ such that*

$$\frac{1}{N} H_N(\boldsymbol{\sigma}^i) \ge \mathcal{Q}(\xi) - \varepsilon \qquad \forall i \in [k].$$

Proposition 3.2.11. *Suppose ξ is topologically trivial. For all $\varepsilon > 0$, there exists $c = c(\xi, \varepsilon)$ such that with probability $1 - e^{-cN}$, there exists $\boldsymbol{\sigma} \in S_N$ such that $\frac{1}{N} H_N(\boldsymbol{\sigma}) \ge \mathcal{Q}(\xi) - \varepsilon$.*

Proposition 3.2.12. *If ξ is strictly FRSB, the conclusion of Proposition 3.2.10 also holds.*

We will prove Propositions 3.2.9 and 3.2.10 in Section 3.3 by the aforementioned truncated second moment argument. Propositions 3.2.11 and 3.2.12 are known from previous work, and we outline their proofs below.

Lemma 3.2.13. *The following holds.*

(a) If ξ is topologically trivial, then $\xi'(1) \ge \xi''(1)$ and $\mathcal{Q}(\xi) = \sqrt{\xi'(1)}$.

(b) If ξ is strictly FRSB, then $\mathcal{Q}(\xi) = \int_0^1 \xi''(q)^{1/2} \, \mathsf{d}q$.

Proof. If ξ is topologically trivial, then (recalling notation of Lemma 3.2.3), $\alpha \equiv 0$, and so $\widehat{\alpha} \equiv L$. Thus

$$G(q) = \xi'(q) - \frac{q}{L^2}.$$

Since Lemma 3.2.3 gives $G(1) = 0$, we have $L^{-2} = \xi'(1)$. Because $\min g \ge 0$, we have

$$0 \le g''(1) = -G'(1) = -\xi''(1) + L^{-2},$$

so $\xi'(1) \ge \xi''(1)$. Moreover, plugging this (L, α) into (3.11) shows $\mathcal{Q}(\xi) = \sqrt{\xi'(1)}$. This proves part (a). If ξ is strictly FRSB, then $g \equiv 0$, so

$$G'(q) = \xi''(q) - \frac{1}{\widehat{\alpha}(q)^2} = 0.$$

This implies $\widehat{\alpha}(q) = \xi''(q)^{-1/2}$. Plugging this $\widehat{\alpha}$ into (3.11) proves part (b). □

Proof of Proposition 3.2.11. By Lemma 3.2.13(a), we have $\xi'(1) \ge \xi''(1)$. Define

$$GS_N \equiv \frac{1}{N} \max_{\boldsymbol{\sigma} \in \mathcal{S}_N} H_N(\boldsymbol{\sigma}). \tag{3.18}$$

[BČNS22, Theorem 1.1] shows via the Kac–Rice formula that with high probability, $GS_N \ge \sqrt{\xi'(1)} - \varepsilon/2$. By concentration of the ground state energy (see Proposition 3.2.15 below), $GS_N \ge \sqrt{\xi'(1)} - \varepsilon$ with probability $1 - e^{-cN}$. Lemma 3.2.13(a) further implies $\sqrt{\xi'(1)} = \mathcal{Q}(\xi)$, concluding the proof. □

Proof of Proposition 3.2.12. We use a randomized version of Subag's Hessian ascent algorithm [Sub21a] (see also [HS25, Section 3.7]). Starting from $\boldsymbol{x}_0 = 0 \in \mathbb{R}^N$, we choose small $\eta = \eta(\varepsilon, \delta) \in 1/\mathbb{N}$ and for $0 \le j < 1/\eta$ construct $\boldsymbol{x}_{j+1}$ from $\boldsymbol{x}_j$ as follows. For $\boldsymbol{x} \in \mathbb{R}^N$ let $S^{(\eta)}(\boldsymbol{x})$ be the span of the top $N\eta$ eigenvectors of $\nabla^2 H_N(\boldsymbol{x})|_{\boldsymbol{x}^\perp}$, with any measurable-in-$\boldsymbol{x}$ tie-breaking procedure. (Here we view $\nabla^2 H_N(\boldsymbol{x})$ as a quadratic form and restrict it to the subspace $\boldsymbol{x}^\perp = \{\boldsymbol{y} \in \mathbb{R}^N \ : \ \langle \boldsymbol{x}, \boldsymbol{y} \rangle = 0\}$.) Let $S_j = S^{(\eta)}(\boldsymbol{x}_j)$, and choose $\boldsymbol{v}_j \in S_j$

uniformly at random from the corresponding unit sphere, independently of all previous choices. Then let $\boldsymbol{x}_{j+1} = \boldsymbol{x}_j + \boldsymbol{v}_j\sqrt{\eta N}$ if $\langle \boldsymbol{v}_j, \nabla H_N(\boldsymbol{x}_j)\rangle \geq 0$, else $\boldsymbol{x}_{j+1} = \boldsymbol{x}_j - \boldsymbol{v}_j\sqrt{\eta N}$. The output of this algorithm is $\boldsymbol{x}_* = \boldsymbol{x}_{1/\eta} \in \mathcal{S}_N$. (Note that $\langle \boldsymbol{x}_j, \boldsymbol{v}_j\rangle = 0$ by construction, thus $\|\boldsymbol{x}_j\|_2^2 = j\eta N$ almost surely.)

Let $\boldsymbol{x}_*, \boldsymbol{x}'_*$ be independent outputs of this algorithm for the same H_N. We claim that with probability $1 - e^{-cN}$ over H_N and the randomness inside both runs of the algorithm:

$$|\langle \boldsymbol{x}_*, \boldsymbol{x}'_*\rangle| \leq \delta N, \tag{3.19}$$

$$H_N(\boldsymbol{x}_*)/N \geq \int_0^1 \xi''(q)^{1/2}\, \mathsf{d}q - \varepsilon \overset{Lem\, 3.2.13(b)}{=} \mathcal{Q}(\xi) - \varepsilon, . \tag{3.20}$$

This claim implies the desired conclusion by taking $e^{cN/3}$ independent runs of the algorithm. The first part (3.19) holds conditionally on $\boldsymbol{x}'_*$: we have $\mathbb{P}[|\langle \boldsymbol{v}_j, \boldsymbol{x}'_*\rangle| \leq \eta N \mid \boldsymbol{x}'_*, \boldsymbol{x}_j] \geq 1 - e^{-cN}$ since $\boldsymbol{v}_j$ is uniform on a $\Omega(N)$ dimensional sphere.

The second part (3.20) is similar to [Sub21a] (see also [HS25, Section 3.7]) so we give an outline. First one notes that $\nabla^2 H_N(\boldsymbol{x})|_{\boldsymbol{x}^\perp}$ is a GOE matrix scaled by $\sqrt{\left(1 - \frac{1}{N}\right)\xi''(\|\boldsymbol{x}\|_2^2)}$ for $\boldsymbol{x}$ independent of H_N (see e.g. [Sub21a, Eq. (3.10)]). Because of the N^2 speed in the large deviation rate function for the bulk spectrum of GOE, combining a $\eta\sqrt{N}$-net of the ball with Proposition 3.2.14 below implies that $\lambda_{N\eta}(\nabla^2 H_N(\boldsymbol{x})|_{\boldsymbol{x}^\perp}) \geq 2\sqrt{\xi''(\|\boldsymbol{x}\|^2/N)} - \varepsilon^2$ uniformly in $\|\boldsymbol{x}\| \leq \sqrt{N}$ with probability $1 - e^{-cN}$. Using Proposition 3.2.14 again to control the Taylor approximation error, one finds with the same probability:

$$H_N(\boldsymbol{x}_{j+1}) - H_N(\boldsymbol{x}_j) \geq \frac{1}{2} \cdot \eta N \cdot \left(2\sqrt{\xi''(\|\boldsymbol{x}_j\|^2/N)} - \varepsilon^2\right), \quad \forall 0 \leq j < 1/\eta.$$

Telescoping gives (3.20), completing the proof. □

3.2.3 Preliminary concentration estimates

For a tensor $\boldsymbol{A} \in (\mathbb{R}^N)^{\otimes k}$, define the operator norm

$$\|\boldsymbol{A}\|_{\mathsf{op}} = \max_{\|\boldsymbol{\sigma}^1\|_2,\dots,\|\boldsymbol{\sigma}^k\|_2 \leq 1} |\langle \boldsymbol{A}, \boldsymbol{\sigma}^1 \otimes \cdots \otimes \boldsymbol{\sigma}^k\rangle| \,.$$

Proposition 3.2.14. *For any model ξ, there exists a constant $c = c(\xi) > 0$ and sequence of constants $(C_k)_{k\geq 0}$ independent of N such that the following holds. Defining the convex set*

$$K_N = \left\{H_N \in \mathscr{H}_N \;:\; \left\|\nabla^k H_N(\boldsymbol{\sigma})\right\|_{\mathsf{op}} \leq C_k N^{1-\frac{k}{2}} \quad \forall k \geq 0, \|\boldsymbol{\sigma}\|_2 \leq \sqrt{N}\right\} \subseteq \mathscr{H}_N,$$

(i) For all N, we have $\mathbb{P}[H_N \in K_N/2] \geq 1 - e^{-cN}$.

(ii) More generally, let $\psi : \mathscr{H}_N \to \mathbb{R}^M$ be an almost surely finite linear map. Then $\mathbb{P}[H_N \in K_N \;:\; \psi(H_N)] \geq 1 - e^{-cN}$ whenever $\mathbb{E}[H_N \;:\; \psi(H_N)] \in K_N/2$.

Proof. Part (i) follows from e.g. [HS25, Proposition 2.3] (modulo dilation of K_N by a factor of two). For part (ii), note that the conditional law of $H_N - \mathbb{E}[H_N \mid \psi(H_N)]$ does not depend on $\psi(H_N)$. Moreover, letting $H_N^{(\psi)} \in \mathscr{H}_N$ be a Hamiltonian with this law, there exists an independent centered Gaussian $H_N^{(\neg\psi)} \in \mathscr{H}_N$ such that the independent sum $H_N^{(\psi)} + H_N^{(\neg\psi)}$ has the law of H_N. Then whenever $\mathbb{E}[H_N \mid \psi(H_N)] \in K_N/2$, we have

$$\begin{aligned}
\mathbb{P}[H_N \in K_N \mid \psi(H_N)] &\geq \mathbb{P}[H_N^{(\psi)} \in K_N/2] \\
&\overset{(\dagger)}{\geq} 2\,\mathbb{P}[H_N^{(\psi)} + H_N^{(\neg\psi)} \in K_N/2] - 1 \\
&= 2\,\mathbb{P}[H_N \in K_N/2] - 1 \\
&\geq 1 - 2e^{-cN}.
\end{aligned}$$

Here (†) follows from the simple observation that by symmetry in law of $H_N^{(\neg\psi)}$ and independence,

$$\mathbb{P}\big[H_N^{(\psi)} + H_N^{(\neg\psi)} \in K_N/2 \mid H_N^{(\psi)} \notin K_N/2\big] \leq 1/2.$$

□

Proposition 3.2.15. *For any model ξ, there exists a constant $c = c(\xi) > 0$ such that F_N (3.3) and GS_N (3.18) satisfy the concentration inequality*

$$\mathbb{P}(|F_N - \mathbb{E}F_N| \geq t), \mathbb{P}(|GS_N - \mathbb{E}GS_N| \geq t) \leq 2\exp(-ct^2 N).$$

Proof. The concentration inequality for F_N is [Pan13b, Theorem 1.2], and that for GS_N is by the Borell-TIS inequality [Bor76, CIS76]. □

3.2.4 Preliminaries on the Kac–Rice formula

For each $\boldsymbol{\sigma} \in S_N$, let $\{e_1(\boldsymbol{\sigma}), \dots, e_N(\boldsymbol{\sigma})\}$ be an orthonormal basis of $\mathbb{R}^N$ with $e_1(\boldsymbol{\sigma}) = \boldsymbol{\sigma}/\sqrt{N}$. Let $\mathcal{T} = \{2, \dots, N\}$. Let $\nabla_{\mathcal{T}} H_N(\boldsymbol{\sigma}) \in \mathbb{R}^{\mathcal{T}}$ denote the projection of $\nabla H_N(\boldsymbol{\sigma}) \in \mathbb{R}^N$ to the space spanned by $\{e_2(\boldsymbol{\sigma}), \dots, e_N(\boldsymbol{\sigma})\}$, and $\nabla^2_{\mathcal{T}\times\mathcal{T}} H_N(\boldsymbol{\sigma}) \in \mathbb{R}^{\mathcal{T}\times\mathcal{T}}$ analogously. Define the radial and tangential derivatives

$$\partial_{\mathsf{rad}} H_N(\boldsymbol{\sigma}) = \langle e_1(\boldsymbol{\sigma}), \nabla H_N(\boldsymbol{\sigma})\rangle, \qquad \nabla_{\mathsf{sp}} H_N(\boldsymbol{\sigma}) = \nabla_{\mathcal{T}} H_N(\boldsymbol{\sigma}).$$

Further, define the Riemannian Hessian

$$\nabla^2_{\mathsf{sp}} H_N(\boldsymbol{\sigma}) = \nabla^2_{\mathcal{T}\times\mathcal{T}} H_N(\boldsymbol{\sigma}) - \frac{1}{\sqrt{N}} \partial_{\mathsf{rad}} H_N(\boldsymbol{\sigma}) I_{\mathcal{T}\times\mathcal{T}}. \tag{3.21}$$

The following lemma describes the joint law of these derivatives for any fixed $\boldsymbol{\sigma} \in S_N$.

Lemma 3.2.16 ([BČNS22, Lemma 3.2])**.** *For any $\boldsymbol{\sigma} \in S_N$, the random variables $\big(H_N(\boldsymbol{\sigma}), \partial_{\mathsf{rad}} H_N(\boldsymbol{\sigma})\big)$ and $\nabla_{\mathsf{sp}} H_N(\boldsymbol{\sigma})$ and $\nabla^2_{\mathcal{T}\times\mathcal{T}} H_N(\boldsymbol{\sigma})$ are independent, with the following laws.*

- *$(H_N(\boldsymbol{\sigma}), \partial_{\mathsf{rad}} H_N(\boldsymbol{\sigma}))$ is a centered Gaussian with covariance*

$$\mathbb{E}(H_N(\boldsymbol{\sigma}), \partial_{\mathsf{rad}} H_N(\boldsymbol{\sigma}))^{\otimes 2} = \begin{bmatrix} N\xi(1) & \sqrt{N}\xi'(1) \\ \sqrt{N}\xi'(1) & \xi'(1) + \xi''(1) \end{bmatrix}.$$

- *$\nabla_{\mathsf{sp}} H_N(\boldsymbol{\sigma})$ is a centered Gaussian with covariance $\xi'(1) I_{N-1}$.*
- *$\nabla^2_{\mathcal{T}\times\mathcal{T}} H_N(\boldsymbol{\sigma})$ is a scaled GOE matrix, with*

$$\mathbb{E}(\nabla^2_{\mathcal{T}\times\mathcal{T}} H_N(\boldsymbol{\sigma}))^2_{i,j} = \frac{(1+\delta_{i,j})\xi''(1)}{N},$$

symmetry across the diagonal, and independent entries on and above the diagonal.

Say $\boldsymbol{\sigma} \in S_N$ is a **critical point** of H_N if $\nabla_{\mathsf{sp}} H_N(\boldsymbol{\sigma}) = \mathbf{0}$, and let Crt denote the set of such points. Further, for $(\boldsymbol{\sigma}, H_N)$-measurable event $\mathcal{E}$, let

$$\mathsf{Crt}(\mathcal{E}) = \{\boldsymbol{\sigma} \in \mathsf{Crt} : (\boldsymbol{\sigma}, H_N) \in \mathcal{E}.\}. \tag{3.22}$$

The Kac–Rice formula [Ric44, Kac48], applied to $\nabla_{\mathsf{sp}} H_N$, states that

$$\mathbb{E}|\mathsf{Crt}(\mathcal{E})| = \int_{S_N} \mathbb{E}\left[|\det \nabla^2_{\mathsf{sp}} H_N(\boldsymbol{\sigma})| \mathbf{1}\{(\boldsymbol{\sigma}, H_N) \in \mathcal{E}\} \,\middle|\, \nabla_{\mathsf{sp}} H_N(\boldsymbol{\sigma}) = \mathbf{0}\right] \varphi_{\nabla_{\mathsf{sp}} H_N(\boldsymbol{\sigma})}(\mathbf{0})\, \mathsf{d}\boldsymbol{\sigma}, \tag{3.23}$$

where φ_X denotes the probability density of the random variable X. In Section 3.3, we will use specific known consequences of this formula, which hold because conditional on $(H_N(\boldsymbol{\sigma}), \partial_{\mathsf{rad}} H_N(\boldsymbol{\sigma}))$, the matrix $\nabla^2_{\mathsf{sp}} H_N(\boldsymbol{\sigma})$ appearing in the determinant is a shifted and scaled GOE matrix.

3.3 Strictly RS and 1RSB Models via truncated second moment

In this section we prove Propositions 3.2.9 and 3.2.10. Subsection 3.3.1 proves Proposition 3.2.9, and the rest of the section is devoted to the proof of Proposition 3.2.10.

3.3.1 Strictly RS models

Let ξ be strictly RS. Recall from Remark 3.2.5 that this implies $\xi'(0) = 0$. Using the notation of Lemma 3.1.2, $x \equiv 1$ so

$$\mathsf{P}(\xi) = \frac{1}{2}\xi(1).$$

Moreover,

$$f(q) = \xi(q) + q + \log(1-q)$$

has unique maximum $q = 0$ on $[0,1]$. Set $\eta > 0$ small depending on δ such that

$$f(q) \le -6\eta \qquad \forall q \in [\delta, 1]. \tag{3.24}$$

For $\boldsymbol{\sigma} \in \mathcal{S}_N$ and $q \in [-1,1]$, let $\mathsf{Band}_q(\boldsymbol{\sigma}) = \{\boldsymbol{\rho} : R(\boldsymbol{\sigma}, \boldsymbol{\rho}) = q\}$.

Definition 3.3.1. A point $\boldsymbol{\sigma}$ is **free energy typical** if for all $|q| \ge \delta$,

$$\Phi(q; \boldsymbol{\sigma}) \equiv \frac{1}{N}\log \int_{\mathsf{Band}_q(\boldsymbol{\sigma})} \exp H_N(\boldsymbol{\rho})\, \mathsf{d}\boldsymbol{\rho} \le \frac{1}{2}\xi(1) - \eta.$$

We denote by $\mathcal{T}_N \subseteq \mathcal{S}_N$ denote the (random) set of free energy typical points, and

$$\widetilde{Z}_N \equiv \int_{\mathcal{T}_N} \exp H_N(\boldsymbol{\sigma}) \mathsf{d}\boldsymbol{\sigma}.$$

We will prove Proposition 3.2.9 by computing two moments of $\widetilde{Z}_N$.

Proposition 3.3.2. *We have* $\mathbb{E}\widetilde{Z}_N \ge (1 - e^{-cN}) \exp(N\xi(1)/2)$.

Proposition 3.3.3. *We have* $\mathbb{E}\widetilde{Z}_N^2 \le \exp(N(\xi(1) + O(\delta)))$.

For fixed $\boldsymbol{\sigma} \in \mathcal{S}_N$, define the **planted model**

$$H_N(\boldsymbol{\rho}) = N\xi(R(\boldsymbol{\rho}, \boldsymbol{\sigma})) + \widehat{H}_N(\boldsymbol{\rho}),$$

where $\widehat{H}_N(\boldsymbol{\rho})$ is a spin glass with mixture ξ. Let $\mathbb{P}^{\boldsymbol{\sigma}}_{\mathsf{Pl}}$ denote the law of this H_N.

The following crucial lemma uses the interpolation bound Proposition 3.2.6 on a band. A similar estimate was observed by Alaoui, Montanari, and the second author in [AMS25, Proposition 3.9] to study shattering.

Lemma 3.3.4. *For fixed* $\boldsymbol{\sigma} \in \mathcal{S}_N$*,* $|q| \ge \delta$*, the following holds. With probability* $1 - e^{-cN}$ *over* $\mathbb{P}^{\boldsymbol{\sigma}}_{\mathsf{Pl}}$*,*

$$\Phi(q; \boldsymbol{\sigma}) \le \frac{1}{2}\xi(1) - 2\eta.$$

Proof. For $\boldsymbol{\rho} \in \mathsf{Band}_q(\boldsymbol{\sigma})$, under $\mathbb{P}^{\boldsymbol{\sigma}}_{\mathsf{Pl}}$ we have

$$H_N(\boldsymbol{\rho}) = \xi(q)N + \widehat{H}_N(\boldsymbol{\rho}).$$

Moreover, on $\mathsf{Band}_q(\boldsymbol{\sigma})$, writing $\boldsymbol{\rho} = q\boldsymbol{\sigma} + \sqrt{1-q^2}\boldsymbol{\tau}$ for $\boldsymbol{\tau} \perp \boldsymbol{\sigma}$, the process $\overline{H}_N(\boldsymbol{\tau}) = \widehat{H}_N(\boldsymbol{\rho})$ has mixture

$$\widetilde{\xi}_q(t) = \xi(q^2 + (1-q^2)t). \tag{3.25}$$

Define the order parameter $\widetilde{x}(t) = \mathbf{1}\{t \ge r\}$ where $r = \frac{q}{1+q}$. By Proposition 3.2.6,

$$\Phi(q; \boldsymbol{\sigma}) \le \xi(q) + \mathsf{P}(\widetilde{x}; \widetilde{\xi}_q, 0) + \frac{1}{2}\log(1-q^2) + o_P(1),$$

where $\frac{1}{2}\log(1-q^2)$ accounts for the volume of $\mathsf{Band}_q(\boldsymbol{\sigma})$ and $o_P(1)$ is a term tending to 0 in probability (under $\mathbb{P}^{\boldsymbol{\sigma}}_{\mathsf{Pl}}$). Note that $\mathsf{P}(\widetilde{x}; \widetilde{\xi}_q, 0)$ depends on $\widetilde{\xi}_q$ through $\widetilde{\xi}_q'$ and $\widetilde{\xi}_q''$ only, and these depend on q only through $|q|$. Moreover $\xi(q) \le \xi(|q|)$. So we may assume without loss of generality that $q > 0$.

We upper bound $\mathsf{P}(\widetilde{x};\widetilde{\xi}_q,0)$ via:

$$\begin{aligned}2\mathsf{P}(\widetilde{x};\widetilde{\xi}_q,0) &\le \widetilde{\xi}_q'(r)(1-r) + \int_r^1 \widetilde{\xi}_q''(t)(1-t)\,\mathsf{d}t + \frac{r}{1-r} + \int_r^1 \frac{\mathsf{d}t}{1-t}\\ &= \widetilde{\xi}_q(1) - \widetilde{\xi}_q(r) + \frac{r}{1-r} + \log(1-r)\\ &= \xi(1) - \xi(q) + q - \log(1+q).\end{aligned}$$

Thus

$$\Phi(q;\boldsymbol{\sigma}) \le \frac{1}{2}\left(\xi(1)+\xi(q)+q+\log(1-q)\right) + o_P(1) \overset{(3.24)}{\le} \frac{1}{2}\xi(1) - 3\eta + o_P(1).$$

The conclusion follows from Proposition 3.2.15, applied to the free energy $\Phi(q;\boldsymbol{\sigma})$. □

Lemma 3.3.5. *For any fixed* $\boldsymbol{\sigma}\in\mathcal{S}_N$, $\mathbb{P}^{\boldsymbol{\sigma}}_{\mathsf{Pl}}[\boldsymbol{\sigma}\in\mathcal{T}_N] \ge 1-e^{-cN}$.

Proof. Suppose the event in Lemma 3.3.4 holds for all $q\in\{\pm\delta,\pm(\delta+1/N),\dots,\pm(\delta+M/N)\}$ for the largest M such that $\delta+M/N\le 1$, and furthermore that $H_N\in K_N$ holds. By Proposition 3.2.14(ii) with $\psi(H_N)=H_N(\boldsymbol{\sigma})$, this occurs with probability $1-e^{-cN}$ after adjusting c. Then, for all $q=\pm(\delta+m/N)$,

$$\Phi(q;\boldsymbol{\sigma}) \le \frac{1}{2}\xi(1) - 2\eta.$$

For all $q\in[\delta+m/N,\delta+(m+1)/N]$, on event K_N,

$$\Phi(q;\boldsymbol{\sigma}) \le \Phi(\delta+m/N;\boldsymbol{\sigma}) + O(N^{-1}) \le \frac{1}{2}\xi(1)-\eta,$$

and similarly for $q\in[-\delta-(m+1)/N,-\delta-m/N]$. This implies $\boldsymbol{\sigma}\in\mathcal{T}_N$. □

Proof of Proposition 3.3.2. For any $\boldsymbol{\sigma}\in\mathcal{S}_N$,

$$\frac{\mathbb{E}\widetilde{Z}_N}{\mathbb{E}Z_N} = \frac{\mathbb{E}[\exp(H_N(\boldsymbol{\sigma}))\mathbf{1}\{\boldsymbol{\sigma}\in\mathcal{T}_N\}]}{\mathbb{E}[\exp(H_N(\boldsymbol{\sigma}))]}.$$

This is the probability of $\boldsymbol{\sigma}\in\mathcal{T}_N$ under the reweighted measure $\widetilde{\mathbb{P}}$ with Radon–Nikodym derivative $\widetilde{\mathbb{P}}/\mathbb{P}\propto \exp(H_N(\boldsymbol{\sigma}))$, and by properties of Gaussian conditioning we have precisely $\widetilde{\mathbb{P}}=\mathbb{P}^{\boldsymbol{\sigma}}_{\mathsf{Pl}}$. Combining with Lemma 3.3.5 yields

$$\frac{\mathbb{E}\widetilde{Z}_N}{\mathbb{E}Z_N} = \mathbb{P}^{\boldsymbol{\sigma}}_{\mathsf{Pl}}(\boldsymbol{\sigma}\in\mathcal{T}_N) = 1-e^{-cN}.$$

The result follows because $\mathbb{E}Z_N=\exp(N\xi(1)/2)$. □

Proof of Proposition 3.3.3. Writing $\widetilde{Z}_N^2$ as a double integral and recalling $\mathcal{T}_N$ from Definition 3.3.1, we have almost surely

$$\begin{aligned}\widetilde{Z}_N^2 &= \iint_{\mathcal{T}_N\times\mathcal{T}_N} \exp(H_N(\boldsymbol{\sigma}^1)+H_N(\boldsymbol{\sigma}^2))\mathsf{d}\boldsymbol{\sigma}^1\mathsf{d}\boldsymbol{\sigma}^2\\ &\le \iint_{\mathcal{S}_N\times\mathcal{S}_N} \exp(H_N(\boldsymbol{\sigma}^1)+H_N(\boldsymbol{\sigma}^2))\mathbf{1}\{R(\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2)\le\delta\}\mathsf{d}\boldsymbol{\sigma}^1\mathsf{d}\boldsymbol{\sigma}^2 + e^{N(\xi(1)/2-\eta)}\int_{\mathcal{S}_N}\exp(H_N(\boldsymbol{\sigma}))\,\mathsf{d}\boldsymbol{\sigma}.\end{aligned}$$

Taking expectations,

$$\mathbb{E}\widetilde{Z}_N^2 \le \int_{-\delta}^{\delta} e^{N(\xi(1)+\xi(q))}(1-q^2)^{-(N-1)/2}\,\mathsf{d}q + e^{N(\xi(1)-\eta)}.$$

The exponential growth rate of the integral is

$$\xi(1) + \max_{q\in[-\delta,\delta]}\left\{\xi(q)+\frac{1}{2}\log(1-q^2)\right\} = \xi(1)+O(\delta)$$

by Lipschitz continuity of the quantity inside the maximum around 0. □

Proof of Proposition 3.2.9. The statement is clearly monotone in δ, so it suffices to prove it for δ suitably small in ε. By the last two propositions and Paley-Zygmund,

$$\mathbb{P}\left(\widetilde{Z}_N \geq \frac{1}{2}\exp(N\xi(1)/2)\right) \geq e^{-O(\delta)N}. \tag{3.26}$$

Suppose this event holds. Let G denote the Gibbs measure of $\widetilde{Z}_N$ (i.e. the Gibbs measure of Z_N conditioned on $\boldsymbol{\sigma} \in \mathcal{T}_N$). Then,

$$\begin{aligned}\int_{\mathsf{Band}_{k,1,\delta}(\mathbf{0})} \exp\left(\sum_{i=1}^k H_N(\boldsymbol{\sigma}^i)\right) \, \mathsf{d}\vec{\boldsymbol{\sigma}} &\geq \int_{\mathsf{Band}_{k,1,\delta}(\mathbf{0})\cap\mathcal{T}_N^k} \exp\left(\sum_{i=1}^k H_N(\boldsymbol{\sigma}^i)\right) \, \mathsf{d}\vec{\boldsymbol{\sigma}} \\ &= \widetilde{Z}_N^k G^{\otimes k}\left(|R(\boldsymbol{\sigma}^i,\boldsymbol{\sigma}^j)| \leq \delta \text{ for all } 1 \leq i < j \leq k\right) \\ &\geq \widetilde{Z}_N^k \left(1 - \binom{k}{2} G^{\otimes 2}(|R(\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2)| \geq \delta)\right).\end{aligned}$$

For any $\boldsymbol{\sigma}^1 \in \mathcal{T}_N$, G-almost surely

$$\begin{aligned}G(|R(\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2)| \geq \delta|\boldsymbol{\sigma}^1) &= \frac{1}{\widetilde{Z}_N}\int_{|q|\geq\delta}\int_{\mathsf{Band}_q(\boldsymbol{\sigma}^1)} \exp(H_N(\boldsymbol{\sigma}^2)) \, \mathsf{d}\boldsymbol{\sigma}^2 \, \mathsf{d}q \\ &\leq \frac{\exp(N(\xi(1)/2-\eta))}{\frac{1}{2}\exp(N\xi(1)/2)} = 2\exp(-\eta N).\end{aligned}$$

For $k \leq e^{\eta N/3}$,

$$1 - \binom{k}{2} G^{\otimes 2}(|R(\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2)| \geq \delta) \geq \frac{1}{2}.$$

Combining the above and letting

$$Z^{(k)} = \frac{1}{kN}\log\int_{\mathsf{Band}_{k,1,\delta}(\mathbf{0})} \exp\left(\sum_{i=1}^k H_N(\boldsymbol{\sigma}^i)\right) \, \mathsf{d}\vec{\boldsymbol{\sigma}},$$

we conclude that

$$\mathbb{P}\left(Z^{(k)} \geq \frac{\xi(1)}{2} - o(1)\right) \geq e^{-O(\delta)N}.$$

Similarly to [Pan13b, Theorem 1.2] we can show the concentration inequality

$$\mathbb{P}(|Z^{(k)} - \mathbb{E}Z^{(k)}| \geq t) \leq \exp(-ct^2N).$$

(In fact a much stronger inequality is true, see Lemma 3.4.6 below.) The result follows from the last two inequalities for δ suitably small in ε. $\square$

3.3.2 Strictly 1RSB models

We turn to the proof of Proposition 3.2.10. Let ξ be strictly 1RSB, and so $\xi'(0) = 0$ by Remark 3.2.5. For most of the proof we will assume that ξ is **not** pure; we then appeal to continuity at the end by slightly perturbing any pure ξ (which will preserve the strict 1RSB property). This simplifies the presentation below as certain formulas degenerate in the pure case, see Remark 3.3.14. We note that some intermediate computations and lemmas resemble those from previous work including [ABČ13, AB13, BSZ20], which made different assumptions on ξ.

Recalling the notation of Lemma 3.2.3, there exists $u > 0$ such that $\alpha \equiv u$, and the order parameter (L, α) is described by the pair (L, u). It can easily be verified that the function $\upsilon : [0, +\infty) \to \mathbb{R}$ given by

$$\upsilon(z) = \frac{(1+z)\log(1+z)}{z^2} - \frac{1}{z}$$

is strictly decreasing with $\lim_{z\to 0^+} \upsilon(z) = \frac{1}{2}$ and $\lim_{z\to\infty} \upsilon(z) = 0$.

Lemma 3.3.6. *Assume ξ is strictly 1RSB. Let z be the unique solution to $\upsilon(z) = \frac{\xi(1)}{\xi'(1)}$ and $y = \sqrt{(1+z)\xi'(1)}$. Then*

$$L = \frac{1+z}{y}, \qquad u = \frac{z}{y}, \qquad \mathcal{Q}(\xi) = \frac{\xi'(1) + z\xi(1)}{y},$$

and for all $q \in [0,1]$,

$$\xi(1) - \xi(q) \geq \xi'(1)\left(\frac{1+z}{z^2}\log\left(1+(1-q)z\right) - \frac{1-q}{z}\right), \tag{3.27}$$

with equality at exactly $q = 0, 1$.

Proof. As $\widehat{\alpha}(q) = L - uq$, we calculate that

$$G(q) = \xi'(q) - \frac{q}{L(L-uq)}, \qquad g(q) = \xi(1) - \xi(q) - \frac{1}{u}\left(\frac{1}{u}\log\frac{L-uq}{L-u} - \frac{1-q}{L}\right).$$

Since $G(1) = g(0) = 0$, we get the system of equations

$$\xi'(1) = \frac{1}{L(L-u)}, \qquad \xi(1) = \frac{1}{u^2}\log\frac{L}{L-u} - \frac{1}{Lu}.$$

Let $z' = \frac{u}{L-u}$, so

$$\frac{\xi(1)}{\xi'(1)} = \frac{L(L-u)}{u^2}\log\frac{L}{L-u} - \frac{L-u}{u} = \upsilon(z').$$

Thus $z' = z$ by monotonicity of υ. Then

$$L = \sqrt{\frac{1+z'}{\xi'(1)}} = \frac{1+z}{y}, \qquad u = L \cdot \frac{z'}{1+z'} = \frac{z}{y}.$$

These formulas and the condition $g(q) \geq 0$ for all q (with equality at precisely $q = 0, 1$) imply (3.27). Finally,

$$\begin{aligned} 2\mathcal{Q}(\xi) = 2\mathcal{Q}(L, \alpha; \xi) &= \int_0^1 \xi''(q)(L-uq)\mathsf{d}q + \int_0^1 \frac{\mathsf{d}q}{L-uq} \\ &= (L-u)\xi'(1) + u\xi(1) + \frac{1}{u}\log(1+z). \end{aligned}$$

Note that $(L-u)\xi'(1) = \frac{\xi'(1)}{y}$, $u\xi(1) = \frac{z\xi'(1)}{y}$, and

$$\frac{1}{u}\log(1+z) = \frac{z\xi'(1)}{y}\left(\frac{(1+z)}{z^2}\log(1+z)\right) = \frac{z\xi'(1)}{y}\left(\frac{1}{z} + \frac{\xi(1)}{\xi'(1)}\right) = \frac{\xi'(1)+z\xi(1)}{y}.$$

This gives the formula for $\mathcal{Q}(\xi)$. $\square$

We also record a simple inequality in the parameters we will use later.

Lemma 3.3.7. *We have $y^2 \geq \xi''(1)$.*

Proof. Because $\min g(q) = 0$ is attained at $q = 1$,

$$0 \leq g''(1) = -G'(1) = -\xi''(1) + (L-u)^{-2} = -\xi''(1) + y^2. \qquad \square$$

Let $\eta_1, \eta_2, \eta_3 > 0$ be small constants to be determined later, where each η_i will be set small in terms of δ, ξ, and $\{\eta_j : j < i\}$ (i.e. informally, $0 < \eta_3 \ll \eta_2 \ll \eta_1 \ll \delta \ll 1$). By (3.27), we may set η_3 such that

$$\xi(1) - \xi(q) \geq \xi'(1)\left(\frac{1+z}{z^2}\log\left(1+(1-q)z\right) - \frac{1-q}{z}\right) + \frac{6y}{z}\eta_3 \tag{3.28}$$

for all $q \in [\delta, 1-\eta_2]$. Set

$$\begin{aligned} E_0 &= \mathcal{Q}(\xi) = \frac{\xi'(1)+z\xi(1)}{y}, \qquad\qquad R_0 = y + \frac{\xi''(1)}{y}, \\ B &= [E_0-\eta_1-\eta_3, E_0-\eta_1+\eta_3] \times [R_0+\eta_1^{3/4}-\eta_3, R_0+\eta_1^{3/4}+\eta_3]. \end{aligned} \tag{3.29}$$

By slight abuse of notation, for $A \subseteq \mathbb{R}^2$, let

$$\mathsf{Crt}(A) = \left\{ \boldsymbol{\sigma} \in \mathsf{Crt} : \left(\frac{1}{N} H_N(\boldsymbol{\sigma}), \frac{1}{\sqrt{N}} \partial_{\mathsf{rad}} H_N(\boldsymbol{\sigma}) \right) \in A \right\}. \tag{3.30}$$

Definition 3.3.8. A point $\boldsymbol{\sigma} \in \mathsf{Crt}(B)$ is **ground state typical** if the following conditions hold.

(i) For all $q \in [\delta, 1-\eta_2]$,

$$\Psi(q; \boldsymbol{\sigma}) \equiv \frac{1}{N} \sup_{\boldsymbol{\rho} \in \mathsf{Band}_q(\boldsymbol{\sigma})} H_N(\boldsymbol{\rho}) \le E_0 - \eta_1 - \eta_3.$$

(ii) H_N does not have any critical points $\boldsymbol{\rho}$ with $R(\boldsymbol{\sigma}, \boldsymbol{\rho}) \ge 1 - \eta_2$.

Denote the set of such points by $\widetilde{\mathsf{Crt}}(B)$.

We will prove Proposition 3.2.10 via the next two propositions, whose proofs comprise the rest of the section.

Proposition 3.3.9. *We have* $\mathbb{E}|\widetilde{\mathsf{Crt}}(B)| \ge e^{\Omega(\eta_1)N}$.

Proposition 3.3.10. *We have* $\mathbb{E}|\widetilde{\mathsf{Crt}}(B)|^2 \le e^{O(\delta)N}$.

Remark 3.3.11. The choice (3.29) of B looks strange at first, because when ξ is a pure model $H_N(\boldsymbol{\sigma})$ and $\partial_{\mathsf{rad}} H_N(\boldsymbol{\sigma})$ are almost surely proportional, so there are a.s. no critical points with energy and radial derivative described by B. This is not a problem for the proof because the parameters η_1, η_2, η_3 can be taken small in ξ, and then the statement of Proposition 3.2.10 is continuous in ξ; see the end of the proof of Proposition 3.2.10.

3.3.3 Ground state typicality is with high probability

The main result of this subsection is the following proposition. In it, we fix $\boldsymbol{\sigma} \in \mathcal{S}_N$ and condition on the event $\{\boldsymbol{\sigma} \in \mathsf{Crt}(B)\}$. Here when conditioning, we refer to the standard regular conditional probability given $(H_N(\boldsymbol{\sigma}), \partial_{\mathsf{rad}} H_N(\boldsymbol{\sigma}))$, which is a linear function of H_N.

Proposition 3.3.12. *If* $(E, R) \in B$, *then*

$$\mathbb{P}\big[\boldsymbol{\sigma} \in \widetilde{\mathsf{Crt}}(B) \mid \big(H_N(\boldsymbol{\sigma}), \partial_{\mathsf{rad}} H_N(\boldsymbol{\sigma}), \nabla_{\mathsf{sp}} H_N(\boldsymbol{\sigma})\big) = (EN, R\sqrt{N}, \mathbf{0})\big] \ge 1 - e^{-cN}.$$

We will prove Proposition 3.3.12 by studying the ground state energy on bands defined by their overlap with $\boldsymbol{\sigma}$, analogously to the replica-symmetric case.

Lemma 3.3.13. *Conditional on* $(H_N(\boldsymbol{\sigma}), \partial_{\mathsf{rad}} H_N(\boldsymbol{\sigma}), \nabla_{\mathsf{sp}} H_N(\boldsymbol{\sigma})) = (EN, R\sqrt{N}, \mathbf{0})$*, the restriction of* H_N *to* $\mathsf{Band}_q(\boldsymbol{\sigma})$ *has law*

$$H_N(\boldsymbol{\rho}) \stackrel{d}{=} \frac{N(q\xi'(q)+z\xi(q))}{y} + N \left\langle v^q, \begin{bmatrix} E - E_0 \\ R - R_0 \end{bmatrix} \right\rangle + \widehat{H}_N(\boldsymbol{\rho}), \tag{3.31}$$

where

$$v^q = \begin{bmatrix} v_E^q \\ v_R^q \end{bmatrix} = \begin{bmatrix} \xi(1) & \xi'(1) \\ \xi'(1) & \xi'(1)+\xi''(1) \end{bmatrix}^{-1} \begin{bmatrix} \xi(q) \\ q\xi'(q) \end{bmatrix} \tag{3.32}$$

and $\widehat{H}_N$ *is a* $(N-1)$*-dimensional spin glass with the following covariance. Write* $\boldsymbol{\rho} = q\boldsymbol{\sigma} + \sqrt{1-q^2}\boldsymbol{\tau}$ *and let* $\overline{H}_N(\boldsymbol{\tau}) = \widehat{H}_N(\boldsymbol{\rho})$*. Then* $\overline{H}_N$ *has mixture*

$$\widetilde{\xi}_q(t) = \xi(q^2+(1-q^2)t) - C - \frac{(1-q^2)\xi'(q)^2}{\xi'(1)} t. \tag{3.33}$$

for some constant C *(which will be irrelevant for our purposes).*

Remark 3.3.14. Note that the matrix in (3.32) has determinant

$$\xi(1)(\xi'(1)+\xi''(1))-\xi'(1)^2 = \left(\sum_{p\ge 2}\gamma_p^2\right)\left(\sum_{p\ge 2}p^2\gamma_p^2\right)-\left(\sum_{p\ge 2}p\gamma_p^2\right)^2.$$

This is nonnegative by Cauchy-Schwarz and strictly positive when ξ is not pure, so the matrix inverse is well-defined. This is one reason we assume ξ is not pure.

Proof. By Lemma 3.2.16, $H_N(\boldsymbol{\sigma})$, $\partial_{\mathsf{rad}}H_N(\boldsymbol{\sigma})$, and $\nabla_{\mathsf{sp}}H_N(\boldsymbol{\sigma})$ are jointly Gaussian with covariance matrix

$$\begin{bmatrix} N\xi(1) & \sqrt{N}\xi'(1) & \mathbf{0}^\top \\ \sqrt{N}\xi'(1) & \xi'(1)+\xi''(1) & \mathbf{0}^\top \\ \mathbf{0} & \mathbf{0} & \xi'(1)I_{N-1} \end{bmatrix}.$$

For $\boldsymbol{\rho}\in\mathsf{Band}_q(\boldsymbol{\sigma})$, we further have

$$\begin{aligned}\mathbb{E}H_N(\boldsymbol{\rho})H_N(\boldsymbol{\sigma}) &= N\xi(q),\\ \mathbb{E}H_N(\boldsymbol{\rho})\partial_{\mathsf{rad}}H_N(\boldsymbol{\sigma}) &= \sqrt{N}q\xi'(q),\\ \mathbb{E}H_N(\boldsymbol{\rho})\nabla_{\mathsf{sp}}H_N(\boldsymbol{\sigma}) &= \xi'(q)P_{\boldsymbol{\sigma}}^\perp\boldsymbol{\rho}.\end{aligned}$$

Thus

$$\begin{aligned}\mathbb{E}[H_N(\boldsymbol{\rho})|H_N(\boldsymbol{\sigma}),\partial_{\mathsf{rad}}H_N(\boldsymbol{\sigma}),\nabla_{\mathsf{sp}}H_N(\boldsymbol{\sigma})] &= \left\langle \begin{bmatrix}\xi(1) & \xi'(1)\\ \xi'(1) & \xi'(1)+\xi''(1)\end{bmatrix}^{-1}\begin{bmatrix}\xi(q)\\ q\xi'(q)\end{bmatrix}, \begin{bmatrix}H_N(\boldsymbol{\sigma})\\ \sqrt{N}\partial_{\mathsf{rad}}H_N(\boldsymbol{\sigma})\end{bmatrix}\right\rangle \\ &\quad + \frac{\xi'(q)}{\xi'(1)}\langle P_{\boldsymbol{\sigma}}^\perp\boldsymbol{\rho},\nabla_{\mathsf{sp}}H_N(\boldsymbol{\sigma})\rangle. \qquad (3.34)\end{aligned}$$

Then $\widehat{H}_N(\boldsymbol{\rho}) = H_N(\boldsymbol{\rho})-\mathbb{E}[H_N(\boldsymbol{\rho})|H_N(\boldsymbol{\rho}),\partial_{\mathsf{rad}}H_N(\boldsymbol{\rho}),\nabla_{\mathsf{sp}}H_N(\boldsymbol{\rho})]$ has covariance

$$\begin{aligned}\frac{1}{N}\mathbb{E}\widehat{H}_N(\boldsymbol{\rho}^1)\widehat{H}_N(\boldsymbol{\rho}^2) &= \xi(R(\boldsymbol{\rho}^1,\boldsymbol{\rho}^2)) - \left\langle\begin{bmatrix}\xi(1) & \xi'(1)\\ \xi'(1) & \xi'(1)+\xi''(1)\end{bmatrix}^{-1}\begin{bmatrix}\xi(q)\\ q\xi'(q)\end{bmatrix},\begin{bmatrix}\xi(q)\\ q\xi'(q)\end{bmatrix}\right\rangle\\ &\quad - \frac{\xi'(q)^2}{\xi'(1)}R(P_{\boldsymbol{\sigma}}^\perp\boldsymbol{\rho}^1,P_{\boldsymbol{\sigma}}^\perp\boldsymbol{\rho}^2).\end{aligned}$$

This proves (3.33). The conclusion (3.31) follows from (3.34), by noting that

$$\begin{aligned}&\begin{bmatrix}\xi(1) & \xi'(1)\\ \xi'(1) & \xi'(1)+\xi''(1)\end{bmatrix}^{-1}\begin{bmatrix}E_0\\ R_0\end{bmatrix}\\ &= \frac{1}{y(\xi(1)(\xi'(1)+\xi''(1))-\xi'(1)^2)}\begin{bmatrix}\xi'(1)+\xi''(1) & -\xi'(1)\\ -\xi'(1) & \xi(1)\end{bmatrix}\begin{bmatrix}z\xi(1)+\xi'(1)\\ (1+z)\xi'(1)+\xi''(1)\end{bmatrix} = \frac{1}{y}\begin{bmatrix}z\\ 1\end{bmatrix}, \qquad (3.35)\end{aligned}$$

and thus

$$\left\langle\begin{bmatrix}\xi(1) & \xi'(1)\\ \xi'(1) & \xi'(1)+\xi''(1)\end{bmatrix}^{-1}\begin{bmatrix}\xi(q)\\ q\xi'(q)\end{bmatrix},\begin{bmatrix}E_0\\ R_0\end{bmatrix}\right\rangle = \frac{q\xi'(q)+z\xi(q)}{y}. \qquad \square$$

The next estimate will prepare us to apply Proposition 3.2.7. We will use the order parameters:

$$\widetilde{L} = \frac{1+(1-q)z}{(1-q^2)y}, \quad \widetilde{\alpha}(t) = u\mathbf{1}\{t\ge r\}, \quad r = \frac{q}{1+q}. \qquad (3.36)$$

Proposition 3.3.15. *For fixed $\boldsymbol{\sigma}\in\mathcal{S}_N$, $E,R\in\mathbb{R}$, $q\in[\delta,1-\eta_2]$,*

$$\frac{q\xi'(q)+z\xi(q)}{y} + \mathcal{Q}(\widetilde{L},\widetilde{\alpha};\widetilde{\xi}_q,0) \le E_0 - 3\eta_3.$$

Proof. By direct computation,

$$2\mathcal{Q}(\widetilde{L},\widetilde{\alpha};\widetilde{\xi}_q,0) = \widetilde{\xi}_q'(r)\widetilde{L} + \int_r^1 \widetilde{\xi}_q''(t)\left(\widetilde{L}-(t-r)u\right)\,\mathsf{d}t + \frac{r}{\widetilde{L}} + \int_r^1 \frac{\mathsf{d}t}{\widetilde{L}-(t-r)u}. \tag{3.37}$$

The first two terms on the right-hand side simplify as

$$\begin{aligned}
&\widetilde{\xi}_q'(r)\widetilde{L} + \int_r^1 \widetilde{\xi}_q''(t)\left(\widetilde{L}-(t-r)u\right)\,\mathsf{d}t \\
&= \widetilde{\xi}_q'(1)\widetilde{L} + (\widetilde{\xi}_q(1)-\widetilde{\xi}_q(r)-(1-r)\widetilde{\xi}_q'(1))u \\
&= (1-q^2)\left(\xi'(1)-\frac{\xi'(q)^2}{\xi'(1)}\right)\widetilde{L} + (\xi(1)-\xi(q)-(1-q)\xi'(1))\,u \\
&= \frac{1}{y\xi'(1)}\left\{(\xi'(1)^2-\xi'(q)^2)(1+(1-q)z) + (\xi(1)-\xi(q)-(1-q)\xi'(1))\,z\xi'(1)\right\} \\
&= \frac{1}{y\xi'(1)}\left\{\xi'(1)^2 + z\xi'(1)(\xi(1)-\xi(q)) - \xi'(q)^2(1+(1-q)z)\right\}.
\end{aligned}$$

The last two terms of (3.37) simplify as

$$\begin{aligned}
\frac{r}{\widetilde{L}} &= \frac{q(1-q)y}{1+(1-q)z} = \frac{q(1-q)(1+z)\xi'(1)}{y(1+(1-q)z)}, \\
\int_r^1 \frac{\mathsf{d}t}{\widetilde{L}-(t-r)u} &= \frac{1}{u}\log\frac{\widetilde{L}}{\widetilde{L}-(1-r)u} = \frac{1}{u}\log(1+(1-q)z) \\
&\overset{(3.28)}{\le} \frac{1}{y}\left(z(\xi(1)-\xi(q)) + (1-q)\xi'(1)\right) - 6\eta_3.
\end{aligned}$$

It thus suffices to show

$$\begin{aligned}
&\frac{q\xi'(q)+z\xi(q)}{y} + \frac{1}{2y\xi'(1)}\left\{\xi'(1)^2 + z\xi'(1)(\xi(1)-\xi(q)) - \xi'(q)^2(1+(1-q)z)\right\} \\
&\quad + \frac{q(1-q)(1+z)\xi'(1)}{2y(1+(1-q)z)} + \frac{1}{2y}\left(z(\xi(1)-\xi(q)) + (1-q)\xi'(1)\right) \le \frac{\xi'(1)+z\xi(1)}{y}.
\end{aligned}$$

The terms in red cancel. Also clearing a factor of y, it remains to show

$$q\xi'(q) - \frac{1+(1-q)z}{2}\cdot\frac{\xi'(q)^2}{\xi'(1)} - \frac{1}{2}\left(q-\frac{q(1-q)(1+z)}{1+(1-q)z}\right)\xi'(1) \le 0.$$

Since $q-\frac{q(1-q)(1+z)}{1+(1-q)z} = \frac{q^2}{1+(1-q)z}$, the desired inequality reduces to the trivial

$$-\frac{1}{2\xi'(1)(1+(1-q)z)}\left(q\xi'(1)-(1+(1-q)z)\xi'(q)\right)^2 \le 0. \qquad \square$$

Proposition 3.3.16. *For fixed $\boldsymbol{\sigma}\in\mathcal{S}_N$, $E,R\in\mathbb{R}$, $q\in[\delta,1-\eta_2]$, the following holds. With probability $1-e^{-cN}$ conditionally on $(H_N(\boldsymbol{\sigma}),\partial_{\mathsf{rad}}H_N(\boldsymbol{\sigma}),\nabla_{\mathsf{sp}}H_N(\boldsymbol{\sigma})) = (EN,R\sqrt{N},\mathbf{0})$:*

$$\Psi(q;\boldsymbol{\sigma}) \le E_0 + \left\langle v^q, \begin{bmatrix}E-E_0\\R-R_0\end{bmatrix}\right\rangle - 2\eta_3.$$

Proof. Lemma 3.3.13, Proposition 3.3.15, and Proposition 3.2.7 imply

$$\Psi(q;\boldsymbol{\sigma}) \le E_0 + \left\langle v^q, \begin{bmatrix}E-E_0\\R-R_0\end{bmatrix}\right\rangle - 3\eta_3 + o_P(1).$$

The result follows by Proposition 3.2.15, applied to the ground state energy $\Psi(q;\boldsymbol{\sigma})$. $\square$

Lemma 3.3.17. *There exist constants $c_1, c_2 > 0$ depending only on δ such that for all $q \in [\delta, 1]$,*

$$v_E^q \geq 1 - c_1(1-q)^2, \qquad v_R^q \leq -c_2(1-q).$$

Proof. We have that

$$\begin{aligned} 1 - v_E^q &= \frac{(\xi(1)-\xi(q))(\xi'(1)+\xi''(1)) - (\xi'(1)-q\xi'(q))\xi'(1)}{\xi(1)(\xi'(1)+\xi''(1))-\xi'(1)^2}, \\ -v_R^q &= \frac{q\xi'(q)\xi(1)-\xi(q)\xi'(1)}{\xi(1)(\xi'(1)+\xi''(1))-\xi'(1)^2}. \end{aligned}$$

Note that

$$\begin{aligned} q\xi'(q)\xi(1) - \xi(q)\xi'(1) &= \left(\sum_{p\geq 2} p\gamma_p^2 q^p\right)\left(\sum_{p\geq 2}\gamma_p^2\right) - \left(\sum_{p\geq 2}\gamma_p^2 q^p\right)\left(\sum_{p\geq 2} p\gamma_p^2\right) \\ &= \sum_{p>p'\geq 2} \gamma_p^2\gamma_{p'}^2 (p-p')(q^{p'}-q^p) \\ &= (1-q)\sum_{p>p'\geq 2}\gamma_p^2\gamma_{p'}^2 q^{p'}(p-p')(1+q+\cdots+q^{p-p'-1}). \end{aligned}$$

The sum is positive and uniformly bounded away from 0 for $q \in [\delta, 1]$. Similarly

$$\begin{aligned} &(\xi(1)-\xi(q))(\xi'(1)+\xi''(1)) - (\xi'(1)-q\xi'(q))\xi'(1) \\ &= \left(\sum_{p\geq 2}\gamma_p^2(1-q^p)\right)\left(\sum_{p\geq 2} p^2\gamma_p^2\right) - \left(\sum_{p\geq 2} p\gamma_p^2(1-q^p)\right)\left(\sum_{p\geq 2} p\gamma_p^2\right) \\ &= \sum_{p>p'\geq 2}\gamma_p^2\gamma_{p'}^2(p-p')\left((1-q^{p'})p - (1-q^p)p'\right) \\ &= (1-q)^2\sum_{p>p'\geq 2}\gamma_p^2\gamma_{p'}^2(p-p')\left(\sum_{r=0}^{p'-1}(r+1)(p-p')q^r + \sum_{r=p'}^{p-2} p'(p-1-r)q^r\right) \end{aligned}$$

and the sum is uniformly bounded above for $q \in [\delta, 1]$. $\square$

Proof of Proposition 3.3.12. Consider $\boldsymbol{\sigma} \in \mathsf{Crt}(B)$ with $H_N(\boldsymbol{\sigma}) = EN$, $\partial_{\mathsf{rad}} H_N(\boldsymbol{\sigma}) = R\sqrt{N}$, so $(E, R) \in B$. We will show both conditions (i) and (ii) hold with conditional probability $1 - e^{-cN}$.

We begin with condition (i), considering a fixed $q \in [\delta, 1-\eta_2]$. Let $E = E_0 - \eta_1 + \iota_1$, $R = R_0 + \eta_1^{3/4} + \iota_2$, where $|\iota_1|, |\iota_2| \leq \eta_3$. We will show that with probability $1 - e^{-cN}$,

$$\Psi(q; \boldsymbol{\sigma}) \leq E_0 - \eta_1 - 2\eta_3. \tag{3.38}$$

By Proposition 3.3.16, it suffices to show

$$\left\langle v^q, \begin{bmatrix} -\eta_1 \\ \eta_1^{3/4} \end{bmatrix}\right\rangle + \left\langle v^q, \begin{bmatrix} \iota_1 \\ \iota_2 \end{bmatrix}\right\rangle \leq -\eta_1. \tag{3.39}$$

By Lemma 3.3.17,

$$\left\langle v^q, \begin{bmatrix} -\eta_1 \\ \eta_1^{3/4} \end{bmatrix}\right\rangle \leq -\eta_1 + c_1(1-q)^2\eta_1 - c_2(1-q)\eta_1^{3/4}.$$

Setting η_1 small enough, we can ensure that $c_1(1-q)^2\eta_1 \leq \frac{1}{2}c_2(1-q)\eta_1^{3/4}$. Since η_3 can be taken small in η_1, this proves (3.39), and (3.38) follows.

Suppose the event (3.38) holds for $q \in \{\delta, \delta+1/N, \ldots, \delta+M/N\}$ for the largest M such that $\delta+M/N \leq 1$, and that $H_N \in K_N$. This occurs with probability $1 - e^{-cN}$ by Proposition 3.2.14(ii). On K_N, for all $q \in [\delta + m/N, \delta + (m+1)/N]$,

$$\Psi(q;\boldsymbol{\sigma}) \leq \Psi(\delta + m/N;\boldsymbol{\sigma}) + O(N^{-1}) \leq E_0 - \eta_1 - 2\eta_3.$$

Thus part (i) holds.

To verify condition (ii), we will argue that with high (conditional) probability, $\boldsymbol{\sigma}$ is a "well" for H_N. Let $\theta : [0,\infty) \to \mathcal{S}_N$ be an arbitrary unit-speed geodesic on $\mathcal{S}_N$ with $\theta(0) = \boldsymbol{\sigma}$, and consider the function $f(t) = H_N(\theta(t\sqrt{N}))/N$. We have $H_N \in K_N$ with conditional probability $1 - e^{-cN}$, and on this event the C^3 norm of f is bounded independently of N. Moreover $f'(0) = 0$ since $\partial_{\mathsf{rad}} H_N(\boldsymbol{\sigma}) = 0$, while

$$f''(0) = \langle \theta'(0), \nabla^2_{\mathsf{sp}} H_N(\boldsymbol{\sigma})\theta'(0)\rangle \leq \lambda_{\max}(\nabla^2_{\mathsf{sp}} H_N(\boldsymbol{\sigma})).$$

Given $(H_N(\boldsymbol{\sigma}), \partial_{\mathsf{rad}} H_N(\boldsymbol{\sigma}))$, the conditional law of $\nabla^2_{\mathsf{sp}} H_N(\boldsymbol{\sigma})$ is (see e.g. [HS23c, Lemma 2.1]):

$$\sqrt{\xi''(1) \cdot \left(1 - \frac{1}{N}\right)} \cdot GOE(N-1) - \partial_{\mathsf{rad}} H_N(\boldsymbol{\sigma}) \cdot I_{N-1}.$$

Hence $\lambda_{\max}(\nabla^2_{\mathsf{sp}} H_N(\boldsymbol{\sigma})) \leq (2\sqrt{\xi''(1)} + \eta_3) - \partial_{\mathsf{rad}} H_N(\boldsymbol{\sigma})$ has conditional probability $1 - e^{-cN}$. We will prove condition (ii) whenever this inequality and $H_N \in K_N$ both hold. By definition of B, we then have:

$$f''(0) \leq (2\sqrt{\xi''(1)} + \eta_3) - \partial_{\mathsf{rad}} H_N(\boldsymbol{\sigma}) \leq 2\sqrt{\xi''(1)} - R_0 - \eta_1^{3/4} + 2\eta_3 \leq -\frac{\eta_1^{3/4}}{2}.$$

The final bound follows because $R_0 = y + \frac{\xi''(1)}{y} \geq 2\sqrt{\xi''(1)}$ by AM-GM, while η_3 is sufficiently small depending on η_1. Recalling that $f'(0) = 0$ and f has bounded C^3 norm, it follows that $f'(t) \neq 0$ for all $t \leq o(\eta_1^{3/4})$. Since $\theta'(0)$ was arbitrary, we conclude that H_N has no other critical point within distance $o(\eta_1^{3/4}\sqrt{N})$ of $\boldsymbol{\sigma}$. In particular for small enough η_2 depending on η_1, H_N has no critical point $\boldsymbol{\sigma}'$ with $R(\boldsymbol{\sigma}, \boldsymbol{\sigma}') \geq 1 - \eta_2$. This completes the proof. □

We will actually use Proposition 3.3.12 via the following natural corollary.

Corollary 3.3.18. *For any $\boldsymbol{\sigma} \in \mathcal{S}_N$ and $(E,R) \in B$:*

$$\begin{aligned}
&\mathbb{E}\Big[|\det \nabla^2_{\mathsf{sp}} H_N(\boldsymbol{\sigma})| \cdot \mathbf{1}\{\boldsymbol{\sigma} \notin \widetilde{\mathsf{Crt}}(B)\} \;\Big|\; \big(H_N(\boldsymbol{\sigma}), \partial_{\mathsf{rad}} H_N(\boldsymbol{\sigma}), \nabla_{\mathsf{sp}} H_N(\boldsymbol{\sigma})\big) = (EN, R\sqrt{N}, \mathbf{0})\Big] \\
&\leq e^{-cN/3}\mathbb{E}\Big[|\det \nabla^2_{\mathsf{sp}} H_N(\boldsymbol{\sigma})| \;\Big|\; \big(H_N(\boldsymbol{\sigma}), \partial_{\mathsf{rad}} H_N(\boldsymbol{\sigma}), \nabla_{\mathsf{sp}} H_N(\boldsymbol{\sigma})\big) = (EN, R\sqrt{N}, \mathbf{0})\Big].
\end{aligned}$$

Proof. Note that the conditional law of $\nabla^2_{\mathsf{sp}} H_N(\boldsymbol{\sigma})$ that of a $GOE(N-1)$ matrix scaled by $\sqrt{\frac{N-1}{N}}$ and shifted by $R \cdot I_{N-1}$. Recalling the notation (3.40), [BBM23, Theorem A.2] implies:

$$\begin{aligned}
\frac{1}{N}\log \mathbb{E}\Big[|\det \nabla^2_{\mathsf{sp}} H_N(\boldsymbol{\sigma})| \;\Big|\; \big(H_N(\boldsymbol{\sigma}), \partial_{\mathsf{rad}} H_N(\boldsymbol{\sigma}), \nabla_{\mathsf{sp}} H_N(\boldsymbol{\sigma})\big) = (EN, R\sqrt{N}, \mathbf{0})\Big] &= \kappa(R) \pm o_N(1), \\
\frac{1}{N}\log \mathbb{E}\Big[|\det \nabla^2_{\mathsf{sp}} H_N(\boldsymbol{\sigma})|^2 \;\Big|\; \big(H_N(\boldsymbol{\sigma}), \partial_{\mathsf{rad}} H_N(\boldsymbol{\sigma}), \nabla_{\mathsf{sp}} H_N(\boldsymbol{\sigma})\big) = (EN, R\sqrt{N}, \mathbf{0})\Big] &= 2\kappa(R) \pm o_N(1).
\end{aligned}$$

(See the end of [HS23c, Proof of Proposition 3.1] for further details.) By conditional Cauchy–Schwarz,

$$\begin{aligned}
&\mathbb{E}\Big[|\det \nabla^2_{\mathsf{sp}} H_N(\boldsymbol{\sigma})| \cdot \mathbf{1}\{\boldsymbol{\sigma} \notin \widetilde{\mathsf{Crt}}(B)\} \;\Big|\; \big(H_N(\boldsymbol{\sigma}), \partial_{\mathsf{rad}} H_N(\boldsymbol{\sigma}), \nabla_{\mathsf{sp}} H_N(\boldsymbol{\sigma})\big) = (EN, R\sqrt{N}, \mathbf{0})\Big] \\
&\leq \mathbb{E}\Big[|\det \nabla^2_{\mathsf{sp}} H_N(\boldsymbol{\sigma})|^2 \;\Big|\; \big(H_N(\boldsymbol{\sigma}), \partial_{\mathsf{rad}} H_N(\boldsymbol{\sigma}), \nabla_{\mathsf{sp}} H_N(\boldsymbol{\sigma})\big) = (EN, R\sqrt{N}, \mathbf{0})\Big]^{1/2} \\
&\quad \times \mathbb{P}\Big[\mathbf{1}\{\boldsymbol{\sigma} \notin \widetilde{\mathsf{Crt}}(B)\} \;\Big|\; \big(H_N(\boldsymbol{\sigma}), \partial_{\mathsf{rad}} H_N(\boldsymbol{\sigma}), \nabla_{\mathsf{sp}} H_N(\boldsymbol{\sigma})\big) = (EN, R\sqrt{N}, \mathbf{0})\Big]^{1/2}.
\end{aligned}$$

Applying Proposition 3.3.12 to the last term gives the claimed estimate. □

3.3.4 Truncated moments of critical point count via Kac-Rice

We will need the following critical point count formulas from [BSZ20]. Let

$$\Sigma = \begin{bmatrix} \xi(1) & \xi'(1) \\ \xi'(1) & \xi'(1)+\xi''(1) \end{bmatrix}, \qquad \Sigma_q = \begin{bmatrix} \xi(1) & \xi(q) & \xi'(1) & q\xi'(q) \\ \xi(q) & \xi(1) & q\xi'(q) & \xi'(1) \\ \xi'(1) & q\xi'(q) & \xi'(1)+\xi''(1) & q\xi'(q)+q^2\xi''(q) \\ q\xi'(q) & \xi'(1) & q\xi'(q)+q^2\xi''(q) & \xi'(1)+\xi''(1) \end{bmatrix}$$

be the covariances of $(\frac{1}{\sqrt{N}}H_N(\boldsymbol{\sigma}), \partial_{\mathsf{rad}}H_N(\boldsymbol{\sigma}))$ and $(\frac{1}{\sqrt{N}}H_N(\boldsymbol{\sigma}), \frac{1}{\sqrt{N}}H_N(\boldsymbol{\rho}), \partial_{\mathsf{rad}}H_N(\boldsymbol{\sigma}), \partial_{\mathsf{rad}}H_N(\boldsymbol{\rho}))$, where $R(\boldsymbol{\sigma},\boldsymbol{\rho}) = q$. Let

$$\rho(\mathsf{d}\lambda) = \frac{1}{2\pi}\sqrt{4-\lambda^2}\mathbf{1}\{|\lambda| \le 2\}\,\mathsf{d}\lambda$$

be the semicircle measure, and define

$$\begin{aligned} \kappa(x) &= \int_{\mathbb{R}} \log|\lambda - x|\rho(\mathsf{d}\lambda) && (3.40)\\ &= \frac{1}{4}x^2 - \frac{1}{2} - \mathbf{1}\{|x|>2\}\left(\frac{1}{4}|x|\sqrt{x^2-4} - \log\left(\frac{\sqrt{x^2-4}+|x|}{2}\right)\right) \\ \Theta(E,R) &= \frac{1}{2} + \frac{1}{2}\log\frac{\xi''(1)}{\xi'(1)} - \frac{1}{2}\left\langle (E,R), \Sigma^{-1}(E,R)\right\rangle + \kappa\left(R/\sqrt{\xi''(1)}\right) && (3.41)\\ \Xi(q,E_1,E_2,R_1,R_2) &= 1 + \frac{1}{2}\log\frac{(1-q^2)\xi''(1)^2}{\xi'(1)^2-\xi'(q)^2} - \frac{1}{2}\left\langle (E_1,E_2,R_1,R_2), \Sigma_q^{-1}(E_1,E_2,R_1,R_2)\right\rangle \\ &\quad + \kappa\left(R_1/\sqrt{\xi''(1)}\right) + \kappa\left(R_2/\sqrt{\xi''(1)}\right). \end{aligned}$$

Similarly to (3.30), for $A \subseteq [-1,1]\times\mathbb{R}^4$ let

$$\mathsf{Crt}_2(A) = \left\{(\boldsymbol{\sigma},\boldsymbol{\rho}) \in \mathsf{Crt}^2 : \left(R(\boldsymbol{\sigma},\boldsymbol{\rho}), \frac{1}{N}H_N(\boldsymbol{\sigma}), \frac{1}{N}H_N(\boldsymbol{\rho}), \frac{1}{\sqrt{N}}\partial_{\mathsf{rad}}H_N(\boldsymbol{\sigma}), \frac{1}{\sqrt{N}}\partial_{\mathsf{rad}}H_N(\boldsymbol{\rho})\right) \in A\right\}.$$

The next lemma, shown by the Kac–Rice formula and Laplace's method, gives the first and second moments for the relevant critical point counts. We note that although only an upper bound is stated below for the second moment, it actually holds with equality as shown in [BBM23, Appendix A].

Lemma 3.3.19 ([BSZ20, Theorems 3.1 and 3.2])**.** *For any product of intervals $A \subseteq \mathbb{R}^2$,*

$$\lim_{N\to\infty}\frac{1}{N}\log\mathbb{E}|\mathsf{Crt}(A)| = \sup_{(E,R)\in A}\Theta(E,R).$$

Furthermore, for any product of intervals $A \subseteq [-1,1]\times\mathbb{R}^4$,

$$\limsup_{N\to\infty}\frac{1}{N}\log\mathbb{E}|\mathsf{Crt}_2(A)| \le \sup_{(q,E_1,E_2,R_1,R_2)\in A}\Xi(q,E_1,E_2,R_1,R_2).$$

Lemma 3.3.20. *We have $\Theta(E_0,R_0) = \Xi(0,E_0,E_0,R_0,R_0) = 0$.*

Proof. Let $x_0 = R_0/\xi''(1)^{1/2}$. Then,

$$x_0 = \frac{y}{\xi''(1)^{1/2}} + \frac{\xi''(1)^{1/2}}{y} \tag{3.42}$$

so by Lemma 3.3.7

$$\sqrt{x_0^2-4} = \frac{y}{\xi''(1)^{1/2}} - \frac{\xi''(1)^{1/2}}{y}. \tag{3.43}$$

Also clearly $x_0 \geq 2$. It follows that

$$\kappa(x_0) = \frac{1}{4}\left(\frac{y}{\xi''(1)^{1/2}} + \frac{\xi''(1)^{1/2}}{y}\right)^2 - \frac{1}{2} - \left\{\frac{1}{4}\left(\frac{y^2}{\xi''(1)} - \frac{\xi''(1)}{y^2}\right) - \log\frac{y}{\xi''(1)^{1/2}}\right\}$$
$$= \frac{\xi''(1)}{2y^2} + \log\frac{y}{\xi''(1)^{1/2}} = \frac{\xi''(1)}{2y^2} + \frac{1}{2}\log(1+z) - \frac{1}{2}\log\frac{\xi''(1)}{\xi'(1)}.$$

By (3.35),

$$-\frac{1}{2}\left\langle (E_0,R_0), \Sigma^{-1}(E_0,R_0)\right\rangle = -\frac{zE_0+R_0}{2y} = -\frac{z\xi'(1)+z^2\xi(1)}{2y^2} - \frac{1}{2} - \frac{\xi''(1)}{2y^2}.$$

Thus,

$$\Theta(E_0,R_0) = \frac{1}{2}\log(1+z) - \frac{z\xi'(1)+z^2\xi(1)}{2y^2}$$
$$= \frac{z^2}{2(1+z)}\left(\frac{(1+z)\log(1+z)}{z^2} - \frac{1}{z} - \frac{\xi(1)}{\xi'(1)}\right) = 0.$$

Clearly $\Xi(0,E_0,E_0,R_0,R_0) = 2\Theta(E_0,R_0)$, which concludes the proof. □

Remark 3.3.21. Since we restrict attention to $\mathsf{Crt}(B)$ in this section, we do not need to verify that $\Theta(E_0,\cdot)$ is actually maximized at R_0. However this is true at least on $R \geq 2\sqrt{\xi''(1)}$ (the range corresponding to local maxima of H_N) and follows from Lemma 3.5.2 later and concavity of Θ. Hence the "annealed complexity" of local maxima at energy E_0 is indeed zero. For the special case of pure models, the appearance of the ground state energy as a threshold for annealed complexity was verified in [ABČ13].

Proof of Proposition 3.3.9. We will show that

$$\Theta(E_0-\eta_1, R_0+\eta_1^{3/4}) = \frac{z\eta_1}{y} + O(\eta_1^{9/8}). \tag{3.44}$$

Note that κ is $C^{3/2}$ on $[2,+\infty)$, with derivative

$$\kappa'(x) = \frac{1}{2}x - \frac{1}{4}\sqrt{x^2-4} - \frac{x^2}{4\sqrt{x^2-4}} + \frac{1}{\sqrt{x^2-4}} = \frac{1}{2}\left(x-\sqrt{x^2-4}\right). \tag{3.45}$$

This implies in particular that $\kappa'\big(a+\frac{1}{a}\big) = 1/a$ for $a \geq 1$. Recalling (3.42), (3.43), and using Taylor's theorem for Hölder continuous functions,

$$\kappa\left(\frac{R_0+\eta_1^{3/4}}{\sqrt{\xi''(1)}}\right) = \kappa\left(\frac{R_0}{\sqrt{\xi''(1)}}\right) + \kappa'\left(\frac{R_0}{\sqrt{\xi''(1)}}\right)\frac{\eta_1^{3/4}}{\sqrt{\xi''(1)}} + O\left((\eta_1^{3/4})^{3/2}\right)$$
$$= \kappa\left(\frac{R_0}{\sqrt{\xi''(1)}}\right) + \frac{\eta_1^{3/4}}{y} + O(\eta_1^{9/8}).$$

The function $f(E,R) = -\frac{1}{2}\langle (E,R), \Sigma^{-1}(E,R)\rangle$ is clearly analytic, so

$$f(E_0-\eta_1, R_0+\eta_1^{3/4}) = f(E_0,R_0) - \langle(-\eta_1,\eta_1^{3/4}), \Sigma^{-1}(E_0,R_0)\rangle + O(\eta_1^{3/2})$$
$$\overset{(3.35)}{=} f(E_0,R_0) + \frac{z\eta_1}{y} - \frac{\eta_1^{3/4}}{y} + O(\eta_1^{3/2}).$$

It follows that

$$\Theta(E_0-\eta_1, R_0+\eta_1^{3/4}) = \Theta(E_0,R_0) + \frac{z\eta_1}{y} + O(\eta_1^{9/8}) \overset{Lem.~3.3.20}{=} \frac{z\eta_1}{y} + O(\eta_1^{9/8}),$$

and thus, by Lemma 3.3.19,

$$\mathbb{E}|\mathsf{Crt}(B)| \geq e^{\Omega(\eta_1)N}.$$

Let $\mathcal{R}(H_N)$ denote the set of ground state typical $\boldsymbol{\sigma} \in S_N$. We apply the Kac–Rice formula (3.23) with event

$$\mathcal{E} = \left\{\boldsymbol{\sigma} \in (\mathsf{Crt}(B) \setminus \widetilde{\mathsf{Crt}}(B))\right\} = \left\{\left(\frac{1}{N} H_N(\boldsymbol{\sigma}), \frac{1}{\sqrt{N}} \partial_{\mathsf{rad}} H_N(\boldsymbol{\sigma})\right) \in B\right\} \cap \{\boldsymbol{\sigma} \notin \mathcal{R}(H_N)\}.$$

By the law of iterated expectation, the Kac–Rice formula gives

$$\begin{aligned}
\mathbb{E}|\mathsf{Crt}(B) \setminus \widetilde{\mathsf{Crt}}(B)| &= \int_{S_N} \mathbb{E}\Bigg[\mathbb{E}\left[|\det \nabla^2_{\mathsf{sp}} H_N(\boldsymbol{\sigma})| \cdot \mathbf{1}\{\boldsymbol{\sigma} \notin \mathcal{R}(H_N)\}\big| H_N(\boldsymbol{\sigma}), \partial_{\mathsf{rad}} H_N(\boldsymbol{\sigma}), \nabla_{\mathsf{sp}} H_N(\boldsymbol{\sigma})\right] \\
&\qquad \times \mathbf{1}\left\{\left(\frac{1}{N} H_N(\boldsymbol{\sigma}), \frac{1}{\sqrt{N}} \partial_{\mathsf{rad}} H_N(\boldsymbol{\sigma})\right) \in B\right\} \Bigg| \nabla_{\mathsf{sp}} H_N(\boldsymbol{\sigma}) = \mathbf{0}\Bigg] \varphi_{\nabla_{\mathsf{sp}} H_N(\boldsymbol{\sigma})}(\mathbf{0}) \, \mathsf{d}\boldsymbol{\sigma} \\
&\overset{Cor.\ 3.3.18}{\le} e^{-cN/3} \int_{S_N} \mathbb{E}\Bigg[\mathbb{E}\left[|\det \nabla^2_{\mathsf{sp}} H_N(\boldsymbol{\sigma})|\big| H_N(\boldsymbol{\sigma}), \partial_{\mathsf{rad}} H_N(\boldsymbol{\sigma}), \nabla_{\mathsf{sp}} H_N(\boldsymbol{\sigma})\right] \\
&\qquad \times \mathbf{1}\left\{\left(\frac{1}{N} H_N(\boldsymbol{\sigma}), \frac{1}{\sqrt{N}} \partial_{\mathsf{rad}} H_N(\boldsymbol{\sigma})\right) \in B\right\} \Bigg| \nabla_{\mathsf{sp}} H_N(\boldsymbol{\sigma}) = \mathbf{0}\Bigg] \varphi_{\nabla_{\mathsf{sp}} H_N(\boldsymbol{\sigma})}(\mathbf{0}) \, \mathsf{d}\boldsymbol{\sigma} \\
&= e^{-cN/3} \mathbb{E}|\mathsf{Crt}(B)|.
\end{aligned}$$

Thus,

$$\mathbb{E}|\widetilde{\mathsf{Crt}}(B)| \ge (1 - e^{-cN/3}) \mathbb{E}|\mathsf{Crt}(B)| \ge e^{\Omega(\eta_1) N}.$$

□

We defer the proofs of the following two lemmas to Subsection 3.3.5.

Lemma 3.3.22. *Let $V \subseteq \mathcal{S}_N$ be a finite set of points with $|V| = M$.*

(a) There exists $V_1 \subseteq V$ such that $|V_1| = M^{\Omega(\delta)}$ and $R(\boldsymbol{\sigma}, \boldsymbol{\rho}) \ge -\delta$ for all $\boldsymbol{\sigma}, \boldsymbol{\rho} \in V_1$.

(b) There exists $V_2 \subseteq V \times V$ such that $|V_2| = \Omega(M^2 \delta)$ and $R(\boldsymbol{\sigma}, \boldsymbol{\rho}) \ge -\delta$ for all $(\boldsymbol{\sigma}, \boldsymbol{\rho}) \in V_2$.

Lemma 3.3.23. *If the function $q \mapsto \xi''(q)^{-1/2}$ is convex on $[0, 1]$, then ξ is strictly 1RSB.*

Proof of Proposition 3.3.10. By definition of ground state typical, a.s. any distinct $\boldsymbol{\sigma}, \boldsymbol{\rho} \in \widetilde{\mathsf{Crt}}(B)$ satisfy $R(\boldsymbol{\sigma}, \boldsymbol{\rho}) \le \delta$. Let $V_2 \subseteq \widetilde{\mathsf{Crt}}(B)^2$ be the set of $(\boldsymbol{\sigma}, \boldsymbol{\rho})$ which furthermore satisfy $R(\boldsymbol{\sigma}, \boldsymbol{\rho}) \ge -\delta$. By Lemma 3.3.22(b), $|\widetilde{\mathsf{Crt}}(B)|^2 \le C\delta^{-1}|V_2|$ for an absolute constant C. Also, a.s. $V_2 \subseteq \mathsf{Crt}_2(B_2)$ where

$$B_2 = [-\delta, \delta] \times [E_0 - \eta_1 - \eta_3, E_0 - \eta_1 + \eta_3]^2 \times [R_0 + \eta_1^{2/3} - \eta_3, R_0 + \eta_1^{2/3} + \eta_3]^2.$$

Thus, by Lemma 3.3.19,

$$\mathbb{E}|\widetilde{\mathsf{Crt}}(B)|^2 \le C\delta^{-1} \mathbb{E}|\mathsf{Crt}_2(B_2)| \le C\delta^{-1} \exp\{N \sup \Xi(B_2) + o(N)\}.$$

It is clear that Ξ is locally Lipschitz near $(0, E_0, E_0, R_0, R_0)$. By Lemma 3.3.20, $\Xi(0, E_0, E_0, R_0, R_0) = 0$, so $\sup \Xi(B_2) \le O(\delta)$. The result follows. □

Proof of Proposition 3.2.10. We first prove the proposition for non-pure ξ, as we have been assuming throughout the section. The statement is monotone in δ, so we may assume δ is small in ε. By Propositions 3.3.9 and 3.3.10 and Paley-Zygmund,

$$\mathbb{P}\left(|\widetilde{\mathsf{Crt}}(B)| \ge \frac{1}{2} \mathbb{E}|\widetilde{\mathsf{Crt}}(B)|\right) \ge e^{-O(\delta) N}. \tag{3.46}$$

Suppose this event holds. By Lemma 3.3.22(a), there exists $V_1 \subseteq \widetilde{\mathsf{Crt}}(B)$ with

$$|V_1| \ge \left(\frac{1}{2} \mathbb{E}|\widetilde{\mathsf{Crt}}(B)|\right)^{O(\delta)} = e^{O(\delta \eta_1 N)}$$

such that $|R(\boldsymbol{\sigma}, \boldsymbol{\rho})| \le \delta$ for all $\boldsymbol{\sigma}, \boldsymbol{\rho} \in V_1$. By the choice of B, all these points have energy at least

$$E_0 - \eta_1 - \eta_3 \ge \mathcal{Q}(\xi) - \varepsilon/2,$$

where we recall $E_0 = \mathcal{Q}(\xi)$ and take η_1, η_3 small in ε. Let

$$X = \frac{1}{N} \sup_{\vec{\boldsymbol{\sigma}} \in \mathsf{Band}_{k,1,\delta}(\mathbf{0})} \inf_{i \in [k]} H_N(\boldsymbol{\sigma}^i).$$

Combining the above shows that for any $k \le e^{O(\delta \eta_1 N)}$,

$$\mathbb{P}\left(X \ge \mathcal{Q}(\xi) - \varepsilon/2\right) \ge e^{-O(\delta)N}.$$

A direct calculation shows that for fixed $\boldsymbol{\sigma}$, $H_N(\boldsymbol{\sigma})$ is $O(N^{1/2})$-Lipschitz in the disorder Gaussians. Since suprema and infima preserve Lipschitz constants, NX is also $O(N^{1/2})$-Lipschitz in the disorder Gaussians. We thus have the concentration inequality

$$\mathbb{P}(|X - \mathbb{E}X| \ge t) \le \exp(-ct^2 N).$$

Combining the last two inequalities implies that (for δ small in ε) $\mathbb{P}(X \ge \mathcal{Q}(\xi) - \varepsilon) \ge 1 - e^{-cN}$. This completes the proof for non-pure ξ.

Finally, we turn to the case where $\xi(q) = \beta^2 q^p$ is pure. Then, $H_N(\boldsymbol{\sigma}) = \beta H_N^{(p)}(\boldsymbol{\sigma})$, where $H_N^{(p)}(\boldsymbol{\sigma}) = \langle \boldsymbol{G}^{(p)}, \boldsymbol{\sigma}^{\otimes p} \rangle$. Consider a perturbation $\widehat{H}_N(\boldsymbol{\sigma}) = \beta H_N^{(p)}(\boldsymbol{\sigma}) + \iota \beta H_N^{(p+1)}(\boldsymbol{\sigma})$, for a fixed $\iota > 0$ chosen small in δ, ε. This has mixture $\widehat{\xi}(q) = \beta^2(q^p + \iota^2 q^{p+1})$. Note that

$$\widehat{\xi}''(q)^{-1/2} = \frac{q^{-(p-2)/2}}{\beta\sqrt{p(p-1)}} \left(1 + \frac{\iota^2(p+1)}{p-1} q\right)^{-1/2}$$

is convex on $[0,1]$, so Lemma 3.3.23 implies $\widehat{\xi}$ is strictly 1RSB. By the result for non-pure ξ, there exists $c = c(\widehat{\xi}, \delta, \varepsilon/2)$ such that for all $k \le e^{cN}$, with probability $1 - e^{-cN}$ there exists $\vec{\boldsymbol{\sigma}} \in \mathsf{Band}_{k,1,\delta}(\mathbf{0})$ such that for all $i \in [k]$,

$$\frac{1}{N}\widehat{H}_N(\boldsymbol{\sigma}^i) \ge \mathcal{Q}(\widehat{\xi}, 0) - \frac{\varepsilon}{2} \ge \mathcal{Q}(\xi) - \frac{\varepsilon}{2}.$$

By Proposition 3.2.14, there exists a constant C such that $\frac{1}{N}\sup_{\boldsymbol{\sigma} \in \mathcal{S}_N} H_N^{(p+1)}(\boldsymbol{\sigma}) \le C$ with probability $1 - e^{-cN}$. On the intersection of these events, for each $i \in [k]$,

$$\frac{1}{N} H_N(\boldsymbol{\sigma}^i) \ge \frac{1}{N}\widehat{H}_N(\boldsymbol{\sigma}^i) - \frac{\iota}{N} H_N^{(p+1)}(\boldsymbol{\sigma}^i) \ge \mathcal{Q}(\xi) - \frac{\varepsilon}{2} - C\iota \ge \mathcal{Q}(\xi) - \varepsilon$$

for ι small in ε. $\square$

3.3.5 Deferred proofs

Proof of Lemma 3.3.22. Consider the graph G with vertex set V where $(\boldsymbol{\sigma}, \boldsymbol{\rho})$ is an edge if $R(\boldsymbol{\sigma}, \boldsymbol{\rho}) < -\delta$. Note that G does not contain a $r = \lceil 1/\delta \rceil$-clique, because if such a clique $U \subseteq V$ existed, then the Gram matrix $[R(\boldsymbol{\sigma}, \boldsymbol{\rho})]_{\boldsymbol{\sigma}, \boldsymbol{\rho} \in U}$ would not be positive semi-definite.

Let $\mathfrak{R}(s,t)$ denote the (s,t) Ramsey number. Recall the classic Ramsey upper bound

$$\mathfrak{R}(s,t) \le \binom{s+t-2}{s-1},$$

which can be proved by applying the inequality $\mathfrak{R}(s,t) \le \mathfrak{R}(s-1,t) + \mathfrak{R}(s,t-1)$ recursively. Thus $\mathfrak{R}(r, M^{\delta/2}) \lesssim M^{(r-1)\delta/2} \ll M$, so G contains a $M^{\delta/2}$-independent set. This proves part (a). Since G avoids an r-clique, by Turán's theorem G avoids at least $\frac{1}{r-1} = O(\delta)$ fraction of edges, proving (b). $\square$

Lemma 3.3.24. *In any model with $\xi'(0) = 0$, we have $0 \in T$.*

Proof. Assume otherwise and let $q \in (0,1]$ be the minimal point in T. Then $g(q) = 0$. If $q < 1$, then q is an interior local minimizer of g, so $0 = g'(q) = -G(q)$; if $q = 1$, then the characterization from Lemma 3.2.3 implies $G(q) = 0$. So in either case $G(q) = 0$. Also from the definition (3.13) of G we have $G(0) = 0$.

Recall that the measure ν given by $\nu([0,s]) = \alpha(s)$ is supported on T. Thus $\alpha \equiv 0$ on $[0,q)$, and so $\widehat{\alpha}$ is constant on $[0,q]$. Therefore G is convex on $[0,q]$. Since $G(0) = G(q) = 0$, this implies $G \le 0$ on $[0,q]$.

Thus $g(0) = g(q) + \int_0^q G(s)\ \mathsf{d}s \le g(q) = 0$. So $0 \in T$, contradicting minimality of q. □

Proof of Lemma 3.3.23. Lemma 3.3.24 implies $0 \in T$, and Lemma 3.2.3 ensures $1 \in T$. The following argument, adapted from [Tal06a, Proposition 2.2], shows that $|T| \le 2$, which then implies $T = \{0,1\}$.

Consider any $q_1, q_2 \in T$ such that $q_1 < q_2$. If $q_i \in (0,1)$, then $G(q_i) = 0$ because q_i is an interior local minimizer of g; if $q_i = 1$ then $G(q_i) = 0$ by Lemma 3.2.3; and if $q_i = 0$ then $G(q_i) = 0$ by definition (3.13) of G. So $G(q_1) = G(q_2) = 0$. Moreover $g(q_1) = g(q_2) = 0$, so $\int_{q_1}^{q_2} G(s)\ \mathsf{d}s = 0$. It follows that there are two points in $q_3, q_4 \in (q_1, q_2)$ such that $G'(q_3) = G'(q_4) = 0$.

We have thus shown that between any two elements of T lies two zeros of G'. However, $G'(q) = \xi''(q) - \frac{1}{\widehat{\alpha}(q)^2}$, so at any zero of G' we have $\xi''(q)^{-1/2} = \widehat{\alpha}(q)$. However $\xi''(q)^{-1/2}$ is convex by assumption, while $\widehat{\alpha}(q)$ is concave by definition, so these functions intersect at most twice. It follows that $|T| \le 2$. □

3.4 Building a model from fundamental types: proof of Theorem 3.1.7

In this section we complete the proof of Theorem 3.1.7. The proof proceeds in two steps:

(1) Using the lower bounds in Propositions 3.2.9 through 3.2.12 and a uniform concentration lemma due to Subag (Lemma 3.4.6 below) we prove Theorem 3.4.4 below, which constructs an ultrametric tree with somewhat more lenient constraints than Theorem 3.1.7.

(2) By pruning this ultrametric tree we arrive at the ultrametric tree in Theorem 3.1.7.

Before giving the full proof, we note the free energy lower bound Corollary 3.1.8 follows by combining what we have done with [Sub24], where the decomposition approach we follow was introduced. Namely [Sub24, Theorem 5] used the existence of many orthogonal replicas to lower bound the free energy by a sum of the ground state on a subsphere $\sqrt{q}\mathcal{S}_N$ plus the free energy of a "band" model centered at a typical point on this subsphere. Using this idea sequentially with $q = q_0, q_1, \dots$ yields Corollary 3.1.8 since each intermediate model takes one of the four fundamental types analyzed previously in this paper.

Our analysis below is conceptually similar to [Sub24], but constructs $e^{\Omega(N)}$ near-orthogonal approximate ground states in each intermediate model, and hence gives a larger tree of pure states than was known to exist previously by any method in this generality. By contrast [Sub24] relies on Chatterjee's superconcentration, which only gives a slowly diverging number of near-orthogonal approximate ground states. This improvement is also what allows us to prove the lower tail large deviations for the ground state have speed at least N^2 (see Subsection 3.5.2).

3.4.1 A tree with local constraints

Lemma 3.4.1. *For any model ξ, the set S (recall (3.6)) is a disjoint union of finitely many closed intervals, possibly including atoms. Moreover $q_D < 1$.*

Proof. The first statement follows from [JT18, Corollary 1.6] and the observation that $\mathfrak{d} = \left(\frac{1}{\sqrt{\xi''}}\right)''$ changes sign finitely many times on $[0,1]$. Indeed, S is precisely the *coincidence set* denoted $\{\eta = \xi\}$, as can be seen from the display between (1.1.1) and (1.1.2) therein.

For the second statement, note that $\widehat{x}(q) = 1 - q$ for q in some interval $[\widehat{q}, 1]$, so $F(q) \le C - \frac{1}{1-q}$ for some C independent of q. Hence $f(q) \le Cq + \log(1-q)$, so $\lim_{q \to 1^-} f(q) = -\infty$, which implies $q_D < 1$. □

Definition 3.4.2. A sequence $q_0, \dots, q_D$ with $0 \le q_0 < \cdots < q_D \le 1$ is an ***S*-refinement** if

$$\partial S \subseteq \{q_0, \dots, q_D\} \subseteq S,$$

where $\partial S = S \cap \overline{S^c}$ is the boundary of S in $\mathbb{R}$. (In particular $q_0 = \inf(S)$ and $q_D = \sup(S)$.)

In the following variant of Definition 3.1.6, the orthogonality constraints are enforced only locally. For $u, v \in \mathbb{T}$, write $u \sim v$ if $u = v$, or u, v are siblings, or one of u, v is the parent of the other.

Definition 3.4.3. Let $k, D \in \mathbb{N}$, $0 \le q_0 < \cdots < q_D \le 1$, $\vec{q} = (q_0, \ldots, q_D)$, and $\delta > 0$. A $(k, D, \vec{q}, \delta)$-locally ultrametric tree is a collection of points $(\boldsymbol{\sigma}^u)_{u \in \mathbb{T}}$ such that (3.7) holds for all $u \sim v$.

We will first prove the following variant of Theorem 3.1.7, where the properties required of the ultrametric tree are relaxed in two ways: ultrametricity will be enforced only locally, and we will lower bound the average energy increment from each node to its children rather than the energy of each node. We will deduce Theorem 3.1.7 from Theorem 3.4.4 in Subection 3.4.3 by pruning this tree.

Theorem 3.4.4. *For any $\delta, \varepsilon > 0$, $D \in \mathbb{N}$, and S-refinement $q_0, \ldots, q_D$, there exists $c > 0$ such that the following holds for any $k \le e^{cN}$. With probability $1 - e^{-cN}$, there is a $(k, D, \vec{q}, \delta)$-locally ultrametric tree $(\boldsymbol{\sigma}^u)_{u \in \mathbb{T}}$ with the following properties.*

(i) Energy of root: $\frac{1}{N} H_N(\boldsymbol{\sigma}^{\emptyset}) \ge E(q_0) - \varepsilon$.

(ii) Parent-to-child energy increments: for each $u \in \mathbb{T} \setminus \mathbb{L}$,

$$\frac{1}{kN} \sum_{i=1}^{k} \left(H_N(\boldsymbol{\sigma}^{ui}) - H_N(\boldsymbol{\sigma}^u) \right) \ge E(q_{|u|+1}) - E(q_{|u|}) - \varepsilon.$$

(iii) Free energy of pure states: for each $u \in \mathbb{L}$,

$$\frac{1}{kN} \log \int_{\mathsf{Band}_{k,1,\delta}(\boldsymbol{\sigma}^u)} \exp\left(\sum_{i=1}^{k} \left(H_N(\boldsymbol{\rho}^i) - H_N(\boldsymbol{\sigma}^u) \right) \right) \, \mathsf{d}\vec{\boldsymbol{\rho}} \ge \mathsf{P}(\xi) - E(q_D) - \varepsilon.$$

3.4.2 Model decomposition into fundamental types

Fix parameters $\delta, \varepsilon, D, (q_0, \ldots, q_D)$ as in Theorem 3.4.4. We take as convention $q_{-1} = 0$, $q_{D+1} = 1$. Define $\xi_{-1}(x) = \xi(q_0 x)$ and, for $0 \le d \le D$,

$$\xi_d(x) = \xi(q_d + (q_{d+1} - q_d)x) - \xi(q_d) - \xi'(q_d)(q_{d+1} - q_d)x. \tag{3.47}$$

The following proposition gives the link between ξ and each ξ_d from the point of view of the Parisi formula. It follows by matching order parameters, see [Sub24, Proposition 11].

Proposition 3.4.5. *The following hold.*

(a) The model ξ_{-1} is topologically trivial and satisfies $\mathcal{Q}(\xi_{-1}) = E(q_0)$. (We treat this part as vacuous if $q_0 = 0$, in which case $\xi_{-1} \equiv 0$.)

(b) For each $0 \le d \le D-1$, ξ_d is either strictly 1RSB or strictly FRSB, and satisfies $\mathcal{Q}(\xi_d) = E(q_{d+1}) - E(q_d)$.

(c) The model ξ_D is strictly RS and satisfies $\mathsf{P}(\xi_D) = \mathsf{P}(\xi) - E(q_D) - \frac{1}{2}\log(1 - q_D)$.

For $0 \le d \le D-1$ and $\|\boldsymbol{\sigma}\|_2 = \sqrt{q_d N}$, define

$$F_{d,k}(\boldsymbol{\sigma}) = \frac{1}{kN} \max_{\vec{\boldsymbol{\rho}} \in \mathsf{Band}_{k, q_{d+1}, \delta}(\boldsymbol{\sigma})} \sum_{i=1}^{k} \left(H_N(\boldsymbol{\rho}^i) - H_N(\boldsymbol{\sigma}) \right),$$

and for $\|\boldsymbol{\sigma}\|_2 = \sqrt{q_D N}$,

$$F_{D,k}(\boldsymbol{\sigma}) = \frac{1}{kN} \log \int_{\mathsf{Band}_{k,1,\delta}(\boldsymbol{\sigma})} \exp\left(\sum_{i=1}^{k} \left(H_N(\boldsymbol{\rho}^i) - H_N(\boldsymbol{\sigma}) \right) \right) \, \mathsf{d}\vec{\boldsymbol{\rho}}.$$

The following **uniform concentration** lemma was proved at finite temperature in [Sub24, Proposition 1], as consequence of the concentration of Lipschitz functions of Gaussians. Its proof at zero temperature is identical.

Lemma 3.4.6. *For all $\varepsilon > 0$, there exists $\delta_0 = \delta_0(\xi, \varepsilon)$ and $k_0 = k_0(\xi, \varepsilon)$ such that for all $\delta \le \delta_0$, $k \ge k_0$ the following holds with probability $1 - e^{-cN}$. For all $0 \le d \le D$ and all $\|\boldsymbol{\sigma}\|_2 = \sqrt{q_d N}$, $|F_{d,k}(\boldsymbol{\sigma}) - \mathbb{E}F_{d,k}(\boldsymbol{\sigma})| \le \varepsilon$.*

Proposition 3.4.7. *There exists $c > 0$ (depending on $\delta, k, D, q_0, \dots, q_D$) such that the following holds.*

(a) *With probability $1 - e^{-cN}$, there exists $\|\boldsymbol{\sigma}\|_2 = \sqrt{q_0 N}$ such that $\frac{1}{N} H_N(\boldsymbol{\sigma}) \ge E(q_0) - \varepsilon$.*

(b) *For all $k \le e^{cN}$, $0 \le d \le D - 1$ and any fixed $\boldsymbol{\sigma}$ with $\|\boldsymbol{\sigma}\|_2 = \sqrt{q_d N}$:*

$$\mathbb{E}F_{d,k}(\boldsymbol{\sigma}) \ge E(q_{d+1}) - E(q_d) - \varepsilon.$$

(c) *For all $k \le e^{cN}$ and any fixed $\boldsymbol{\sigma}$ with $\|\boldsymbol{\sigma}\|_2 = \sqrt{q_D N}$:*

$$\mathbb{E}F_{D,k}(\boldsymbol{\sigma}) \ge \mathsf{P}(\xi) - E(q_D) - \varepsilon.$$

Proof. Since $H_N^{(-1)}(\boldsymbol{\rho}) \equiv H_N(\sqrt{q_0}\boldsymbol{\rho})$ has mixture ξ_{-1}, part (a) follows by Propositions 3.4.5(a) and 3.2.11.

Next we prove part (b). For $0 \le d \le D - 1$, and fixed $\|\boldsymbol{\sigma}\|_2 = \sqrt{q_d N}$, the model

$$H_N^{(d,\boldsymbol{\sigma})}(\boldsymbol{\rho}) = H_N(\sqrt{q_{d+1} - q_d}\boldsymbol{\rho} + \boldsymbol{\sigma}) - H_N(\boldsymbol{\sigma}) - \sqrt{q_{d+1} - q_d}\langle \nabla H_N(\boldsymbol{\sigma}), \boldsymbol{\rho} \rangle,$$

restricted to the band $\boldsymbol{\sigma}^\perp = \{\boldsymbol{\rho} \in \mathcal{S}_N : R(\boldsymbol{\rho}, \boldsymbol{\sigma}) = 0\}$ is a $(N-1)$-dimensional model with mixture ξ_d. By Proposition 3.4.5(b), this model is either strictly 1RSB or strictly FRSB. By Propositions 3.2.10 and 3.2.12, combined with concentration via Borell-TIS, we thus find:

$$\frac{1}{kN}\mathbb{E} \max_{\vec{\boldsymbol{\rho}} \in \mathsf{Band}_{k,1,\delta}(\mathbf{0}) \cap (\boldsymbol{\sigma}^\perp)^k} \sum_{i=1}^k H_N^{(d,\boldsymbol{\sigma})}(\boldsymbol{\rho}^i) \ge E(q_{d+1}) - E(q_d) - \varepsilon.$$

Still with $\boldsymbol{\sigma}$ fixed, let $\vec{\boldsymbol{\rho}}_*$ attain the maximum in the previous display. Then we have the inequality chain

$$\begin{aligned}\mathbb{E}F_{d,k}(\boldsymbol{\sigma}) &\ge \frac{1}{kN}\mathbb{E} \max_{\vec{\boldsymbol{\rho}} \in \mathsf{Band}_{k,1,\delta}(\mathbf{0}) \cap (\boldsymbol{\sigma}^\perp)^k} \sum_{i=1}^k \left(H_N(\sqrt{q_{d+1} - q_d}\boldsymbol{\rho}^i + \boldsymbol{\sigma}) - H_N(\boldsymbol{\sigma}) \right) \\ &\ge \frac{1}{kN}\mathbb{E} \sum_{i=1}^k \left(H_N(\sqrt{q_{d+1} - q_d}\boldsymbol{\rho}_*^i + \boldsymbol{\sigma}) - H_N(\boldsymbol{\sigma}) \right) \\ &\overset{(*)}{=} \frac{1}{kN}\mathbb{E} \sum_{i=1}^k H_N^{(d,\boldsymbol{\sigma})}(\boldsymbol{\rho}_*^i) \ge E(q_{d+1}) - E(q_d) - \varepsilon,\end{aligned}$$

where the step $(*)$ uses that $H_N^{(d,\boldsymbol{\sigma})}$ is independent of $\nabla H_N(\boldsymbol{\sigma})$ as a process. This proves part (b).

The proof of (c) is similar. The model $H_N^{(D,\boldsymbol{\sigma})}$ restricted to $\boldsymbol{\sigma}^\perp$ is a $(N-1)$-dimensional model with mixture ξ_D. By Proposition 3.4.5(c), this model is strictly RS with respect to the normalized $N-2$ dimensional Hausdorff measure $\widetilde{\mathcal{H}}_{N-2}$ on $\boldsymbol{\sigma}^\perp$. By Propositions 3.2.9 and 3.4.5(c), with probability $1 - e^{-cN}$:

$$\frac{1}{kN} \log \int_{\mathsf{Band}_{k,1,\delta^2}(\mathbf{0}) \cap (\boldsymbol{\sigma}^\perp)^k} \exp\left(\sum_{i=1}^k H_N^{(D,\boldsymbol{\sigma})}(\boldsymbol{\rho}^i) \right) \, \mathsf{d}\widetilde{\mathcal{H}}_{N-2}^k(\vec{\boldsymbol{\rho}}) \ge \mathsf{P}(\xi) - E(q_D) - \frac{1}{2}\log(1 - q_D) - \varepsilon/3.$$

Moreover since $H_N \in K_N$ with probability $1 - e^{-cN}$, we easily find that with high probability, the restricted free energy with respect to the original uniform measure on $\mathcal{S}_N$ obeys a similar bound:

$$\frac{1}{kN} \log \int_{\mathsf{Band}_{k,1,\delta}(\boldsymbol{\sigma})} \exp\left(\sum_{i=1}^k H_N(\boldsymbol{\rho}^i) - H_N(\boldsymbol{\sigma}) \right) \, \mathsf{d}\vec{\boldsymbol{\rho}} \ge \mathsf{P}(\xi) - E(q_D) - \varepsilon/2.$$

Here the term $\frac{1}{2}\log(1 - q_D)$ disappeared from rescaling. By Lipschitz concentration of the left-hand side,

$$\frac{1}{kN}\mathbb{E} \int_{\mathsf{Band}_{k,1,\delta}(\boldsymbol{\sigma})} \exp\left(\sum_{i=1}^k H_N(\boldsymbol{\rho}^i) - H_N(\boldsymbol{\sigma}) \right) \, \mathsf{d}\vec{\boldsymbol{\rho}} \ge \mathsf{P}(\xi) - E(q_D) - 2\varepsilon/3.$$

This concludes the proof. ∎

Proof of Theorem 3.4.4. It suffices to prove the theorem for $k = e^{cN}$ and $\delta \le \delta_0(\xi, \varepsilon)$, as the statement is clearly monotone in k and δ. As k is growing in N, $k \ge k_0(\xi, \varepsilon)$ and Lemma 3.4.6 holds.

By Proposition 3.4.7, for $d \le D-1$ and any fixed $\|\boldsymbol{\sigma}\|_2 = \sqrt{q_d N}$, $\|\boldsymbol{\sigma}\|_2 = \sqrt{q_D N}$, respectively,

$$\mathbb{E} F_{d,k}(\boldsymbol{\sigma}) \ge E(q_{d+1}) - E(q_d) - \varepsilon/2, \qquad \mathbb{E} F_{D,k}(\boldsymbol{\sigma}) \ge \mathsf{P}(\xi) - E(q_D) - \varepsilon/2.$$

By Lemma 3.4.6, with probability $1 - e^{-cN}$, for **all** $\|\boldsymbol{\sigma}\|_2 = \sqrt{q_d N}$, $\|\boldsymbol{\sigma}\|_2 = \sqrt{q_D N}$, respectively,

$$F_{d,k}(\boldsymbol{\sigma}) \ge E(q_{d+1}) - E(q_d) - \varepsilon, \qquad F_{D,k}(\boldsymbol{\sigma}) \ge \mathsf{P}(\xi) - E(q_D) - \varepsilon. \tag{3.48}$$

By Proposition 3.4.7(a), with probability $1 - e^{-cN}$ there exists $\left\|\boldsymbol{\sigma}^{\emptyset}\right\|_2 = \sqrt{q_0 N}$ such that $\frac{1}{N} H_N(\boldsymbol{\sigma}^{\emptyset}) \ge E(q_0) - \varepsilon$. Starting from this point, we can construct the remaining $\boldsymbol{\sigma}^u$ using (3.48). $\square$

3.4.3 Pruning the relaxed tree

We apply Theorem 3.4.4 with parameters $(\delta^2/2D^4, \varepsilon/2(D+1))$ in place of (δ, ε). Let $c > 0$ be given by this theorem and $k = e^{cN}$. Then, with probability $1 - e^{-cN}$, there is a $(k, D, \vec{q}, \delta)$-locally ultrametric tree $(\boldsymbol{\sigma}^u)_{u \in \mathbb{T}}$, where $\mathbb{T} = \mathbb{T}(k, D)$, with properties (i), (ii), (iii) (where δ, ε are replaced by $\delta^2/2D^4, \varepsilon/2(D+1)$). Throughout this subsection, assume this event holds and K_N from Proposition 3.2.14 holds.

Let $c_0 = c/2D$ and $k'' = e^{c_0 N}$. We will show that for a subtree $\mathbb{T}'' \cong \mathbb{T}(k'', D)$ of $\mathbb{T}$, $(\boldsymbol{\sigma}^u)_{u \in \mathbb{T}''}$ has the properties described in Theorem 3.1.7. We obtain $\mathbb{T}''$ from $\mathbb{T}$ by two steps of pruning: we first ensure all energies are suitably large (Proposition 3.4.8), and then that global overlap constraints are satisfied (Proposition 3.4.9).

Proposition 3.4.8. *For an absolute constant $C = C(\xi)$ and $k' = \frac{\varepsilon}{CD} e^{cN}$, there exists a subtree $\mathbb{T}' \cong \mathbb{T}(k', D)$ of $\mathbb{T}$ such that the following holds. For each $u \in \mathbb{T}' \setminus \mathbb{L}$ and all $i \in [k]$ such that $ui \in \mathbb{T}'$,*

$$\frac{1}{N}\left(H_N(\boldsymbol{\sigma}^{ui}) - H_N(\boldsymbol{\sigma}^u)\right) \ge E(q_{|u|+1}) - E(q_{|u|}) - \varepsilon/(D+1). \tag{3.49}$$

Proof. We will construct $\mathbb{T}'$ by breadth-first exploration starting from the root: at every non-leaf u we encounter, we will find k' children of it such that (3.49) holds.

Consider one such u, and abbreviate $\Delta_i = \frac{1}{N}\left(H_N(\boldsymbol{\sigma}^{ui}) - H_N(\boldsymbol{\sigma}^u)\right)$, $\Delta E_{|u|} = E(q_{|u|+1}) - E(q_{|u|})$. By property (ii) of Theorem 3.4.4,

$$\frac{1}{k}\sum_{i=1}^{k} \Delta_i \ge \Delta E_{|u|} - \varepsilon/2(D+1).$$

On event K_N, $\sup_{\|\boldsymbol{x}\|_2 \le \sqrt{N}} |H_N(\boldsymbol{\sigma})| \le C_0 N$, so deterministically $|\Delta_i| \le 2C_0$ for all $i \in [k]$. By Markov's inequality on $\mathrm{unif}([k])$,

$$\begin{aligned}\frac{1}{k}\left|i \in [k] : \Delta_i \le \Delta E_{|u|} - \varepsilon/(D+1)\right| &= \frac{1}{k}\left|i \in [k] : 2C_0 - \Delta_i \ge 2C_0 - \Delta E_{|u|} - \varepsilon/(D+1)\right| \\ &\le \frac{2C_0 - \Delta E_{|u|} - \varepsilon/(D+1)}{2C_0 - \Delta_i \le 2C_0 - \Delta E_{|u|} - \varepsilon/2(D+1)},\end{aligned}$$

and thus

$$\frac{1}{k}\left|i \in [k] : \Delta_i \ge \Delta E_{|u|} - \varepsilon/(D+1)\right| \ge \frac{\varepsilon/2D}{2C_0 - \Delta_i \le 2C_0 - \Delta E_{|u|} - \varepsilon/2(D+1)} \ge \frac{\varepsilon}{4C_0(D+1)}.$$

Setting $C = 8C_0$, we conclude that we can find k' children of u such that (3.49) holds. $\square$

Proposition 3.4.9. *There is a subtree $\mathbb{T}'' \cong \mathbb{T}(k'', D)$ of $\mathbb{T}'$ such that for any distinct parent-child pairs (u, ui), (v, vj) in $\mathbb{T}''$ (where possibly $u = v$), $|R(\boldsymbol{\sigma}^{ui} - \boldsymbol{\sigma}^u, \boldsymbol{\sigma}^{vj} - \boldsymbol{\sigma}^v)| \le \delta/D^2$.*

Proof. We will construct $\mathbb{T}''$ by breadth-first exploration starting from the root. We will abbreviate $\boldsymbol{x}^{ui} = \boldsymbol{\sigma}^{ui} - \boldsymbol{\sigma}^{u}$. We maintain a set

$$\mathcal{C} = \{\boldsymbol{x}^{ui} : (u, ui) \text{ is a parent-child pair in } \mathbb{T}''\},$$

and will maintain the invariant that $|R(\boldsymbol{x}, \boldsymbol{y})| \le \delta/D^2$ for any distinct $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{C}$. At every non-leaf u we encounter in the exploration, we will find k'' children of it such that, when the corresponding k'' parent-child pairs are added to $\mathbb{T}''$, this invariant continues to hold. Note that at all times,

$$|\mathcal{C}| \le k'' + (k'')^2 + \cdots + (k'')^D \le 2(k'')^D = 2e^{cN/2}.$$

Consider the step in this procedure where we choose children for node u. Let $I_u = \{i \in [k] : ui \in \mathbb{T}'\}$, so $|I_u| = k'$. For $\boldsymbol{y} \in \mathcal{C}$, let $I_u^+(\boldsymbol{y}), I_u^-(\boldsymbol{y})$ be the sets of $i \in [k]$ such that $R(\boldsymbol{y}, \boldsymbol{x}^{ui}) > \delta/D^2$ and $R(\boldsymbol{y}, \boldsymbol{x}^{ui}) < -\delta/D^2$. We **claim** that at any such step, $|I_u^+(\boldsymbol{y})| \le 2D^4/\delta^2$ and $|I_u^-(\boldsymbol{y})| \le 2D^4/\delta^2$. This claim suffices, since it implies

$$\left| I_u \setminus \bigcup_{\boldsymbol{y} \in \mathcal{C}} (I_u^+(\boldsymbol{y}) \cup I_u^-(\boldsymbol{y})) \right| \ge k' - 2e^{cN/2} \cdot \frac{4D^4}{\delta^2} \gg k''.$$

Indeed, we may choose any k'' elements i in this set and add the corresponding ui to $\mathbb{T}''$, which preserves the required invariant by definition.

It remains to prove the above claim, which we do now. We bound only $|I_u^+(\boldsymbol{y})|$ since the case of $|I_u^-(\boldsymbol{y})|$ is analogous. For all $i \in I_u^+(\boldsymbol{y})$, write

$$\boldsymbol{x}^{ui} = \frac{R(\boldsymbol{y}, \boldsymbol{x}^i)}{R(\boldsymbol{y}, \boldsymbol{y})} \boldsymbol{y} + \boldsymbol{\tau}^{ui}$$

for $\boldsymbol{\tau}^{ui} \perp \boldsymbol{y}$. Because $\mathbb{T}$ is a $(k, D, \vec{q}, \delta^2/2D^4)$-locally ultrametric tree, $|R(\boldsymbol{x}^{ui}, \boldsymbol{x}^{uj})| \le \delta^2/2D^4$ for all $i \neq j$. Thus, for all distinct $i, j \in I_u^+(\boldsymbol{y})$.

$$\frac{\delta^2}{2D^4} \ge R(\boldsymbol{x}^{ui}, \boldsymbol{x}^{uj}) = \frac{R(\boldsymbol{y}, \boldsymbol{x}^i) R(\boldsymbol{y}, \boldsymbol{x}^j)}{R(\boldsymbol{y}, \boldsymbol{y})} + R(\boldsymbol{\tau}^{ui}, \boldsymbol{\tau}^{uj}) \ge \frac{\delta^2}{D^4} + R(\boldsymbol{\tau}^{ui}, \boldsymbol{\tau}^{uj}),$$

where we use that $R(\boldsymbol{y}, \boldsymbol{y}) \le 1$. Thus $R(\boldsymbol{\tau}^{ui}, \boldsymbol{\tau}^{uj}) \le -\delta^2/2D^4$. However, $R(\boldsymbol{\tau}^{ui}, \boldsymbol{\tau}^{ui}) \le R(\boldsymbol{x}^{ui}, \boldsymbol{x}^{ui}) \le 1$. Thus $|I_u^+(\boldsymbol{y})| \le 2D^4/\delta^2$, as if not the Gram matrix of $(\boldsymbol{\tau}^{ui})_{i \in I_u^+(\boldsymbol{y})}$ would not be positive semi-definite. $\square$

Proof of Theorem 3.1.7. We will show $\mathbb{T}''$ satisfies the desired properties. First, for any $u, v \in \mathbb{T}''$, let $|u| = d_1$, $|v| = d_2$, and let $(\emptyset = u_0, u_1, u_2, \ldots, u_{d_1} = u)$, $(\emptyset = v_0, v_1, \ldots, v_{d_2} = v)$ be the ancestor paths of u, v. Also let $\ell = u \wedge v$, so $u_\ell = v_\ell$ is the least common ancestor of u, v. Then

$$R(\boldsymbol{\sigma}^u, \boldsymbol{\sigma}^v) = \sum_{i=0}^{d_1 - 1} \sum_{j=0}^{d_2 - 1} R(\boldsymbol{\sigma}^{u_{i+1}} - \boldsymbol{\sigma}^{u_i}, \boldsymbol{\sigma}^{v_{j+1}} - \boldsymbol{\sigma}^{v_j}). \tag{3.50}$$

The sub-sum corresponding to $0 \le i = j < \ell$ equals

$$\sum_{i=0}^{\ell-1} R(\boldsymbol{\sigma}^{u_{i+1}} - \boldsymbol{\sigma}^{u_i}, \boldsymbol{\sigma}^{u_{j+1}} - \boldsymbol{\sigma}^{u_j}) = \sum_{i=0}^{\ell-1} (q_{i+1} - q_i) = q_\ell,$$

while the remaining terms of (3.50) are bounded by δ/D^2 in absolute value by Proposition 3.4.9. Thus

$$|R(\boldsymbol{\sigma}^u, \boldsymbol{\sigma}^v) - q_{u \wedge v}| \le D^2 \cdot \delta/D^2 = \delta,$$

so $(\boldsymbol{\sigma}^u)_{u \in \mathbb{T}''}$ is a $(k'', D, \vec{q}, \delta)$-ultrametric tree. By property (i) of Theorem 3.4.4 and (3.49), for all $u \in \mathbb{T}''$ with $|u| = d$,

$$\frac{1}{N} H_N(\boldsymbol{\sigma}^u) = \frac{1}{N} H_N(\boldsymbol{\sigma}^\emptyset) + \sum_{i=0}^{d-1} \frac{1}{N} \left(H_N(\boldsymbol{\sigma}^{u_{i+1}}) - H_N(\boldsymbol{\sigma}^{u_i}) \right) \ge E(q_d) - (d+1)\varepsilon/(D+1) \ge E(q_d) - \varepsilon.$$

Thus property (i) of Theorem 3.1.7 holds. Property (ii) of Theorem 3.1.7 follows immediately from property (iii) of Theorem 3.4.4, as this property is monotone in k'. $\square$

Proof of Corollary 3.1.9. The proof is essentially identical to Theorem 3.1.7 but without the strictly RS part. Thus we just give an outline. Using [JT17, Theorem 1.13] and analyticity of ξ, it follows that Lemma 3.4.1 also holds for T. With the obvious definition, let $q_0, \ldots, q_D = 1$ be a T-refinement. Proposition 3.4.5 remains true with the same proof, except that $\xi_D = 0$ is now trivial. The remainder of the proof is as before. $\square$

3.5 Large deviations for the ground state

Here we make a brief study of large deviations for the ground state energy $GS_N = \max_{\boldsymbol{\sigma}\in\mathcal{S}_N} H_N(\boldsymbol{\sigma})$. [LACTFLD24] recently investigated this problem using the replica method, obtaining very interesting but non-rigorous results. It was predicted that for 1RSB models (without external field), the upper tail has rate function given by a natural Kac–Rice upper bound, referred to as "replica-symmetric" behavior therein. We verify this prediction in Subsection 3.5.1 by adapting the interpolation-enhanced truncation from Section 3.3. In Subsection 3.5.2 we employ Corollary 3.1.9 to show the lower tail speed transitions to $\Omega(N^2)$ below $\mathcal{Q}(\xi - \gamma_1^2 t)$ for general mixtures. As mentioned in Remark 3.5.9, the super-linearity in the lower tail is closely connected to the Dotsenko–Franz–Mézard conjecture [DFM94, Tal07, Jag17].

3.5.1 Upper tail for 1RSB models

We assume in this subsection that $\gamma_1 = 0$ and ξ is 1RSB, i.e. the minimizer (L, α) of (3.12) satisfies $\alpha \equiv u$. Unlike Section 3.3, we do not assume **strict** 1RSB, but parameters such as y, z, E_0, R_0 still retain the same definitions. We will also assume throughout this subsection that ξ is not pure. Similarly to Proposition 3.2.10, pure ξ can be handled by adding small perturbation terms to ξ, e.g. chosen small enough so the perturbation Hamiltonian has maximum absolute value at most δN with probability $1 - e^{N/\delta}$ (such perturbations have essentially no effect even in a large deviation sense).

The main computation is again encapsulated in controlling conditional band models as described in Lemma 3.3.13. Note that the only term in (3.31) that depends on (E, R) is

$$N \left\langle v^q, \begin{bmatrix} E - E_0 \\ R - R_0 \end{bmatrix} \right\rangle$$

for $v^q = (v_E^q, v_R^q)$ defined in (3.32).

Proposition 3.5.1. *For any ξ, ε with $\gamma_1 = 0$, there exists δ such that for all $q \in [\varepsilon, 1-\varepsilon]$,*

$$v_E^q \in [\delta, 1-\delta], \tag{3.51}$$

$$v_R^q \le -\delta. \tag{3.52}$$

Proof. We easily compute

$$v_E^q = \frac{\xi(q)\xi'(1) + \xi(q)\xi''(1) - q\xi'(q)\xi'(1)}{\xi(1)\xi'(1) + \xi(1)\xi''(1) - \xi'(1)^2}; \qquad v_R^q = \frac{q\xi'(q)\xi(1) - \xi(q)\xi'(1)}{\xi(1)\xi'(1) + \xi(1)\xi''(1) - \xi'(1)^2}.$$

(Recall from Remark 3.3.14 that the denominators are strictly positive as long as ξ is not pure.) Note that

$$\frac{\mathsf{d}}{\mathsf{d}q}\frac{q\xi'(q)}{\xi(q)} = \frac{\xi(q)(\xi'(q) + q\xi''(q)) - q\xi'(q)^2}{\xi(q)^2} > 0 \tag{3.53}$$

by Cauchy–Schwarz (as in Remark 3.3.14), and so $q \mapsto \frac{q\xi'(q)}{\xi(q)}$ is strictly increasing. Thus $\frac{q\xi'(q)\xi(1)}{\xi(q)} - \xi'(1)$ is negative and bounded away from 0 on $q \in [\varepsilon, 1-\varepsilon]$. Since $\xi(q)$ is also bounded away from 0 on this interval, this implies (3.52). Moreover, since $q \mapsto \frac{q\xi'(q)}{\xi(q)}$ is increasing,

$$\xi(q)\xi'(1) + \xi(q)\xi''(1) - q\xi'(q)\xi'(1) \ge \xi(q)\Big(\xi'(1) + \xi''(1) - \frac{\xi'(1)^2}{\xi(1)}\Big),$$

i.e. $v_E^q \geq \xi(q)/\xi(1) \geq \delta$. It now suffices to show v_E^q is strictly increasing in q. Differentiating and rearranging, it suffices to show

$$\frac{\xi''(1)}{\xi'(1)} \overset{?}{>} \frac{q\xi''(q)}{\xi'(q)}.$$

This holds because, by a calculation analogous to (3.53), $q \mapsto \frac{q\xi''(q)}{\xi'(q)}$ is increasing. □

Given ξ, we let $\mathcal{R} = \mathbb{R} \times [2\sqrt{\xi''(1)}, \infty)$. Critical points with $(E, R) \in \mathcal{R}$ will correspond to possible local maxima in Kac–Rice. Recalling (3.41), it is easy to see that Θ is strictly concave and continuously differentiable on $\mathcal{R}$. (Indeed the integral definition of κ immediately implies strict concavity outside the support of ρ.) For all $E \in \mathbb{R}$, define

$$R_*(E) = \underset{R \geq 2\sqrt{\xi''(1)}}{\arg\max}\, \Theta(E, R), \qquad \Theta_*(E) = \Theta(E, R_*(E)).$$

It is easy to see that both are finite, since the matrix Σ in (3.41) is positive definite.

Lemma 3.5.2. *For 1RSB ξ, we have $\frac{\partial}{\partial R}\Theta(E_0, R_0) = 0$.*

Proof. Recall (3.41) and (3.45). Since $R_0 = y + \frac{\xi''(1)}{y}$ and $y \geq \sqrt{\xi''(1)}$ by Lemma 3.3.7, we find:

$$\begin{aligned} \nabla\Theta(E_0, R_0) &= -\Sigma^{-1} \begin{bmatrix} E_0 \\ R_0 \end{bmatrix} + \begin{bmatrix} 0 \\ \xi''(1)^{-1/2}\kappa'(R_0/\sqrt{\xi''(1)}) \end{bmatrix} \\ &= -\Sigma^{-1} \begin{bmatrix} (\xi'(1) + z\xi(1))/y \\ y + (\xi''(1)/y) \end{bmatrix} + \begin{bmatrix} 0 \\ 1/y \end{bmatrix}. \end{aligned}$$

We would like to show the second entry in this vector vanishes, and (using Cramer's rule) it is given by

$$\frac{\xi'(1)^2 + z\xi(1)\xi'(1) - \xi(1)y^2 - \xi(1)\xi''(1)}{y\det(\Sigma)} + \frac{1}{y}.$$

Recalling from Lemma 3.3.6 that $y = \sqrt{(1+z)\xi'(1)}$, the conclusion follows. □

Lemma 3.5.3. *If ξ is 1RSB, then $R_*(E)$ is continuous and strictly increasing on $[E_0, \infty)$ with $R_*(E_0) = R_0$. Moreover $\Theta_*(E)$ is continuous and strictly decreasing with $\Theta_*(E_0) = 0$ and $\lim_{E\to\infty} \Theta_*(E) = -\infty$.*

Proof. Let

$$M(R) = \max_{E \in \mathbb{R}} \Theta(E, R).$$

This is easily seen to be C^1 on $[2\sqrt{\xi''}, \infty)$ and smooth on the interior, and inherits concavity from Θ. Since $\Theta(E, R)$ is a strictly concave quadratic function of (E, R) plus a strictly concave function of R, it can be written as

$$\Theta(E, R) = M(R) - K_1(E - K_2R)^2,$$

for K_1, K_2 depending only on ξ. Further, one easily finds that $-M(R) \asymp R^2$ for large R since κ grows sublinearly. Hence $M'(R)$ is a strictly decreasing, continuous function with $\lim_{R\to\infty} M'(R) = -\infty$.

$R_*(E)$ is the unique solution in $[2\sqrt{\xi''}, \infty)$ to

$$M'(R_*) = 2K_1K_2(K_2R_* - E), \tag{3.54}$$

assuming such a solution exists (if not, one would have the boundary solution $R_* = 2\sqrt{\xi''(1)}$.) Lemma 3.5.2 implies $R_0 = R_*(E_0)$. By monotonicity arguments (or inspecting a diagram), it follows that for all $E \geq E_0$, (3.54) admits a solution $R_*(E)$ which is continuous and strictly decreasing in E. It similarly follows that $\Theta_*(E)$ is continuous and strictly decreasing. Finally to show $\lim_{E\to\infty} \Theta_*(E) = -\infty$, note that since Σ is positive definite one has $\Theta(E, R) \leq -\varepsilon(|E| + |R|)^2 + \log(|R|)$ for some $\varepsilon = \varepsilon(\xi) > 0$. □

Similarly to Section 3.3, for $0 < \eta_4 \ll \eta_3 \ll \eta_2 \ll \eta_1 \ll 1$ small depending on some fixed $E > E_0$ let

$$\widetilde{B}(E) = [E - \eta_3, E + \eta_3] \times [R_*(E) - \eta_3, R_*(E) + \eta_3].$$

Recalling (3.30), we say $\boldsymbol{\sigma} \in \mathsf{Crt}(\widetilde{B}(E))$ is **large deviation typical** if H_N has no critical points $\boldsymbol{\rho}$ with $R(\boldsymbol{\sigma}, \boldsymbol{\rho}) \geq 1 - \eta_1$ and

$$\Psi(q;\boldsymbol{\sigma}) = \sup_{\boldsymbol{\rho} \in \mathsf{Band}_q(\boldsymbol{\sigma})} H_N(\boldsymbol{\rho}) \leq E - \eta_3, \quad \forall q \in [-\eta_4, 1 - \eta_1]. \tag{3.55}$$

Lemma 3.5.4. *For any $E > E_0$ and with small $0 < \eta_3 \ll \eta_2 \ll \eta_1 \ll 1$, given that $\boldsymbol{\sigma} \in \mathsf{Crt}(\widetilde{B}(E))$, $\boldsymbol{\sigma}$ is large deviation typical with conditional probability at least $1 - e^{-cN}$.*

Proof. Let $E_{\boldsymbol{\sigma}} = H_N(\boldsymbol{\sigma})/N$ and $R_{\boldsymbol{\sigma}} = \partial_{\mathsf{rad}} H_N(\boldsymbol{\sigma})/\sqrt{N}$. For η_3 small enough we must have $E_{\boldsymbol{\sigma}} - E_0 \geq (E - E_0)/2 > 0$ and $R_{\boldsymbol{\sigma}} - R_0 \geq (R - R_0)/2 > 0$ since $\boldsymbol{\sigma} \in \mathsf{Crt}(\widetilde{B}(E))$.

The former condition that H_N has no critical points $\boldsymbol{\rho}$ with $|R(\boldsymbol{\sigma}, \boldsymbol{\rho})| \geq 1 - \eta_1$ follows by Proposition 3.2.14 applied to the conditionally random part of the Hamiltonian $\widehat{H}_N$, exactly as in the proof of Proposition 3.3.12(ii). (Since $R_{\boldsymbol{\sigma}} > R_0 = y + \frac{\xi''}{y} \geq 2\sqrt{\xi''}$, which was also the case in that proof.)

Now we show (3.55). Combining Proposition 3.5.1 and Lemma 3.5.3, and using (η_1, η_2) for (ε, δ) in the former,

$$\left\langle v^q, \begin{bmatrix} E_{\boldsymbol{\sigma}} - E_0 \\ R_{\boldsymbol{\sigma}} - R_0 \end{bmatrix} \right\rangle \leq v_E^q (E_{\boldsymbol{\sigma}} - E_0) \leq (1 - \eta_2)(E_{\boldsymbol{\sigma}} - E_0), \quad \forall q \in [0, 1 - \eta_1]. \tag{3.56}$$

Next, note that if one replaces the $-2\eta_3$ term with $+\eta_3$, the proof of Proposition 3.3.15 goes through for (possibly non-strictly) 1RSB models and all $q \in [0, 1 - \eta_1]$. Moreover as usual it can be made simultaneous for all q by using Proposition 3.2.14 to union bound over a finite set of q. Thus conditional on $\boldsymbol{\sigma} \in \mathsf{Crt}(\widetilde{B}(E))$ and the values $(E_{\boldsymbol{\sigma}}, R_{\boldsymbol{\sigma}})$, with probability $1 - e^{-cN}$ we have for all $q \in [0, 1 - \eta_1]$ simultaneously:

$$\Psi(q;\boldsymbol{\sigma}) \leq E_0 + \left\langle v^q, \begin{bmatrix} E_{\boldsymbol{\sigma}} - E_0 \\ R_{\boldsymbol{\sigma}} - R_0 \end{bmatrix} \right\rangle + \eta_3 \overset{(3.56)}{\leq} E_0 + (1 - \eta_2)(E_{\boldsymbol{\sigma}} - E_0) + \eta_3 \leq E_{\boldsymbol{\sigma}} - 2\eta_3.$$

Finally Proposition 3.2.14 implies $\Psi(q;\boldsymbol{\sigma})$ is $C(\xi, E_{\boldsymbol{\sigma}}, R_{\boldsymbol{\sigma}})$-Lipschitz on $[-\eta_4, 0]$ with probability $1 - e^{-cN}$. Thus we find, as desired, that with conditional probability $1 - e^{-cN}$,

$$\Psi(q;\boldsymbol{\sigma}) \leq E_{\boldsymbol{\sigma}} - \eta_3, \quad \forall q \in [-\eta_4, 1 - \eta_1]. \qquad \square$$

Recalling (3.22), let $\mathsf{LMAX} \subseteq \mathsf{Crt}$ denote the set of local maxima of H_N.

Proposition 3.5.5. *For any ξ and $\varepsilon > 0$, there exists $c(\varepsilon) > 0$ such that for N large enough,*

$$\mathbb{E}\left[\left|\mathsf{LMAX} \cap \{\boldsymbol{\sigma} \ : \ \partial_{\mathsf{rad}} H_N(\boldsymbol{\sigma}) \leq 2\sqrt{\xi''(1)} - \varepsilon\}\right|\right] \leq e^{-c(\varepsilon)N^2}.$$

Proof. Similarly to Corollary 3.3.18, by combining Cauchy–Schwarz and the Kac–Rice formula it suffices to show that for all $R \leq 2\sqrt{\xi''(1)} - \varepsilon$,

$$\mathbb{P}\Big[\boldsymbol{\sigma} \in \mathsf{LMAX} \ \Big| \ \big(\partial_{\mathsf{rad}} H_N(\boldsymbol{\sigma}), \nabla_{\mathsf{sp}} H_N(\boldsymbol{\sigma})\big) = (R\sqrt{N}, \mathbf{0})\Big] \leq e^{-c'(\varepsilon)N^2}.$$

Recalling (3.21), this follows easily by the large deviation principle for the bulk spectrum of a GOE matrix, which has speed N^2 [BG97a]. $\square$

We are ready to determine the rate function for the upper tail of GS_N in 1RSB models.

Theorem 3.5.6. *Assume ξ is 1RSB. Then $\max(GS_N, E_0)$ obeys a large deviation principle on $[E_0, \infty)$ with speed N and good rate function $-\Theta_*(E)$.*

Proof. Since $\Theta_*(E)$ decreases continuously from $\Theta_*(E_0) = 0$ to $-\infty$, and exponential tightness is clear by e.g. Borell–TIS, it suffices[2] to show

$$\lim_{\eta \downarrow 0} \lim_{N\to\infty} \frac{1}{N} \log \mathbb{P}[GS_N \in [E-\eta, E+\eta]] \stackrel{?}{=} \Theta_*(E), \quad \forall E > E_0. \tag{3.57}$$

Thus, fix $E > E_0$ and let

$$\mathsf{Crt}_{typ}(\widetilde{B}(E)) \sqcup \mathsf{Crt}_{atyp}(\widetilde{B}(E)) = \mathsf{Crt}(\widetilde{B}(E))$$

respectively denote the large deviation typical and atypical critical points of H_N. For the large deviation upper bound, recall from Lemma 3.5.3 that on $[E,\infty) \times [2\sqrt{\xi''(1)}, \infty)$ the function Θ is maximized at $(E, R_*(E))$ with value $\Theta_*(E)$. We claim that the expected number of local maxima $\boldsymbol{\sigma} \in \mathsf{LMAX}$ satisfying $H_N(\boldsymbol{\sigma})/N \geq [E, \bar{E}]$ is at most $\exp(N\Theta_*(E) + o(N))$ for any $\bar{E} < \infty$ independent of N. Indeed, Proposition 3.5.5 shows that points $R \leq 2\sqrt{\xi''(1)} - \varepsilon$ contribute a negligible amount. Sending $\varepsilon \to 0$ slowly with N and applying Lemma 3.3.19 (and continuity of Θ) yields the claim. Since the global maximum of H_N is of course a local maximum, this together with exponential tightness of the ground state yields the upper bound.

For the lower bound, exactly as in Proposition 3.3.18 and its use in proving Proposition 3.3.9, we may deduce from Lemma 3.5.4 that

$$\mathbb{E}|\mathsf{Crt}_{atyp}(\widetilde{B}(E))| \leq e^{-cN/3}\mathbb{E}|\mathsf{Crt}(\widetilde{B}(E))|.$$

In particular

$$\mathbb{E}|\mathsf{Crt}_{typ}(\widetilde{B}(E))| \geq \mathbb{E}|\mathsf{Crt}(\widetilde{B}(E))|/2 \geq \exp(-N\Theta_*(E) \pm o(N))/2.$$

By definition, any two distinct large deviation typical points have overlap at most $-\eta_4$. Hence there are almost surely at most $2\eta_4^{-2}$ large deviation typical points in total, for any H_N (because their Gram matrix of overlaps must be positive semi-definite). Therefore

$$\begin{aligned}\mathbb{P}[|\mathsf{Crt}_{typ}(\widetilde{B}(E))| \geq 1] &\geq \eta_4^2\mathbb{E}|\mathsf{Crt}_{typ}(\widetilde{B}(E))|/2 \geq \eta_4^2 \exp(-N\Theta_*(E) \pm o(N))/4 \\ &\geq \exp(-N\Theta_*(E) \pm o(N)).\end{aligned}$$

By definition, if $|\mathsf{Crt}_{typ}(\widetilde{B}(E))| \geq 1$ then $GS_N \geq E - \eta_3$. Since η_3 was arbitrarily small, we obtain

$$\liminf_{\eta \downarrow 0} \lim_{N\to\infty} \frac{1}{N} \log \mathbb{P}[GS_N \geq E - \eta] \geq \Theta_*(E), \quad \forall E > E_0.$$

Since we already established the large deviation upper bound with strictly increasing rate function $-\Theta_*$ as well as exponential tightness, the previous display implies (3.57) as desired. □

3.5.2 Transition to quadratic speed

In this subsection, ξ is a general model, not necessarily 1RSB. We show in Theorem 3.5.8 that the large deviations of GS_N are of speed $O(N)$ above $\mathcal{Q}(\xi - \gamma_1^2 t)$ and $\Omega(N^2)$ below. The following exact orthogonalization lemma is crucial to show the latter result.

Lemma 3.5.7. *Fix small constants $c, \delta > 0$. Suppose $\boldsymbol{\sigma}^1, \dots, \boldsymbol{\sigma}^k \in \mathcal{S}_N$ for $k = e^{cN}$ satisfy $|R(\boldsymbol{\sigma}^i, \boldsymbol{\sigma}^j)| \leq \delta$ for all $1 \leq i < j \leq k$. Then for $c' > 0$ depending only on c, δ and for N large enough, there exists a subset $A \subseteq [k]$ of size $|A| \geq c'N$ and points $\{\widetilde{\boldsymbol{\sigma}}^a\}_{a\in A}$ such that:*

$$\begin{aligned} R(\widetilde{\boldsymbol{\sigma}}^a, \widetilde{\boldsymbol{\sigma}}^{a'}) &= 0, \quad \forall a \neq a' \in A, \\ \|\widetilde{\boldsymbol{\sigma}}^a - \boldsymbol{\sigma}^a\|_2 &\leq \delta^{0.01}\sqrt{N}. \end{aligned} \tag{3.58}$$

[2] Given exponential tightness, [AGZ10, Theorem D.4 and Corollary D.6] show (3.57) implies a large deviation principle on $[E_0 + \varepsilon, \infty)$ for any $\varepsilon > 0$. The large deviation principle easily extends to $[E_0, \infty)$ due to the aforementioned properties of Θ_*.

Proof. Let $A \subseteq [k]$ be any maximal subset such that there exist $\{\widetilde{\boldsymbol{\sigma}}^a\}_{a\in A}$ obeying (3.58), and assume for sake of contradiction that $|A| < c'N$. For $i \in [k]$, let $\widehat{\boldsymbol{\sigma}}_i$ be the projection of $\boldsymbol{\sigma}_i$ onto $\mathsf{span}\big(\{\widetilde{\boldsymbol{\sigma}}^a\}_{a\in A}\big)$. By maximality of A, we have $\|\widehat{\boldsymbol{\sigma}}_i\|_2 \geq \delta^{0.1}\sqrt{N}$ for all i. With $\mathcal{B}(A)$ the radius $\sqrt{N}$ ball in $\mathsf{span}\big(\{\widetilde{\boldsymbol{\sigma}}^a\}_{a\in A}\big)$, let $\hat{\mathcal{B}} = \mathcal{B}(A)\backslash\delta^{0.1}\mathcal{B}(A)$. Then, for a universal $C > 0$, $\hat{\mathcal{B}}(A)$ admits a covering by $\exp(C|A|\log(1/\delta))$ radius $\delta\sqrt{N}$ balls, whose centers are disjoint from $\delta^{0.1}\mathcal{B}(A)/2$. Since we assumed $|A| < c'N$, we have (for small enough c')

$$\exp(C|A|\log(1/\delta)) \leq \exp(cN/3).$$

Hence by the pigeonhole principle, there exists $J \subseteq [k]\backslash A$ with $|J| \geq \exp(cN/3)$ such that the points $\{\widehat{\boldsymbol{\sigma}}^j\}_{j\in J}$ in $\widehat{\mathcal{B}}$ are all contained in a radius $\delta\sqrt{N}$ ball centered at some $\widehat{\boldsymbol{\sigma}}^J \in \mathsf{span}\big(\{\widetilde{\boldsymbol{\sigma}}^a\}_{a\in A}\big)$ with $\|\widehat{\boldsymbol{\sigma}}^J\| \geq \delta^{0.1}\sqrt{N}/2$. Therefore

$$\frac{1}{|J|}\sum_{j\in J}\langle \boldsymbol{\sigma}^j, \widehat{\boldsymbol{\sigma}}^J\rangle = \frac{1}{|J|}\sum_{j\in J}\langle \widehat{\boldsymbol{\sigma}}^j, \widehat{\boldsymbol{\sigma}}^J\rangle \geq \frac{1}{|J|}\sum_{j\in J}\left(\langle \widehat{\boldsymbol{\sigma}}^J, \widehat{\boldsymbol{\sigma}}^J\rangle - \|\widehat{\boldsymbol{\sigma}}^j - \widehat{\boldsymbol{\sigma}}^J\|_2 \cdot \|\widehat{\boldsymbol{\sigma}}^J\|_2\right) \geq \delta^{0.3}N.$$

On the other hand, since we assumed $|R(\boldsymbol{\sigma}^i, \boldsymbol{\sigma}^j)| \leq \delta$, we have

$$\left\|\frac{1}{|J|}\sum_{j\in J}\boldsymbol{\sigma}^j\right\|_2^2 = \frac{1}{|J|^2}\sum_{j_1,j_2\in J}\langle \boldsymbol{\sigma}^{j_1}, \boldsymbol{\sigma}^{j_2}\rangle \leq \frac{1}{|J|^2}\left(|J| + |J|^2\delta N\right) \leq 2\delta N.$$

Since $\|\widehat{\boldsymbol{\sigma}}^J\|_2 \leq \sqrt{N}$, Cauchy–Schwarz gives the desired contradiction. □

Recall γ_1 is the weight of the degree-1 interactions in (3.1), and $\xi'(0) = \gamma_1^2$. For $h \geq 0$, let

$$\xi^{\gamma_1\leftarrow h}(t) = \xi(t) - \gamma_1^2 t + h^2 t$$

denote ξ with this interaction weight replaced by h. In particular $\xi^{\gamma_1\leftarrow 0}(t) = \xi(t) - \gamma_1^2 t$.

Theorem 3.5.8. *For any $E > \mathcal{Q}(\xi^{\gamma_1\leftarrow 0})$,*

$$\limsup_{\varepsilon\to 0}\limsup_{N\to\infty} -\frac{1}{N}\log\mathbb{P}[GS_N \in [E-\varepsilon, E+\varepsilon]] \leq C_1(\xi, E). \tag{3.59}$$

On the other hand, for any $\varepsilon > 0$,

$$\liminf_{N\to\infty} -\frac{1}{N^2}\log\mathbb{P}[GS_N \leq \mathcal{Q}(\xi^{\gamma_1\leftarrow 0}) - \varepsilon] \geq C_2(\xi, \varepsilon). \tag{3.60}$$

Proof. We first prove (3.59), assuming further that $\xi'(0) > 0$. By Theorem 3.2.2 and Proposition 3.2.14, $\mathcal{Q}(\xi^{\gamma_1\leftarrow h})$ is uniformly Lipschitz in h. Moreover clearly $\lim_{h\to\infty}\mathcal{Q}(\xi^{\gamma_1\leftarrow h}) = \infty$. Thus $E = \mathcal{Q}(\xi^{\gamma_1\leftarrow h})$ for some $h > 0$. Let $\boldsymbol{g} = (g_1, \ldots, g_N)$ be the vector of degree 1 disorder coefficients in (3.1). For small $\eta > 0$,

$$\frac{\gamma_1\|\boldsymbol{g}\|_2}{\sqrt{N}} \in [h-\eta, h+\eta]$$

occurs with probability at least $e^{-C_1(\xi,E)N}$. Conditional on this event, $\mathbb{P}[GS_N \in [E-\varepsilon, E+\varepsilon]] \geq 1/2$. This proves (3.59) if $\xi'(0) > 0$.

Next, suppose $\xi'(0) = 0$. Assuming $\gamma_p > 0$, we consider the conditional behavior of H_N on the large deviation event $E_{N,x}$ that the order p coefficient $g_{1,1,\ldots,1}$ (as in (3.1)) satisfies

$$g_{1,1,\ldots,1} = x\sqrt{N}.$$

Using Proposition 3.2.14 to discretize S_N into bands with fixed first coordinate, and applying the zero temperature Parisi formula to each band, it easily follows that, with $\widetilde{\xi}_q$ as in (3.25),

$$GS(x) \equiv \lim_{N\to\infty}\mathbb{E}[GS_N \mid E_{N,x}] = \sup_{0\leq q\leq 1}\left\{\mathcal{Q}(\widetilde{\xi}_q) + q^p x\right\}.$$

Moreover this limit holds locally uniformly in x. In particular $GS(x)$ is continuous and strictly increasing, and (since $\gamma_1 = 0$) $GS(0) = \mathcal{Q}(\xi)$. Hence for some $x_*(E), \delta > 0$ we have $GS(x) \in [E - \varepsilon/2, E + \varepsilon/2]$ for all $x \in [x_*(E) - \delta, x_*(E) + \delta]$. Then by Borell-TIS, for N sufficiently large, we conclude (3.59) from:

$$\mathbb{P}[GS_N \in E - \varepsilon, E + \varepsilon] \geq \mathbb{P}[N^{-1/2} g_{1,1,\dots,1} \in [x_*(E) - \delta, x_*(E) + \delta]]/2 \geq e^{-C(\xi,E)N}.$$

We turn to the proof of (3.60). We start from Corollary 3.1.9, applied to

$$H_{N,\geq 2}(\boldsymbol{\sigma}) = H_N(\boldsymbol{\sigma}) - \gamma_1 \langle \boldsymbol{g}, \boldsymbol{\sigma} \rangle,$$

which is a Hamiltonian with mixture $\xi^{\gamma_1 \leftarrow 0}$. This implies the high-probability existence of $\boldsymbol{\sigma}^1, \dots, \boldsymbol{\sigma}^k \in \mathcal{S}_N$ for $k = e^{cN}$ such that $|R(\boldsymbol{\sigma}^i, \boldsymbol{\sigma}^j)| \leq \delta \ll \varepsilon$ for all $1 \leq i < j \leq k$ and $H_{N,\geq 2}(\boldsymbol{\sigma}^i) \geq \mathcal{Q}(\xi^{\gamma_1 \leftarrow 0}) - \frac{\varepsilon}{4}$. On this event, Lemma 3.5.7 ensures the existence of $A \subseteq [k]$ with $|A| \geq c'N$ and points $\{\widetilde{\boldsymbol{\sigma}}^a\}_{a \in A}$ obeying (3.58). On the event of Proposition 3.2.14 we then have

$$\frac{1}{c'N^2} \max_{\vec{\boldsymbol{\rho}} \in \mathsf{Band}_{c'N,1,0}(\mathbf{0})} \sum_{i=1}^{c'N} H_{N,\geq 2}(\boldsymbol{\rho}^i) \geq \min_{\widetilde{\boldsymbol{\sigma}}^a} H_{N,\geq 2}(\widetilde{\boldsymbol{\sigma}}^a)/N \geq \mathcal{Q}(\xi^{\gamma_1 \leftarrow 0}) - \frac{\varepsilon}{2}.$$

Similarly to [Sub24, Proposition 1], the left-hand side above is sub-Gaussian with standard deviation proxy $O(N^{-2})$. With $c'' = c''(\xi, c, c', \delta) > 0$ a small constant, we find that with probability $1 - e^{-c''N^2}$,

$$\frac{1}{c'N^2} \max_{\vec{\boldsymbol{\rho}} \in \mathsf{Band}_{c'N,1,0}(\mathbf{0})} \sum_{i=1}^{c'N} H_{N,\geq 2}(\boldsymbol{\rho}^i) \geq \mathcal{Q}(\xi^{\gamma_1 \leftarrow 0}) - \frac{3\varepsilon}{4}.$$

Let $(\widehat{\boldsymbol{\rho}}^1, \dots, \widehat{\boldsymbol{\rho}}^{c'N})$ attain the maximum on the left-hand side (and depend measurably on $H_{N,\geq 2}$).

Finally we add back in the external field $\gamma_1 \langle \boldsymbol{g}, \boldsymbol{\sigma} \rangle$. Note that $\{\widetilde{\boldsymbol{\sigma}}^a\}_{a \in A}$ can be chosen (in some measurable way) depending only on $H_{N,\geq 2}$, so we may take $\boldsymbol{g}$ independent of $\{\widehat{\boldsymbol{\rho}}^i\}_{1 \leq i \leq c'N}$. Then $\sum_{i \leq c'N} \gamma_1 \langle \boldsymbol{g}, \widehat{\boldsymbol{\rho}}^i \rangle$ is conditionally a centered Gaussian with variance $c'N^2$, so it has absolute value smaller than $\varepsilon c'N^2/4$ with probability $1 - e^{-c''N^2}$. On this event and that of the preceding display, we find as desired:

$$\max_{\boldsymbol{\sigma} \in \mathcal{S}_N} H_N(\boldsymbol{\sigma})/N \geq \frac{1}{c'N^2} \max_{\vec{\boldsymbol{\rho}} \in \mathsf{Band}_{c'N,1,0}(\mathbf{0})} \sum_{i=1}^{c'N} H_N(\boldsymbol{\rho}^i) \geq \Big(\frac{1}{c'N^2} \sum_{i=1}^{c'N} H_{N,\geq 2}(\widehat{\boldsymbol{\rho}}^i) \Big) - \frac{\varepsilon}{4} \geq \mathcal{Q}(\xi^{\gamma_1 \leftarrow 0}) - \varepsilon. \qquad \square$$

Remark 3.5.9. Given a weaker version of Corollary 3.1.9 with $k_N \leq e^{o(N)}$, one finds $|A| \geq \Omega(\log k_N)$ in Lemma 3.5.7, which implies speed $\Omega(N \log k_N)$. In particular, Chatterjee's "multiple peaks" property [Cha14] suffices to obtain super-linear speed for the lower tail. At positive temperature, the super-linearity in the lower tail was essentially predicted by Dotsenko-Franz-Mézard in [DFM94] and proved in [Tal07, Jag17] using a similar "orthogonal structures" idea; see also [Che23a] which derives it from superconcentration. The main new feature of Theorem 3.5.8 is the quadratic rate, which is best possible when $\gamma_2 \neq 0$. It is natural to conjecture that $N^{p_{\min}}$ is the correct lower tail speed for $p_{\min} = \min\{p \,:\, \gamma_p > 0\}$.

Chapter 4

Strong topological trivialization of multi-species spherical spin glasses

Abstract – We study the landscapes of multi-species spherical spin glasses. Our results determine the phase boundary for annealed trivialization of the number of critical points, and establish its equivalence with a quenched *strong topological trivialization* property. Namely in the "trivial" regime, the number of critical points is constant, all are well-conditioned, and all *approximate* critical points are close to a true critical point. As a consequence, we deduce that Langevin dynamics at sufficiently low temperature has logarithmic mixing time.

Our approach begins with the Kac–Rice formula. We characterize the annealed trivialization phase by explicitly solving a suitable multi-dimensional variational problem, obtained by simplifying certain asymptotic determinant formulas from [BBM23, McK24]. To obtain more precise quenched results, we develop general purpose techniques to avoid sub-exponential correction factors and show non-existence of *approximate* critical points.Many of the results are new even in the 1-species case.

4.1 Introduction

This paper studies the landscapes of certain random, non-convex functions $H_N : \mathbb{R}^N \to \mathbb{R}$, namely the Hamiltonians of spherical spin glasses. Mean-field spin glass models were introduced in [SK75] to study disordered magnetic materials, and subsequently studied in many papers including [Rue87, CS92, CHS93]. Of particular note, Parisi predicted the free energy through the phenomenon of replica symmetry breaking in [Par79], which was later proved by [Tal06b, Tal06a] following decades of progress.

The spin glasses we focus on will feature $r \geq 1$ species, with the domain of H_N given by a product of r high-dimensional spheres. Such models include (in the Ising case) the $r = 2$ bipartite SK model [KC75, KS85, FKS87a, FKS87b] which has received recent attention due to connections with neural networks [BGG10, ABG+12, HPG18]. For $r > 1$, a basic understanding of the low-temperature statics is still missing in general: due to a breakdown of the crucial interpolation method [GT02], it is not even known that the limiting ground state energy exists in general despite much recent work [BCMT15, Pan15, Mou21, Mou23, BL20, BS22a, Sub21b, Sub23b, Kiv23].

We follow the landscape complexity approach pioneered by [Fyo04], studying the set of critical points using techniques such as the Kac–Rice formula. We focus in particular on *topological trivialization*: the transition of the number of critical points from exponential to constant beyond a critical external field strength. As further detailed below, precise understanding of this phenomenon faces several challenges. First, while one expects a dimension-independent number of critical points under a strong external field, the reduced symmetry from multiple species means tools for computing expected counts of critical points are accurate only to leading exponential order. Second, these leading order terms must be accessed implicitly through the solution to a vector Dyson equation. Third, even perfect knowledge of the critical points does not suffice to understand *approximate critical points* with *small* gradient (e.g. low temperature Gibbs samples) which might be far from any genuine critical point. Fourth, the Kac–Rice formula only gives annealed expectations, so the phase boundary it suggests might not correspond to any quenched property. These challenges will lead us to develop new techniques which enhance the Kac–Rice formula and yield a more complete description of the landscape even in the one-species setting.

Landscape complexity Our starting point is the Kac–Rice formula introduced in [Ric44, Kac48] (see [AT09, Chapter 11] for a textbook treatment). In general this formula allows one to compute moments for the number of critical points, local optima, and similar quantities for smooth Gaussian processes on manifolds. It has been employed, usually to obtain annealed counts of critical points, in many settings including spiked tensor models [BMMN19, RBBC19, FMM21, CFM23, ABL22], non-gradient vector fields [CKLDP97, FK16, Fyo16, Gar17, BFK21, Kiv24, Sub23a], polymer models [FLDRT18], Euler characteristics [TA03], generalized linear models [MBB20], and the elastic manifold [BBM24]. Typically the most complicated term in the Kac–Rice integrand is the expected determinant of a large random matrix.

For spherical spin glasses, the important works [ABČ13, AB13] calculated the annealed exponential growth rates for the number of critical points of various indices and energy levels. Matching second moment estimates for pure models were established in [Sub17a], see also [AG20, SZ21]. These yielded in some cases an elementary proof of the Parisi formula at zero temperature, as well as further geometric results on the Gibbs measures [SZ17, Sub17b, BSZ20, BJ24]. Annealed asymptotics for the multi-species setting were obtained in [McK24], with a matching second moment computation in the pure case by [Kiv23].

Our primary aim will be to identify the topologically trivial phase where the landscape contains a (dimension-free) constant number of critical points, and to understand it in detail. This was done for the *annealed* complexity of single-species spin glasses in [FLD14, Fyo15, BČNS22], which showed that in the trivial regime the only critical points are the unique global maximum and minimum (with high probability). In these works and many others mentioned above, the high degree of symmetry is crucial: it ensures the random matrices appearing in Kac–Rice computations are from the Gaussian Orthogonal Ensemble, for which exact determinantal formulas are available. Recently the work [BBM23] gave broadly applicable tools for random matrix determinants with less symmetry. Their work enables quite general Kac–Rice computations, with the caveat that the results hold only to leading exponential order (i.e. with an extra $e^{o(N)}$ factor in dimension N).

Through the example of multi-species spin glasses, we aimed to study the following three meta-questions on random landscapes which do not seem to have been addressed in the literature. While the first is somewhat tailored to the multi-species setting and the recent work [BBM23], we are unaware of rigorous results toward the latter two in any of the models above.

(1) For non-symmetric models with Hessians more complicated than GOE, does the topologically trivial regime still exhibit a dimension-free number of critical points? Or is the $e^{o(N)}$ upper bound on annealed complexity the end of the story?

(2) With or without symmetry, does the phase boundary of annealed topological trivialization have genuine significance? Or can the regime of *quenched* topological trivialization be strictly larger?

(3) Does topological trivialization imply rapid convergence for optimization algorithms such as Langevin dynamics? Or might regions with small but non-zero gradient lead to arbitrarily slow convergence?

The last question in particular was highlighted in the recent book chapter [RF23], which ends:

> *Finally, we find it appropriate to conclude this chapter by recalling that getting a refined information on the landscape topology and geometry can hopefully shed light and guide us into the comprehension of the dynamical evolution of the complex systems associated to it: establishing quantitatively this connection between landscape and dynamics is the underlying goal of the landscape program, and thus the most relevant perspective.*

Our results We make progress on all three of the above questions. As the first step, we establish in Theorem 4.1.11 the annealed phase boundary for topological trivialization in the sense of leading exponential order. Already these annealed estimates determine the ground state energy in the topologically trivial phase. Answering Question (1), we go further and show throughout the topologically trivial regime that the number of critical points in an r-species spin glass is exactly 2^r, which is the minimum possible for any Morse function on a product of r spheres. Moreover the landscape trivializes in a quantitative sense: each of the 2^r critical points has dimension-free condition number, and all *approximate* critical points (with small gradient) are close to one of them. We call this confluence of properties **strong topological trivialization** (see Definition 4.1.12), and show that for any landscape satisfying it, the mixing time for low temperature Langevin dynamics is $O(\log N)$. This addresses Question (3) above. Conversely in the topologically non-trivial phase, our companion paper [HS24] explicitly constructs exponentially many well-separated approximate critical points (see Proposition 4.1.14 below). This implies *quenched* failure of *strong* topological trivialization whenever the annealed complexity is non-trivial, partially addressing Question (2).

Proof techniques Our computations using the Kac–Rice formula rely on asymptotics for expectations of random determinants computed in [BBM23] via the vector Dyson equation, in particular Corollary 1.9.A therein. We determine the expected number of critical points to leading exponential order, in particular identifying the trivial regime of annealed complexity $e^{o(N)}$.

First we discuss several new challenges arising in the Kac–Rice computations as compared to the single-species setting. Whereas the random determinants arising for one species are of a GOE matrix plus a scalar multiple of the identity, with multiple species the GOE is replaced by a more general Gaussian block matrix. Before the present work, the exponential-order growth rates of the relevant determinants were only known in the form of an integral $\Psi = \int \log|\gamma| \mathsf{d}\mu(\gamma)$ for a measure μ whose Stieltjes transform solves a vector Dyson equation (see (4.44)). While this integral can be explicitly evaluated in the single-species case because μ is a shift of the semicircle law, in general this representation is far from explicit. Addressing this challenge, we find a closed-form formula for this integral (Lemma 4.4.4), expressed in terms of the solution to the vector Dyson equation. This formula makes the Kac–Rice calculations reasonably explicit and may be of independent interest.

The annealed complexity is now given by the maximum of an r-dimensional complexity functional $F : \mathbb{R}^r \to \mathbb{R}$ whose main term is this Ψ. Determining the maximum of F is also complicated and requires more than finding its stationary points. Even for one species, F has both concave and convex regions and is C^1 but not C^2. With $r > 1$ species, the maximization of F becomes a multi-dimensional optimization problem where the number of concave and nonconcave regions in $\mathbb{R}^r$ grows exponentially with r; see Subsection 4.4.3 and Figure 4.4.1. In general, F has about 3^r stationary points (see Lemma 4.4.5), 2^r of which eventually yield critical points of H_N. To show that the remaining stationary points do not maximize F, we construct at each of these points an explicit direction along which the Hessian $\nabla^2 F$ is positive. The construction of this direction requires an understanding of the *solution* to the vector Dyson equation, which ranges over a non-trivial subset of $\mathbb{C}^r$ that we characterize (Lemma 4.4.8). To justify the necessary calculations, we also prove (in Theorem 4.A.2) new joint continuity properties of the vector Dyson equation; these extend results of [AEK17a, AEK20] in the case of finitely many blocks.

Having determined the trivial regime for annealed complexity, we next turn a more precise understanding of this regime, for example aiming to show the number of critical points is exactly 2^r with high probability. It follows from the annealed estimates that all critical points are well-conditioned and must have one of 2^r "types" corresponding to maxima of F (see Definition 4.6.1). Here the type of a critical point essentially determines its overlap with the external field (which is an r-dimensional vector), as well as its energy and Hessian spectrum. This motivates the following natural strategy. For each of the 2^r types we restrict attention to a lower-dimensional band having the correct overlap with the external field (thus containing all critical points of that type), and search for critical points of the restriction of H_N to this band by repeating this process. The conditional law of H_N on such a band is again a spherical multi-species spin glass in the topologically trivial regime, and the relevant band shrinks in diameter each step. Thus we would hope to eventually show that all critical points of each fixed type are close together. Since the Kac–Rice estimates imply all critical points are well-conditioned, there can only be at most 1 inside any small region. Hence this would imply a 2^r upper bound for the number of critical points.

Unfortunately this approach does not make sense on its face because critical points are *brittle*. In particular the set of critical points of H_N restricted to a lower dimensional band might be unrelated to the set of critical points on an *open neighborhood* of said band. To overcome this difficulty, we establish in Theorem 4.5.2 a way to pass from annealed upper bounds for exact critical points to high-probablity non-existence of *approximate* critical points (with small gradient). Because the notion of approximate critical point is more robust, the shrinking bands argument above can be salvaged, thus proving the 2^r upper bound. The fact that all 2^r critical points actually exist then follows by the Morse inequalities from differential topology. The aforementioned non-existence of approximate critical points far from any exact critical point also falls out of the shrinking bands argument.

We note that the ability to control approximate critical points via Kac–Rice estimates seems quite powerful and should have further applications to random landscapes. In addition to the shrinking bands argument above, we also needed Theorem 4.5.2 to show strict positivity of the annealed complexity in the complementary "non-trivial" regime, see Subsection 4.5.3. Further, in Section 4.7 we use these ideas to derive energy estimates for "approximate local maxima" in single-species spherical spin glasses without external field, via the annealed thresholds $E_\infty^\pm$ of [AB13]. We have already applied these estimates to obtain bounds on the energy attained by low temperature Langevin dynamics [Sel24b] and the algorithmic threshold energy for Lipschitz optimization [HS23a, HS24].

Connections to algorithms As mentioned previously, our landscape results yield non-asymptotic algorithmic consequences in the "trivial" regime. We show in Theorem 4.1.16 that low temperature Langevin dynamics rapidly enters a small neighborhood of the global maximum and remains there for an exponentially long time, even from disorder-dependent initialization. The proof relies on recent work by one of us [Sel24b] to ensure the dynamics does not get stuck in saddle points. Thanks to the local concavity of H_N around its global maximum, we also deduce in this theorem that low temperature Langevin dynamics undergoes total variation mixing within $O(\log N)$ time. These results follow in a black-box way from the strong topological trivialization property discussed above.

Being an optimization algorithm, Langevin dynamics can find only the global maximum or minimum. However an equally natural "critical point following" algorithm explained in Subsection 4.6.6 suffices to find all 2^r critical points. Here one first locates the critical point of the desired "type" under an amplified external field, and then follows its movement as the external field strength is gradually decreased. Well-conditioning of critical points ensures that this movement is stable and easy to follow for any model in the "trivial" regime. This can be seen as a variant of "state following" [BFP97, ZK10, SCK$^+$12]; see also [BSZ20, Proposition 9.1] and [SFL19] for similar ideas.

Finally, the phase boundary for topological trivialization coincides with a transition in the structure of algorithmically reachable states we recently identified in [HS23a, HS24]. These works study the optimization of H_N using algorithms, viewed as functions of the disorder coefficients, with Lipschitz constant independent of N. Roughly speaking, it is shown that the reachable points for the best such optimization algorithms have the structure of a continuously branching ultrametric tree, and both approximate message passing and a second-order ascent algorithm generalizing that of [Sub21a] (and using a correlated ensemble of Hamiltonians) find these points. The algorithmic tree is rooted at a random location correlated with the external field (which is simply the origin when the external field vanishes), and branches orthogonally outward until reaching the boundary of the state space. When the external field is large enough, the algorithmic tree degenerates; the root moves all the way to the boundary of the state space and no branching occurs. We show in this paper that said degeneracy coincides with topological trivialization. In the "nontrivial" regime, the algorithmic tree is non-degenerate and [HS24] uses the above algorithms to construct e^{cN} well-separated approximate critical points, yielding the quenched non-trivialization discussed previously. Conversely in the "trivial" regime, [HS24] gives a signed generalization of the root-finding approximate message passing iteration which locates all 2^r critical points by implementing the previously mentioned recursive-bands argument as an algorithm.

4.1.1 Model description

Fix a finite set $\mathscr{S} = \{1, \dots, r\}$ and weights $\vec{\lambda} = (\lambda_1, \dots, \lambda_r) \in \mathbb{R}_{>0}^{\mathscr{S}}$ with $\sum_{s\in\mathscr{S}} \lambda_s = 1$. For each positive integer N, fix a deterministic partition $\{1, \dots, N\} = \bigsqcup_{s\in\mathscr{S}} \mathcal{I}_s$ with $N_s/N = \lambda_{N,s}$ and $\lim_{N\to\infty} \lambda_{N,s} = \lambda_s$ for $N_s = |\mathcal{I}_s|$. For $s \in \mathscr{S}$ and $\boldsymbol{x} \in \mathbb{R}^N$, let $\boldsymbol{x}_s \in \mathbb{R}^{\mathcal{I}_s}$ denote the restriction of $\boldsymbol{x}$ to coordinates $\mathcal{I}_s$. We consider the product-of-spheres state space

$$\mathcal{S}_N = \left\{ \boldsymbol{x} \in \mathbb{R}^N : \|\boldsymbol{x}_s\|_2^2 = \lambda_s N \quad \forall\, s \in \mathscr{S} \right\}. \tag{4.1}$$

For each $k \geq 1$ fix a symmetric tensor

$$\Gamma^{(k)} = (\gamma_{s_1,\ldots,s_k})_{s_1,\ldots,s_k \in \mathscr{S}} \in (\mathbb{R}^r_{\geq 0})^{\otimes k}$$

with $\sum_{k\geq 1} 2^k \|\Gamma^{(k)}\|_\infty < \infty$, and let $\boldsymbol{G}^{(k)} \in (\mathbb{R}^N)^{\otimes k}$ be a tensor with i.i.d. standard Gaussian entries. For $A \in (\mathbb{R}^{\mathscr{S}})^{\otimes k}$, $B \in (\mathbb{R}^N)^{\otimes k}$, define $A \diamond B \in (\mathbb{R}^N)^{\otimes k}$ to be the tensor with entries

$$(A \diamond B)_{i_1,\ldots,i_k} = A_{s(i_1),\ldots,s(i_k)} B_{i_1,\ldots,i_k}, \tag{4.2}$$

where $s(i)$ denotes the $s \in \mathscr{S}$ such that $i \in \mathcal{I}_s$. We consider the mean-field multi-species spin glass Hamiltonian

$$\begin{aligned} H_N(\boldsymbol{\sigma}) &= \sum_{k\geq 1} \frac{1}{N^{(k-1)/2}} \langle \Gamma^{(k)} \diamond \boldsymbol{G}^{(k)}, \boldsymbol{\sigma}^{\otimes k} \rangle \\ &= \sum_{k\geq 1} \frac{1}{N^{(k-1)/2}} \sum_{i_1,\ldots,i_k=1}^{N} \gamma_{s(i_1),\ldots,s(i_k)} \boldsymbol{G}^{(k)}_{i_1,\ldots,i_k} \sigma_{i_1} \cdots \sigma_{i_k} \end{aligned} \tag{4.3}$$

with inputs $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_N) \in \mathcal{S}_N$. For $\boldsymbol{\sigma}, \boldsymbol{\rho} \in \mathcal{S}_N$, define the species s overlap and overlap vector

$$R_s(\boldsymbol{\sigma}, \boldsymbol{\rho}) = \frac{\langle \boldsymbol{\sigma}_s, \boldsymbol{\rho}_s \rangle}{\lambda_s N}, \qquad \vec{R}(\boldsymbol{\sigma}, \boldsymbol{\rho}) = \big(R_1(\boldsymbol{\sigma}, \boldsymbol{\rho}), \ldots, R_r(\boldsymbol{\sigma}, \boldsymbol{\rho})\big). \tag{4.4}$$

Let $\odot$ denote coordinate-wise product. For $\vec{x} = (x_1, \ldots, x_r) \in \mathbb{R}^{\mathscr{S}}$, let

$$\begin{aligned} \xi(\vec{x}) &= \sum_{k\geq 1} \langle \Gamma^{(k)} \odot \Gamma^{(k)}, (\vec{\lambda} \odot \vec{x})^{\otimes k} \rangle \\ &= \sum_{k\geq 1} \sum_{s_1\ldots,s_k \in \mathscr{S}} \gamma^2_{s_1,\ldots,s_k} (\lambda_{s_1} x_{s_1}) \cdots (\lambda_{s_k} x_{s_k}). \end{aligned}$$

The random function H_N can also be described as the Gaussian process on $\mathbb{R}^N$ with covariance

$$\mathbb{E} H_N(\boldsymbol{\sigma}) H_N(\boldsymbol{\rho}) = N \xi(\vec{R}(\boldsymbol{\sigma}, \boldsymbol{\rho})).$$

It will be useful to define, for $s \in \mathscr{S}$,

$$\xi^s(\vec{x}) = \lambda_s^{-1} \partial_{x_s} \xi(\vec{x}), \tag{4.5}$$

$$\xi' = \nabla \xi(\vec{1}) \in \mathbb{R}^r, \tag{4.6}$$

$$\xi'' = \nabla^2 \xi(\vec{1}) \in \mathbb{R}^{r\times r}. \tag{4.7}$$

We will often write $\mathrm{diag}(\xi')$ for the $r \times r$ matrix with (s,s) entry ξ'_s, and similarly for other vectors. Finally, most of our results require the following generic non-degeneracy condition for ξ.

Assumption 4.1.1. ξ is **non-degenerate** if $\Gamma^{(1)}, \Gamma^{(2)}, \Gamma^{(3)} > 0$ holds entry-wise. For fixed $\vec{\lambda}$, a family of mixture functions ξ is **uniformly non-degenerate** if the sums $\sum_{k\geq 1} 2^k \|\Gamma^{(k)}\|_\infty$ are uniformly bounded above, and for some $\varepsilon > 0$, we have $\Gamma^{(1)}, \Gamma^{(2)}, \Gamma^{(3)} \geq \varepsilon$ entry-wise for all ξ in the family.

4.1.2 Basic notations and conventions

Here we detail some notations that will be useful to understand the statements in the next subsection.

Definition 4.1.2. For probability measures μ, ν on a metric space $(\mathcal{X}, d)$, and $p \in [1, \infty]$, we denote by $\mathbb{W}_p(\mu, \nu)$ the Wasserstein distance

$$\mathbb{W}_p(\mu, \nu) = \left(\inf_{\Pi \in \mathcal{C}(\mu,\nu)} \mathbb{E}_\Pi \Big[d(\boldsymbol{x}, \boldsymbol{y})^p \Big] \right)^{1/p},$$

where the infimum is over all couplings $(\boldsymbol{x}, \boldsymbol{y}) \sim \Pi$ with marginals $\boldsymbol{x} \sim \mu$ and $\boldsymbol{y} \sim \nu$. (For $p = \infty$, the distance is the essential supremum of $d(\boldsymbol{x}, \boldsymbol{y})$ under the coupling.) Unless otherwise specified, $(\mathcal{X}, d)$ will always be $\mathbb{R}^n$ for some $n \geq 1$ with the standard Euclidean metric.

Definition 4.1.3. The Hausdorff distance between sets $S_1, S_2 \subseteq \mathbb{R}$ is given by

$$d_{\mathcal{H}}(S_1, S_2) \equiv \max\left(\max_{s_1\in S_1} d(s_1, S_2), \max_{s_2\in S_2} d(S_1, s_2)\right). \tag{4.8}$$

Here $d(s_1, S_2) = \inf_{s_2\in S_2} d(s_1, s_2)$ is the usual point-to-set distance.

Given a symmetric $n \times n$ matrix M, we denote by $\boldsymbol{\lambda}_{\min}(M), \boldsymbol{\lambda}_{\max}(M)$ its minimum and maximum eigenvalue, and by $\boldsymbol{\lambda}_k(M)$ its k-th largest eigenvalue. Using $\mathcal{P}(\mathbb{R})$ to denote the space of probability measures on $\mathbb{R}$, denote by

$$\mathsf{spec}(M) = \{\boldsymbol{\lambda}_k(M) \,:\, k \in [n]\} \subseteq \mathbb{R}, \qquad \widehat{\mu}(M) = \frac{1}{n}\sum_{k=1}^{n} \delta_{\boldsymbol{\lambda}_k(M)} \in \mathcal{P}(\mathbb{R}), \tag{4.9}$$

the empirical spectral support and measure of M. Let

$$\mathop{\mathsf{spec}}_{H_N}(\boldsymbol{x}) = \mathsf{spec}\left(\nabla^2_{\mathsf{sp}} H_N(\boldsymbol{x})\right), \qquad \widehat{\mu}_{H_N}(\boldsymbol{x}) = \widehat{\mu}\big(\nabla^2_{\mathsf{sp}} H_N(\boldsymbol{x})\big) \tag{4.10}$$

be the corresponding objects for the Riemannian Hessian defined just below. We will always use non-bolded $\vec{\lambda} = (\lambda_s)_{s\in\mathscr{S}}$ to denote the species weights as in Subsection 4.1.1.

Next we define the *radial derivative* and *Riemannian gradient and Hessian H_N*. Throughout the paper we assume $\mathcal{I}_1 = \{1, \ldots, m_1\}$, $\mathcal{I}_2 = \{m_1+1, \ldots, m_2\}$, and so on. Let $\mathcal{R} = \{m_1, \ldots, m_r\}$ and $\mathcal{T} = [N]\setminus\mathcal{R}$. For each $\boldsymbol{\sigma} \in \mathcal{S}_N$, we pick an orthonormal basis $\big\{e_1(\boldsymbol{\sigma}), \ldots, e_N(\boldsymbol{\sigma})\big\}$ of $\mathbb{R}^N$ so that $\{e_i(\boldsymbol{\sigma}) : i \in \mathcal{I}_s\}$ constitutes an orthonormal basis of $\mathbb{R}^{\mathcal{I}_s}$, and $\boldsymbol{\sigma}_s = \sqrt{\lambda_s N} e_{m_s}(\boldsymbol{\sigma})$. For $S \subseteq [N]$, let $\nabla_S H_N(\boldsymbol{\sigma}) \in \mathbb{R}^S$ denote the restriction of $\nabla H_N(\boldsymbol{\sigma}) \in \mathbb{R}^N$ to the coordinates in S (in the orthonormal basis $\{e_1(\boldsymbol{\sigma}), \ldots, e_N(\boldsymbol{\sigma})\}$), and define $\nabla^2_{S\times S} H_N \in \mathbb{R}^{S\times S}$ analogously. The radial derivative is $\nabla_{\mathcal{R}} H_N(\boldsymbol{\sigma})$; it will be convenient to define a rescaled radial derivative $\nabla_{\mathsf{rad}} H_N(\boldsymbol{\sigma}) = N^{-1/2}\nabla_{\mathcal{R}} H_N(\boldsymbol{\sigma})$ so that the below formulas become dimension-free. Then, define the matrices

$$\Lambda = \mathrm{diag}(\vec{\lambda}) \in \mathbb{R}^{r\times r}, \qquad A = \mathrm{diag}(\xi') + \xi'' \in \mathbb{R}^{r\times r}. \tag{4.11}$$

The following standard fact relates the Riemannian gradient and Hessian of H_N to the Euclidean gradient and Hessian and can be taken as a definition.

Fact 4.1.4. *Let $\nabla_{\mathsf{sp}} H_N(\boldsymbol{\sigma})$, $\nabla^2_{\mathsf{sp}} H_N(\boldsymbol{\sigma})$ denote the Riemannian gradient and Hessian of H_N in $\mathcal{S}_N$. Then,*

$$\nabla_{\mathsf{sp}} H_N(\boldsymbol{\sigma}) = \nabla_{\mathcal{T}} H_N(\boldsymbol{\sigma}), \qquad \nabla^2_{\mathsf{sp}} H_N(\boldsymbol{\sigma}) = \nabla^2_{\mathcal{T}\times\mathcal{T}} H_N(\boldsymbol{\sigma}) - \mathrm{diag}(\Lambda^{-1/2}\nabla_{\mathsf{rad}} H_N(\boldsymbol{\sigma}) \diamond \mathbf{1}_{\mathcal{T}}).$$

Explicitly, the curvature term $\mathrm{diag}(\Lambda^{-1/2}\nabla_{\mathsf{rad}} H_N(\boldsymbol{\sigma}) \diamond \mathbf{1}_{\mathcal{T}})$ is a diagonal matrix $D \in \mathbb{R}^{\mathcal{T}\times\mathcal{T}}$ where for all $i \in \mathcal{T} \cap \mathcal{I}_s$,

$$D_{i,i} = \frac{1}{\sqrt{\lambda_s}}(\nabla_{\mathsf{rad}} H_N(\boldsymbol{\sigma}))_s = \frac{1}{\sqrt{\lambda_s N}}\partial_{m_s} H_N(\boldsymbol{\sigma}).$$

We now define approximate critical points and ground states. We remark that all ε-approximate ground states are δ-approximate critical points for some $\delta(\varepsilon)$ tending to 0 with ε, assuming H_N lies in the exponentially high probability set K_N defined in Proposition 4.2.4.

Definition 4.1.5. A point $\boldsymbol{x} \in \mathcal{S}_N$ is an **ε-approximate critical point** if $\|\nabla_{\mathsf{sp}} H_N(\boldsymbol{x})\|_2 \le \varepsilon\sqrt{N}$, and an **$\varepsilon$-approximate ground state** if $H_N(\boldsymbol{x}) + \varepsilon N \ge \max_{\boldsymbol{\sigma}\in\mathcal{S}_N} H_N(\boldsymbol{\sigma})$. We will sometimes abbreviate these by **ε-critical point** and **ε-ground state**. Moreover, $\boldsymbol{x}$ is **C-well conditioned** if $\nabla^2_{\mathsf{sp}} H_N(\boldsymbol{x})$ has all eigenvalues in $\pm[C^{-1}, C]$. Finally, $\boldsymbol{x}$ is an **(ε, C)-well conditioned critical point** if it is both an ε-critical point and C-well conditioned.

Finally, we will often say that an event holds with probability $1 - e^{-cN}$. In these cases, unless specified otherwise, c is a small constant which may depend on all other relevant N-independent constants.

4.1.3 Main results on strong topological trivialization

In this subsection we state our main results. Theorem 4.1.11 below shows that super-solvable models have $e^{o(N)}$ critical points and identifies their possible asymptotic energies, correlation with $\boldsymbol{G}^{(1)}$, and Hessian spectrum. Theorem 4.1.13 considerably refines this statement for strictly super-solvable models, establishing *strong* topological trivialization: the number of critical points equals 2^r, all are well-conditioned, and all approximate critical points are close to one of them.

Definition 4.1.6. Assume ξ is non-degenerate. We say ξ is **super-solvable** if $\operatorname{diag}(\xi') \succeq \xi''$, where $\succeq$ denotes the Loewner (positive semi-definite) partial order, **strictly super-solvable** if $\operatorname{diag}(\xi') \succ \xi''$. Similarly we say ξ is **strictly sub-solvable** if $\operatorname{diag}(\xi') \not\succeq \xi''$. We say ξ is **solvable** if $\operatorname{diag}(\xi') - \xi'' \succeq 0$ is singular.

Remark 4.1.7. The condition $\operatorname{diag}(\xi') \succeq \xi''$ coincides with the condition for degeneracy of the algorithmic ultrametric trees identified in [HS23a, Theorem 3]. In one species, this condition recovers the annealed trivialization condition determined by [BČNS22, Theorem 1.1] and coincides with the condition for zero-temperature replica symmetry [CS17, Proposition 1].

Remark 4.1.8. The super-solvability condition above also appeared naturally in [HS23a] via the analysis of an ODE describing the algorithmic threshold for optimizing multi-species spherical spin glass Hamiltonians. In that paper, super-solvability is a property of points in $\mathbb{R}^r_{\geq 0}$ for fixed ξ, and this ODE has different behaviors on super-solvable and sub-solvable regions of $\mathbb{R}^r_{\geq 0}$. In the present work, ξ is super-solvable if, in the language of [HS23a], $\vec{1}$ is super-solvable for ξ.

Remark 4.1.9. We note that in both [HS23a, HS24], the linear external field term in H_N is a deterministic vector $\boldsymbol{h}$ rather than the random $\boldsymbol{G}^{(1)}$; both papers immediately apply to the model (4.3) by conditioning on $\boldsymbol{G}^{(1)}$. On the other hand, the upper bounds we obtain on annealed complexity become slightly stronger when the external field is random.

The 2^r critical points of a strictly super-solvable model will correspond naturally to sign patterns $\vec{\Delta} \in \{-1,1\}^r$, which determine whether the critical point is positively or negatively correlated with the external field in each species. For each $\vec{\Delta}$, we define its associated energy, overlap, and radial derivative:

$$\begin{aligned} E(\vec{\Delta}) &= \sum_{s\in\mathscr{S}} \Delta_s \sqrt{\lambda_s \xi'_s}\,, \\ \vec{R}(\vec{\Delta}) &= \left(\frac{\Delta_s \gamma_s}{\sqrt{\xi'_s}}\right)_{s\in\mathscr{S}}, \\ \vec{x}(\vec{\Delta}) &= \left(\Delta_s\sqrt{\xi'_s} + \sum_{s'\in\mathscr{S}} \Delta_{s'}\sqrt{\frac{\lambda_{s'}}{\lambda_s}}\cdot\frac{\xi''_{s,s'}}{\sqrt{\xi'_{s'}}}\right)_{s\in\mathscr{S}}. \end{aligned} \tag{4.12}$$

Below we refer to certain probability measures $\mu(\vec{x}(\vec{\Delta})) \in \mathcal{P}(\mathbb{R})$, which are the limiting spectral measures for certain random block matrices, defined using the vector Dyson equation in (4.39). We also refer to their supports $S(\vec{\Delta}) = \mathsf{supp}\left(\mu(\vec{x}(\vec{\Delta}))\right) \subseteq \mathbb{R}$ which are finite unions of intervals (see (4.43)). We will use Definitions 4.1.2 and 4.1.3 as well as ∇^2_{sp} and ∇_{rad} as defined in Fact 4.1.4.

Definition 4.1.10. Let $\varepsilon > 0$ and $\vec{\Delta} \in \{-1,1\}^r$. A point $\boldsymbol{x} \in \mathcal{S}_N$ is $(\varepsilon,\vec{\Delta})$**-good** if:

$$\left|\frac{1}{N}H_N(\boldsymbol{x}) - E(\vec{\Delta})\right| \leq \varepsilon, \tag{4.13}$$

$$\left\|\vec{R}(\boldsymbol{G}^{(1)},\boldsymbol{x}) - \vec{R}(\vec{\Delta})\right\|_\infty \leq \varepsilon \tag{4.14}$$

$$\left\|\nabla_{\mathsf{rad}} H_N(\boldsymbol{x}) - \vec{x}(\vec{\Delta})\right\|_\infty \leq \varepsilon \tag{4.15}$$

$$\mathbb{W}_2\left(\widehat{\mu}_{H_N}(\boldsymbol{x}), \mu(\vec{x}(\vec{\Delta}))\right) \leq \varepsilon \quad \text{and} \quad d_{\mathcal{H}}\left(\mathsf{spec}(\nabla^2_{\mathsf{sp}} H_N(\boldsymbol{x})), S(\vec{\Delta})\right) \leq \varepsilon\,. \tag{4.16}$$

A point is ε**-good** if it is $(\varepsilon,\vec{\Delta})$-good for some $\vec{\Delta} \in \{-1,1\}^r$. Let $\mathcal{Q}(\varepsilon,\vec{\Delta}), \mathcal{Q}(\varepsilon) \subseteq \mathcal{S}_N$ be the sets of $(\varepsilon,\vec{\Delta})$-good and ε-good points, respectively.

Define $\mathsf{Crt}_N^{\mathsf{tot}} = \mathsf{Crt}_N^{\mathsf{tot}}(H_N)$ to be the set of critical points of H_N, and $\mathsf{Crt}_N^{\mathsf{good},\varepsilon} = \mathsf{Crt}_N^{\mathsf{tot}} \cap \mathcal{Q}(\varepsilon)$ and $\mathsf{Crt}_N^{\mathsf{bad},\varepsilon} = \mathsf{Crt}_N^{\mathsf{tot}} \setminus \mathcal{Q}(\varepsilon)$. The following theorem shows that solvability defines the phase boundary for annealed topological trivialization. It further shows that in strictly super-solvable models, all critical points are ε-good and thus belong to one of 2^r types corresponding to $\vec{\Delta} \in \{-1,1\}^r$, defined by Definition 4.1.10.

Theorem 4.1.11. *(a) If ξ is super-solvable, then*

$$\lim_{N\to\infty} \frac{1}{N} \log \mathbb{E}|\mathsf{Crt}_N^{\mathsf{tot}}| = 0. \tag{4.17}$$

(b) If ξ is strictly super-solvable, for all $\varepsilon > 0$, there exists $c = c(\xi,\varepsilon) > 0$ such that

$$\limsup_{N\to\infty} \frac{1}{N} \log \mathbb{E}|\mathsf{Crt}_N^{\mathsf{bad},\varepsilon}| \le -c. \tag{4.18}$$

(c) On the other hand, if ξ is strictly sub-solvable, then

$$\lim_{N\to\infty} \frac{1}{N} \log \mathbb{E}|\mathsf{Crt}_N^{\mathsf{tot}}| > 0. \tag{4.19}$$

Our next result states that H_N has *exactly* 2^r critical points with high probability when ξ is strictly super-solvable: one for each type $\vec{\Delta} \in \{-1,1\}^r$. Moreover all approximate critical points are near a true critical point, and (as a consequence) all approximate ground states are near the true ground state (recall Definition 4.1.5). We find it helpful to abstract some of these results into the following definition, which could easily be extended to other manifolds besides $\mathcal{S}_N$. (The first requirement, smoothness at the natural scale, holds with high probability by Proposition 4.2.4.)

Definition 4.1.12. We say the function $H_N : \mathcal{S}_N \to \mathbb{R}$ is (C,ε,ι)-**strongly topologically trivial** if:

(i) $\|\nabla^k H_N(\boldsymbol{\sigma})\|_{\mathsf{op}} \le C N^{1-\frac{k}{2}}$ for $k \in \{0,1,2,3\}$.

(ii) $|\mathsf{Crt}_N^{\mathsf{tot}}| \le C$.

(iii) All critical points of H_N are C-well-conditioned.

(iv) All critical points $\boldsymbol{x}$ of H_N besides the unique global maximum satisfy $\boldsymbol{\lambda}_{N/C}(\nabla^2_{\mathsf{sp}} H_N(\boldsymbol{x})) \ge 1/C$.

(v) All ε-approximate critical points of H_N are within distance $\iota\sqrt{N}$ of a critical point.

We say the sequence of functions $(H_N)_{N\ge 1}$ is C-**strongly topologically trivial** if for any $\iota > 0$, for $\varepsilon > 0$ sufficiently small, all but finitely many are (C,ε,ι)-strongly topologically trivial. We say the sequence is **strongly topologically trivial** if the previous condition holds for some finite C.

We note that from conditions (i) and (iii), it follows that if $\boldsymbol{x}$ is a critical point for H_N and $\|\widetilde{\boldsymbol{x}} - \boldsymbol{x}\| \ge C^{-4}\sqrt{N}$, then $\widetilde{\boldsymbol{x}}$ is not a C^{-10}-approximate critical point. Hence if condition (v) holds for ι which is small depending on C, then one can actually take $\iota = O(\varepsilon)$. The next main result shows that if ξ is strictly super-solvable then the sequence $(H_N)_{N\ge1}$ is almost surely topologically trivial. In fact, we precisely describe the 2^r critical points of H_N.

Theorem 4.1.13. *If ξ is strictly super-solvable, then the following holds with probability $1 - e^{-cN}$. For sufficiently small $\varepsilon > 0$, H_N has exactly one critical point $\boldsymbol{x}_{\vec{\Delta}}$ satisfying (4.13) through (4.16) for each $\vec{\Delta} \in \{-1,1\}^r$. Moreover for all $\iota > 0$ there exists $\varepsilon > 0$ such that:*

(a) All ε-critical points of H_N are $C(\vec{\lambda},\xi)$-well conditioned and lie in the disjoint union

$$\bigcup_{\vec{\Delta}\in\{-1,1\}^r} B_{\iota\sqrt{N}}(\boldsymbol{x}_{\vec{\Delta}}).$$

(b) The number of positive eigenvalues of $\nabla^2_{\mathsf{sp}} H_N(\boldsymbol{x}_{\vec{\Delta}})$ is exactly $\sum_{s\in\mathscr{S}} N_s \cdot 1_{\Delta_s=-1}$.

(c) *All ε-ground states lie in $B_{\iota\sqrt{N}}(\boldsymbol{x}_{\vec{1}})$.*

On the other hand, strong topological trivialization becomes false for strictly sub-solvable ξ (with probability $1-e^{-cN}$). This exhibits a natural *quenched* phase transition coinciding with the annealed transition of $\mathbb{E}|\mathsf{Crt}_N^{\mathsf{tot}}|$ in Theorem 4.1.11. Indeed our companion paper explicitly constructs exponentially many $\sqrt{N}/C$-separated approximate critical points whenever ξ is strictly sub-solvable, which contradicts parts (ii), (v) of Definition 4.1.12 when $\iota \le 1/(2C)$. The precise result is quoted below. (Note however that we give no quenched lower bounds on the number of *exact* critical points when ξ is strictly sub-solvable, which would be interesting to obtain.)

Proposition 4.1.14 ([HS24, Proposition 3.6])**.** *For any strictly sub-solvable ξ there exists $C(\vec{\lambda},\xi) > 0$ such that for any $\varepsilon > 0$ there exists $\delta > 0$ such that with probability $1-e^{-cN}$ the following holds. There exist $M = e^{\delta N}$ distinct ε-approximate critical points $\boldsymbol{x}_1,\dots,\boldsymbol{x}_M \in \mathcal{S}_N$ such that $\|\boldsymbol{x}_i - \boldsymbol{x}_j\|_2 \ge \sqrt{N}/C$ for all $1 \le i < j \le M$.*

4.1.4 Consequences for Langevin dynamics

Here we obtain dynamical consequences from our landscape results, showing that for any strongly topologically trivial landscape, low temperature Langevin dynamics mixes in $O(\log N)$ time. The main ingredient is a recent result by one of us [Sel24b] showing that, roughly speaking, low temperature Langevin dynamics can get stuck only in approximate local maxima.

Definition 4.1.15. Given a Hamiltonian $H_N : \mathcal{S}_N \to \mathbb{R}$, initial condition $\boldsymbol{X}(0) \in \mathcal{S}_N$, and $\beta \ge 0$, the β-Langevin dynamics $\boldsymbol{X}(t)$ driven by standard $\mathbb{R}^N$-valued Brownian motion $\boldsymbol{B}(t)$ is the process solving the stochastic differential equation

$$\mathsf{d}\boldsymbol{X}(t) = \left(\beta\nabla_{\mathsf{sp}}H_N(\boldsymbol{X}(t)) - \sum_{s\in\mathscr{S}} \frac{N_s-1}{2\lambda_s N}\boldsymbol{X}_s(t)\right)\mathsf{d}t + P^{\perp}_{\boldsymbol{X}(t)}\sqrt{2}\,\mathsf{d}\boldsymbol{B}(t). \tag{4.20}$$

Here $P^{\perp}_{\boldsymbol{X}(t)}$ is the rank $N-r$ projection matrix onto the orthogonal complement of $\operatorname{span}\big(\boldsymbol{X}_1(t),\dots,\boldsymbol{X}_r(t)\big)$.

Theorem 4.1.16. *Fix $\vec{\lambda}$ and let $(H_N)_{N\ge1}$ be a C-strongly topologically trivial sequence of functions $H_N : \mathcal{S}_N \to \mathbb{R}$ with unique local maximum $\boldsymbol{x}_* = \boldsymbol{x}_{*,N}$. (E.g. H_N as above, with $\boldsymbol{x}_* = \boldsymbol{x}(\vec{1})$.) Then:*

(a) *For any $\varepsilon > 0$, if $\beta \ge \beta_0(\vec{\lambda},\varepsilon,C)$ and $T \ge T_0(\vec{\lambda},\varepsilon,C)$ and N are sufficiently large, β-Langevin dynamics started from any $\boldsymbol{X}(0) \in \mathcal{S}_N$ satisfies with probability $1-e^{-cN}$:*

$$\inf_{t\in[T,T+e^{cN}]} H_N(\boldsymbol{X}(t)) \ge H_N(\boldsymbol{x}_*) - \varepsilon N, \tag{4.21}$$

$$\inf_{t\in[T,T+e^{cN}]} \|\boldsymbol{X}(t) - \boldsymbol{x}_*\|_2 \le \varepsilon\sqrt{N}. \tag{4.22}$$

(b) *For $\beta \ge \beta_0(\vec{\lambda},C)$, the β-Langevin dynamics has total variation mixing time $O(\log N)$.*

Proof. **Part (a):** We use [Sel24b, Theorem 1.2], the proof of which easily extends to finite products of spheres as considered here (see Remark 2.8 therein, and note that Proposition 4.2.4 below ensures the needed C-boundedness). The implication is that to prove (4.21) it suffices to show that for $\eta \le \eta_0(\xi,\vec{\lambda},\varepsilon) \le \varepsilon/2$ small enough (playing the role of ε therein), $E_*^{(\eta)}$ as defined in [Sel24b, Definition 2] satisfies

$$E_*^{(\eta)} \ge \frac{H_N(\boldsymbol{x}_*)}{N} - \frac{\varepsilon}{2}.$$

By said definition, this holds if no $\boldsymbol{\sigma} \in \mathcal{S}_N$ satisfies all of the following properties:

(1) $\|\nabla_{\mathsf{sp}}H_N(\boldsymbol{\sigma})\|_2 \le \eta\sqrt{N}$.

(2) The Hessian $\nabla^2_{\mathsf{sp}}H_N(\boldsymbol{\sigma})$ satisfies $\boldsymbol{\lambda}_{\lfloor \eta N\rfloor}\big(\nabla^2_{\mathsf{sp}}H_N(\boldsymbol{\sigma})\big) \le \eta$.

(3) $H_N(\boldsymbol{\sigma}) \le H_N(\boldsymbol{x}_*) - \frac{\varepsilon N}{2}$.

By the definition of strong topological trivialization, for small η condition (1) implies that $\boldsymbol{\sigma} \in B_{\iota\sqrt{N}}(\boldsymbol{x})$ for some critical point $\boldsymbol{x}$, where $\iota \to 0$ as $\eta \to 0$. Condition (2) implies $\boldsymbol{x} = \boldsymbol{x}_*$: otherwise Definition 4.1.12(iv) implies $\boldsymbol{\lambda}_{N/C}(\nabla^2_{\mathsf{sp}} H_N(\boldsymbol{x})) \ge 1/C$, so Definition 4.1.12(i) with $k = 3$ implies (for small enough ι) that $\boldsymbol{\lambda}_{N/C}(\nabla^2_{\mathsf{sp}} H_N(\boldsymbol{x})) \ge 1/(2C)$, which is a contradiction for $\eta_0 < 1/(2C)$. Thus $\boldsymbol{\sigma} \in B_{\iota\sqrt{N}}(\boldsymbol{x}_*)$, which (for small η) contradicts condition (3).

We conclude that (4.21) holds, and (4.22) follows (with a different choice of ε) thanks to Theorem 4.1.13.

Part (b): Let us choose ε small enough that $\boldsymbol{\lambda}_{\max}\big(\nabla^2_{\mathsf{sp}} H_N(\boldsymbol{x})\big) \le 0$ is negative semi-definite on $B_{\varepsilon\sqrt{N}}(\boldsymbol{x}_*)$; this is possible thanks to parts (i) (with $k = 3$) and (iii) of Definition 4.1.12. Also choose β large enough that Part (a) of this theorem applies with this choice of ε.

Let $\mathcal{C} \subseteq \mathcal{S}_N$ be the product of diameter $\varepsilon\sqrt{N}/r$ spherical caps inside each $(\mathcal{S}_{N,s})_{1\le s\le r}$ centered at $\boldsymbol{x}_*$. Each of the r factors is a convex Riemannian manifold with boundary (see e.g. [Kro79]), hence so is $\mathcal{C}$. We consider the reflected Langevin dynamics with inward normal reflection in $\mathcal{C}$, as constructed in e.g. [CZ17, Section 2.1]. Let P_t be the transition kernel for ordinary Langevin dyamics, and $\widetilde{P}_t$ that of the reflected dynamics.

Recall that $\mathcal{S}_N$ has uniformly positive Ricci curvature (as N varies) and $\nabla^2_{\mathsf{sp}} H_N(\boldsymbol{x})$ is negative semidefinite on all of $\mathcal{C}$. It follows that the Gibbs measure $\mathsf{d}\nu_\beta(\boldsymbol{x}) = Z_{N,\beta}^{-1} e^{\beta H_N(\boldsymbol{x})}\mathsf{d}\boldsymbol{x}$ also has uniformly positive Ricci curvature within $\mathcal{C}$ (see e.g. [GJ19, Proposition 22]). Let $\widetilde{\nu}_\beta$ be the Gibbs measure ν_β of H_N conditioned to lie in $\mathcal{C}$. By [Wan14, Theorem 3.3.2] and convexity of the manifold $\mathcal{C}$, it follows that the reflected Langevin dynamics inside $\mathcal{C}$ contracts exponentially in (Riemannian) Wasserstein distance: for any probability measures ρ_0, ρ_0' on $\mathcal{C}$,

$$\mathbb{W}_2(\widetilde{P}_t\rho_0, \widetilde{P}_t\rho_0') \le e^{-ct}\mathbb{W}_2(\rho_0, \rho_0') \le e^{-ct}\sqrt{N}, \qquad c = c(\vec{\lambda}, \xi) > 0 \tag{4.23}$$

Finally, we combine (4.23) with [BGL01, Lemma 4.2] for a small time δ (denoted T therein) to find (using Pinsker's inequality in the first step)

$$\|\widetilde{P}_{t+\delta}\rho_0 - \widetilde{\nu}_\beta\|_{TV}^2 \le \mathsf{Ent}\big(\widetilde{P}_{t+\delta}\rho_0 \| \widetilde{\nu}_\beta\big) \le CNe^{-2ct}. \tag{4.24}$$

([BGL01, Lemma 4.2] is stated for Euclidean space, but all proof ingredients remain available by [Wan14, Theorem 3.3.2].)

On the other hand, it follows from Part (a) that from any initial $\boldsymbol{X}(0) \in \mathcal{S}_N$, with probability $1 - e^{-cN}$, we have $\boldsymbol{X}(0) \in \mathcal{C}$ for $T \le t \le e^{cN}$. Taking the convention that $\widetilde{P}_t\delta_{\boldsymbol{y}} = \delta_{\boldsymbol{y}}$ for $\boldsymbol{y} \in \mathcal{S}_N \backslash \mathcal{C}$, we write for any $\boldsymbol{x} \in \mathcal{S}_N$:

$$\begin{aligned}\|P_{2t}\delta_{\boldsymbol{x}} - \nu_\beta\|_{TV} &\le \|P_{2t}\delta_{\boldsymbol{x}} - \widetilde{P}_t(P_t\delta_{\boldsymbol{x}})\|_{TV} + \|\widetilde{P}_t(P_t\delta_{\boldsymbol{x}}) - \widetilde{\nu}_\beta\|_{TV} + \|\widetilde{\nu}_\beta - \nu_\beta\|_{TV} \\ &\le e^{-cN} + CNe^{-ct} + e^{-cN}.\end{aligned}$$

Here the bound on $\|P_{2t}\delta_{\boldsymbol{x}} - \widetilde{P}_t(P_t\delta_{\boldsymbol{x}})\|_{TV}$ follows from the definition of $\mathcal{C}$, which ensures that with probability e^{-cN}, the ordinary and reflected Langevin dynamics (with shared Brownian motion) agree for t units of time when started from $\boldsymbol{X}(t)$. The bound on $\|\widetilde{\nu}_\beta - \nu_\beta\|_{TV}$ follows from Part (a).

In particular, for $t/(\log N)$ at least a large constant, we deduce that for N large enough,

$$\sup_{\boldsymbol{x}\in\mathcal{S}_N} \|P_{2t}\delta_{\boldsymbol{x}} - \nu_\beta\|_{TV} \le 1/4$$

which concludes the proof. □

Remark 4.1.17. The above proof established that any C-strongly topologically trivial sequence H_N satisfies the Bakry-Emery conditions on a small neighborhood of the global optimum for all $\beta \ge 0$. This implies exponential concentration of overlaps (via concentration of Lipschitz functions). Namely for any $\varepsilon > 0$, with $\boldsymbol{\sigma}, \widetilde{\boldsymbol{\sigma}} \overset{i.i.d}{\sim} \nu_\beta$ for β large, there exists $\vec{q}_* \in [0,1]^r$ depending on H_N such that

$$\mathbb{P}[\|\vec{R}(\boldsymbol{\sigma}, \widetilde{\boldsymbol{\sigma}}) - \vec{q}_*\|_\infty \le \varepsilon] \ge 1 - e^{-cN}.$$

For spin glass Hamiltonians distributed according to some strictly super-solvable ξ, the value $\vec{q}_*$ can be chosen deterministically depending only on $(\vec{\lambda},\xi,\beta)$. This is because the restricted free energies

$$\frac{1}{N}\log \iint_{\substack{\boldsymbol{\sigma},\widetilde{\boldsymbol{\sigma}}\in\mathcal{S}_N,\\ \|\vec{R}(\boldsymbol{\sigma},\widetilde{\boldsymbol{\sigma}})-\vec{q}\|_\infty\le\varepsilon}} \exp(\beta(H_N(\boldsymbol{\sigma})+H_N(\widetilde{\boldsymbol{\sigma}}))\ \mathsf{d}\boldsymbol{\sigma}\mathsf{d}\widetilde{\boldsymbol{\sigma}}.$$

concentrate exponentially for any $\vec{q}$ and $\varepsilon>0$.

To take $\vec{q}$ independent of N one may fix a large $T>0$, and deterministic $\boldsymbol{X}(0)\in\mathcal{S}_N$ for each N, and combine the following two observations. First $\mathbb{E}\mathbb{W}_2(P_T\delta_{\boldsymbol{X}(0)},\nu_\beta)\le Ce^{-cT}\sqrt{N}+e^{-cN}$ since (4.23) holds with exponentially good probability. Second, one may show via two-replica multi-species analogs of the Cugliandolo–Kurchan equations [CHS93, CK94, BDG06, DGM07] that $\vec{R}(\boldsymbol{X}(T),\widetilde{\boldsymbol{X}}(T))$ concentrates exponentially around an N-independent value $\vec{q}_*(T)$ when $\boldsymbol{X}(T),\widetilde{\boldsymbol{X}}(T)$ are driven by independent Brownian motions for the same H_N. Indeed from [Sel24b, Lemma 3.1 and Subsection 4.2] it suffices to prove these Cugliandolo–Kurchan equations for soft spherical Langevin dynamics, which follows mechanically from the approach of [CCM21]. From these observations, taking $T\to\infty$ after $N\to\infty$, this implies that overlaps concentrate around $\vec{q}_*$.

Remark 4.1.18. We argued above that the following holds for strictly super-solvable ξ with probability $1-e^{-cN}$: the expected hitting time of a radius $\delta\sqrt{N}$ neighborhood $B_{\delta\sqrt{N}}(\boldsymbol{x}(\vec{1}))$ of the global optimum of H_N is at most $C(\vec{\lambda},\xi,\delta)$ for β sufficiently large, uniformly in $\boldsymbol{X}(0)$. Combining these two facts, the Lyapunov function technique of [BBCG08] implies that the Gibbs measure ν_β has Poincaré constant at most $C(\vec{\lambda},\xi)$ for β sufficiently large. The relevant Lyapunov function $L:\mathcal{S}_N\to\mathbb{R}_{\ge 1}$ is essentially an exponential moment of the hitting time of $B_{\delta\sqrt{N}}(\boldsymbol{x}(\vec{1}))$; see [LE23, Proposition 9.13 and Appendix B.4] for a detailed derivation of this implication.

Acknowledgments

Thanks to Paul Bourgade, Jiří Černý, Sinho Chewi, László Erdős, Mufan Li, David Benjamin Lim, Benjamin McKenna, Eliran Subag, and the anonymous referees for helpful discussions, suggestions and comments.

4.2 Further preliminaries

Below we provide further notations and background. Subsection 4.2.1 will be assumed throughout the entire paper. Subsections 4.2.2 and 4.2.3 will be used primarily in Section 4.4.

4.2.1 Geometry of $\mathcal{S}_N$

Definition 4.2.1. A linear subspace $U\subseteq\mathbb{R}^N$ is **species-aligned** if it is the direct sum of subspaces $U_s\subseteq\mathbb{R}^{\mathcal{I}_s}$, for $s\in\mathscr{S}$.

For $\boldsymbol{z}\in\mathbb{R}^N$ or a species-aligned subspace $U\subseteq\mathbb{R}^N$, we define

$$\boldsymbol{z}^\perp=\left\{\boldsymbol{x}\in\mathbb{R}^N\ :\ \vec{R}(\boldsymbol{z},\boldsymbol{x})=\vec{0}\right\},\qquad U^\perp=\left\{\boldsymbol{x}\in\mathbb{R}^N\ :\ \vec{R}(\boldsymbol{u},\boldsymbol{x})=\vec{0}\ \forall\boldsymbol{u}\in U\right\}. \tag{4.25}$$

Recalling the definitions in Fact 4.1.4, we now explicitly describe the law of the local behavior of H_N around a given $\boldsymbol{\sigma}\in\mathcal{S}_N$.

Lemma 4.2.2. *Fix $\boldsymbol{\sigma}\in\mathcal{S}_N$. The random variables $\nabla_{\mathcal{T}}H_N(\boldsymbol{\sigma})$, $\nabla^2_{\mathcal{T}\times\mathcal{T}}H_N(\boldsymbol{\sigma})$, and $(H_N(\boldsymbol{\sigma}),\vec{R}(\boldsymbol{G}^{(1)},\boldsymbol{\sigma}),\nabla_{\mathsf{rad}}H_N(\boldsymbol{\sigma}))$ are mutually independent Gaussians with the following distributions.*

(a) Tangential derivative: for each $i\in\mathcal{T}$, $\partial_i H_N(\boldsymbol{\sigma})\sim\mathcal{N}(0,\xi^{s(i)})$ and these are independent across i.

(b) Tangential Hessian: $\boldsymbol{W}=\nabla^2_{\mathcal{T}\times\mathcal{T}}H_N(\boldsymbol{\sigma})$ is a symmetric random matrix with independent centered Gaussian entries on and above the diagonal, where

$$\mathbb{E}[W_{i,j}^2]=\frac{(1+\delta_{i,j})\xi''_{s(i),s(j)}}{N\lambda_{s(i)}\lambda_{s(j)}}\,. \tag{4.26}$$

(c) *Energy, 1-spin overlap, and radial derivative:* $(H_N(\boldsymbol{\sigma}), \vec{R}(\boldsymbol{G}^{(1)}, \boldsymbol{\sigma}), \nabla_{\mathsf{rad}} H_N(\boldsymbol{\sigma}))$ *is a centered Gaussian vector with covariance satisfying*

$$\mathbb{E}\left[H_N(\boldsymbol{\sigma})^2\right] = N\xi(\vec{1}), \tag{4.27}$$

$$\mathbb{E}\left[\vec{R}(\boldsymbol{G}^{(1)}, \boldsymbol{\sigma})\vec{R}(\boldsymbol{G}^{(1)}, \boldsymbol{\sigma})^\top\right] = N^{-1}\Lambda^{-1}, \tag{4.28}$$

$$\mathbb{E}\left[\nabla_{\mathsf{rad}} H_N(\boldsymbol{\sigma})\nabla_{\mathsf{rad}} H_N(\boldsymbol{\sigma})^\top\right] = N^{-1}\Lambda^{-1/2}A\Lambda^{-1/2}, \tag{4.29}$$

$$\mathbb{E}\left[H_N(\boldsymbol{\sigma})\nabla_{\mathsf{rad}} H_N(\boldsymbol{\sigma})\right] = \Lambda^{-1/2}\xi', \tag{4.30}$$

$$\mathbb{E}\left[\vec{R}(\boldsymbol{G}^{(1)}, \boldsymbol{\sigma})\nabla_{\mathsf{rad}} H_N(\boldsymbol{\sigma})^\top\right] = N^{-1}\mathrm{diag}(\Gamma^{(1)})\Lambda^{-1}. \tag{4.31}$$

As a consequence,

$$\mathbb{E}\left[H_N(\boldsymbol{\sigma})|\nabla_{\mathsf{rad}} H_N(\boldsymbol{\sigma})\right] = N(\xi')^\top A^{-1}\Lambda^{1/2}\nabla_{\mathsf{rad}} H_N(\boldsymbol{\sigma}), \tag{4.32}$$

$$\mathbb{E}\left[\vec{R}(\boldsymbol{G}^{(1)}, \boldsymbol{\sigma})|\nabla_{\mathsf{rad}} H_N(\boldsymbol{\sigma})\right] = \mathrm{diag}(\Gamma^{(1)})\Lambda^{-1/2}A^{-1}\Lambda^{1/2}\nabla_{\mathsf{rad}} H_N(\boldsymbol{\sigma}), \tag{4.33}$$

$$\mathrm{Var}\left[H_N(\boldsymbol{\sigma})|\nabla_{\mathsf{rad}} H_N(\boldsymbol{\sigma})\right] = N\left(\xi(\vec{1}) - (\xi')^\top A^{-1}\xi'\right). \tag{4.34}$$

Proof. Due to the symmetry of the sphere it suffices to verify these statements for $\boldsymbol{\sigma}$ equal to the "r-tuple north pole," i.e. $\sigma_{m_s} = \sqrt{\lambda_s N}$ for all $m_s \in \mathcal{R}$, and $\sigma_i = 0$ for all $i \in \mathcal{T}$. Then $\nabla_{\mathcal{T}} H_N(\boldsymbol{\sigma})$, $\nabla^2_{\mathcal{T}\times\mathcal{T}} H_N(\boldsymbol{\sigma})$, and $(H_N(\boldsymbol{\sigma}), \vec{R}(\boldsymbol{G}^{(1)}, \boldsymbol{\sigma}), \nabla_{\mathsf{rad}} H_N(\boldsymbol{\sigma}))$ can be evaluated as explicit linear combinations of disorder coefficients, which readily implies the stated covariance structure. In particular, they are mutually independent because the sets of disorder coefficients contributing to them are disjoint. As an example calculation (see also [McK24, Section 3.3]),

$$\partial_{m_s} H_N(\boldsymbol{\sigma}) = \lambda_s^{-1/2}\sum_{p\geq 1}\sum_{s_1,\dots,s_p} n_s(s_1,\dots,s_p)\gamma_{s_1,\dots,s_p}\boldsymbol{G}^{(p)}_{m_{s_1},\dots,m_{s_p}}\sqrt{\lambda_{s_1}\lambda_{s_2}\cdots\lambda_{s_p}},$$

where $n_s(s_1,\dots,s_p)$ is the number of times s appears in $s_1,\dots,s_p$. This readily implies that

$$\begin{aligned}
\mathbb{E}\left[\partial_{m_s} H_N(\boldsymbol{\sigma})^2\right] &= \lambda_s^{-1}\sum_{p\geq 1}\sum_{s_1,\dots,s_p} n_s(s_1,\dots,s_p)^2\gamma^2_{s_1,\dots,s_p} = \lambda_s^{-1}(\xi'_s + \xi''_{s,s}), \\
\mathbb{E}\left[\partial_{m_s} H_N(\boldsymbol{\sigma})\partial_{m_{s'}} H_N(\boldsymbol{\sigma})\right] &= \lambda_s^{-1/2}\lambda_{s'}^{-1/2}\sum_{p\geq 1}\sum_{s_1,\dots,s_p} n_s(s_1,\dots,s_p)n_{s'}(s_1,\dots,s_p)\gamma^2_{s_1,\dots,s_p} \\
&= \lambda_s^{-1/2}\lambda_{s'}^{-1/2}\xi''_{s,s'},
\end{aligned}$$

which implies (4.29). The rest of (4.27) through (4.31) are verified similarly. The formula (4.32) is verified from the standard fact

$$\begin{aligned}
&\mathbb{E}\left[H_N(\boldsymbol{\sigma})\,|\,\nabla_{\mathsf{rad}} H_N(\boldsymbol{\sigma})\right] \\
&= \mathbb{E}\left[H_N(\boldsymbol{\sigma})\nabla_{\mathsf{rad}} H_N(\boldsymbol{\sigma})\right]^\top \mathbb{E}\left[\nabla_{\mathsf{rad}} H_N(\boldsymbol{\sigma})\nabla_{\mathsf{rad}} H_N(\boldsymbol{\sigma})^\top\right]^{-1}\nabla_{\mathsf{rad}} H_N(\boldsymbol{\sigma}),
\end{aligned}$$

and (4.33) follows similarly. Finally (4.34) follows from

$$\mathrm{Var}\left[H_N(\boldsymbol{\sigma})|\nabla_{\mathsf{rad}} H_N(\boldsymbol{\sigma})\right] = \mathbb{E}\left[H_N(\boldsymbol{\sigma})^2\right] - \mathbb{E}\left[\mathbb{E}\left[H_N(\boldsymbol{\sigma})\,|\,\nabla_{\mathsf{rad}} H_N(\boldsymbol{\sigma})\right]^2\right]. \qquad \square$$

Fact 4.2.3. *The volume of $\mathcal{S}_N$ w.r.t. the $(N-r)$-dimensional Hausdorff measure $\mathcal{H}^{N-r}$ satisfies*

$$\frac{1}{N}\log\mathcal{H}^{N-r}(\mathcal{S}_N) = \frac{1+\log(2\pi)}{2} + o_N(1).$$

Proof. By Stirling's approximation, the volume of $\sqrt{N}\mathbb{S}^{N-1}$ is

$$\frac{2\pi^{N/2}N^{(N-1)/2}}{\Gamma(N/2)} = e^{o(N)}\frac{(\pi N)^{N/2}}{(N/2e)^{N/2}} = e^{o(N)}(2\pi e)^{N/2}.$$

Thus the volume of $\mathcal{S}_N$ is

$$\mathsf{Vol}(\mathcal{S}_N) = e^{o(N)} \prod_{s\in\mathscr{S}} (2\pi e)^{\lambda_s N/2} = e^{o(N)}(2\pi e)^{N/2}\,. \qquad \square$$

Let $\mathscr{H}_N$ denote the space of possible Hamiltonians H_N, which we identify as (infinite-dimensional) vectors consisting of their disorder coefficients $(\boldsymbol{G}^{(p)})_{p\geq 1}$ concatenated in an arbitrary but fixed order. Also let $S_N = \{\boldsymbol{x}\in\mathbb{R}^N : \|\boldsymbol{x}\|_2^2 = N\}$ and for any tensor $\boldsymbol{A}\in(\mathbb{R}^N)^{\otimes k}$, define the operator norm

$$\|\boldsymbol{A}\|_{\mathsf{op}} = \max_{\|\boldsymbol{\sigma}^1\|_2,\dots,\|\boldsymbol{\sigma}^k\|_2\leq 1} |\langle \boldsymbol{A}, \boldsymbol{\sigma}^1\otimes\cdots\otimes\boldsymbol{\sigma}^k\rangle|\,.$$

Proposition 4.2.4 ([HS23a, Proposition 1.13])**.** *For any ξ there exists $c>0$, a sequence $(K_N)_{N\geq 1}$ of symmetric convex sets $K_N\subseteq\mathscr{H}_N$, and constant $C=C(\xi)$, such that the following holds.*

(a) $\P[H_N\in K_N]\geq 1-e^{-cN}$;

(b) For all $H_N\in K_N$, $k\leq 3$, and $\boldsymbol{x},\boldsymbol{y}\in\mathcal{S}_N$,

$$\left\|\nabla^k H_N(\boldsymbol{x})\right\|_{\mathsf{op}} \leq C N^{1-\frac{k}{2}}, \tag{4.35}$$

$$\left\|\nabla^k H_N(\boldsymbol{x}) - \nabla^k H_N(\boldsymbol{y})\right\|_{\mathsf{op}} \leq C N^{\frac{1-k}{2}} \|\boldsymbol{x}-\boldsymbol{y}\|_2. \tag{4.36}$$

Proposition 4.2.5. *For symmetric matrices $M, M'\in\mathbb{R}^{r\times r}$ we have (recall (4.8)):*

$$d_{\mathcal{H}}\big(\mathsf{spec}(M),\mathsf{spec}(M')\big) \leq \mathbb{W}_\infty\big(\widehat{\mu}(M),\widehat{\mu}(M')\big) \leq \|M-M'\|_{\mathsf{op}}.$$

In particular for $H_N\in K_N$ and all $\boldsymbol{x},\boldsymbol{y}\in\mathcal{S}_N$:

$$d_{\mathcal{H}}\left(\mathsf{spec}(\nabla^2_{\mathsf{sp}} H_N(\boldsymbol{x})), \mathsf{spec}(\nabla^2_{\mathsf{sp}} H_N(\boldsymbol{y}))\right) \leq \frac{C}{\sqrt{N}}\|\boldsymbol{x}-\boldsymbol{y}\|_2.$$

Proof. The first part is immediate from the Weyl inequalities. For the second part,

$$\begin{aligned}
&\left\|\nabla^2_{\mathsf{sp}} H_N(\boldsymbol{x}) - \nabla^2_{\mathsf{sp}} H_N(\boldsymbol{y})\right\|_{\mathsf{op}} \\
&\leq \left\|\nabla^2_{\mathcal{T}\times\mathcal{T}} H_N(\boldsymbol{x}) - \nabla^2_{\mathcal{T}\times\mathcal{T}} H_N(\boldsymbol{y})\right\|_{\mathsf{op}} \\
&\qquad + \left\|\mathrm{diag}(\Lambda^{-1/2}(\nabla_{\mathsf{rad}} H_N(\boldsymbol{x}) - \nabla_{\mathsf{rad}} H_N(\boldsymbol{y}))\diamond \mathbf{1}_{\mathcal{T}})\right\|_{\mathsf{op}} \\
&\leq \left\|\nabla^2 H_N(\boldsymbol{x}) - \nabla^2 H_N(\boldsymbol{y})\right\|_{\mathsf{op}} + \frac{1}{\sqrt{N\min\vec{\lambda}}}\|\mathrm{diag}(\nabla_{\mathcal{R}} H_N(\boldsymbol{x}) - \nabla_{\mathcal{R}} H_N(\boldsymbol{y}))\|_{\mathsf{op}}\,.
\end{aligned}$$

The final term is bounded by

$$\begin{aligned}
\|\mathrm{diag}(\nabla_{\mathcal{R}} H_N(\boldsymbol{x}) - \nabla_{\mathcal{R}} H_N(\boldsymbol{y}))\|_{\mathsf{op}} &\leq \|\nabla H_N(\boldsymbol{x}) - \nabla H_N(\boldsymbol{y})\|_2 \\
&= \|\nabla H_N(\boldsymbol{x}) - \nabla H_N(\boldsymbol{y})\|_{\mathsf{op}}.
\end{aligned}$$

The result now follows from the first part and Proposition 4.2.4. $\square$

4.2.2 Elementary linear algebra

Definition 4.2.6. A symmetric matrix $M\in\mathbb{R}^{r\times r}$ is **diagonally signed** if $M_{i,i}\geq 0$ and $M_{i,j}<0$ for all distinct $i,j\in[r]$.

Lemma 4.2.7. *If $M\in\mathbb{R}^{r\times r}$ is diagonally signed, then the minimal eigenvalue $\boldsymbol{\lambda}_{\min}(M)$ has multiplicity 1, and the corresponding eigenvector $\vec{w}$ has strictly positive entries. Moreover,*

$$\boldsymbol{\lambda}_{\min}(M) = \sup_{\vec{v}\succ\vec{0}} \min_{s\in\mathscr{S}} \frac{(M\vec{v})_s}{v_s}\,.$$

Proof. [HS23a, Proposition 4.3] shows the final equality, and the proof therein shows that any minimal eigenvector $\vec{w}$ of M must have strictly positive entries. Since M is symmetric its eigenvectors are orthogonal, so $\vec{w}$ is unique. $\square$

Lemma 4.2.8. *If $M \in \mathbb{R}^{r\times r}$ is diagonally signed, $M \succeq 0$, and $M' \in \mathbb{R}^{r\times r}$ is defined by $M'_{i,j} = |M_{i,j}|$, then $M' \succeq 0$.*

Proof. By Lemma 4.2.7, the minimal eigenvector $\vec{w}$ of M has strictly positive entries. Let $\boldsymbol{\lambda}_{\min}(M) = t \geq 0$. The equation $M\vec{w} = t\vec{w}$ implies that for any $s \in \mathscr{S}$,

$$(M_{s,s} - t)w_s + \sum_{s' \neq s} M_{s,s'} w_{s'} = 0 \qquad \Longrightarrow \qquad M_{s,s} = t + \sum_{s' \neq s} |M_{s,s'}| \frac{w_{s'}}{w_s}\,.$$

Thus for any $\vec{x} \in \mathbb{R}^r$,

$$\langle \vec{x}, M'\vec{x} \rangle = t\|\vec{x}\|_2^2 + \sum_{s \neq s'} |M_{s,s'}| \left(\sqrt{\frac{w_{s'}}{w_s}} x_s + \sqrt{\frac{w_s}{w_{s'}}} x_{s'} \right)^2 \geq 0\,. \qquad \square$$

Corollary 4.2.9. *If ξ is super-solvable, then (recall (4.11)) $A \succ 0$.*

Proof. Applying Lemma 4.2.8 to $M = \mathrm{diag}(\xi') - \xi'' \succeq 0$ shows $M' = A - 2\mathrm{diag}(\xi'') \succeq 0$, where $\mathrm{diag}(\xi'') \in \mathbb{R}^{r\times r}$ denotes the diagonal matrix with the same diagonal entries as ξ''. By Assumption 4.1.1, $A \succ 0$. $\square$

4.2.3 Random matrix theory

Our calculations will involve standard notions from random matrix theory. For a probability measure $\mu \in \mathcal{P}(\mathbb{R})$, and with $\mathbb{H}$ the open complex upper half-space, its Stiejtles transform $m : \mathbb{H} \to \mathbb{C}$ is the holomorphic function

$$m(z) = \int \frac{\mu(\mathsf{d}\gamma)}{\gamma - z}, \qquad z \in \mathbb{H}. \tag{4.37}$$

If μ is compactly supported with piecewise smooth density $\rho(x)$, it is well known (see e.g. [AGZ10, Chapter 2.4]) that m extends continuously to $\mathbb{R}$ at all points of smoothness with $m(x) = \Im(\pi\rho(x))$. Here and throughout, we use $\Re(\cdot)$ and $\Im(\cdot)$ respectively to denote the real and imaginary parts of a complex scalar or vector. Throughout this paper it will also be understood that $m(x) = \lim_{z\in\mathbb{H}, z\to x} m(z)$.

Next, fixing $\vec{x} \in \mathbb{R}^r$, let $m_s(z) = m_s(z; \vec{x}) \in \mathbb{H}$ solve the vector Dyson equation

$$1 + \left(z + \frac{x_s}{\sqrt{\lambda_s}} + \sum_{s' \in \mathscr{S}} \frac{\xi''_{s,s'}}{\lambda_s} m_{s'}(z) \right) m_s(z) = 0\,, \qquad z \in \mathbb{H}\,. \tag{4.38}$$

For each s, let μ_s be such that m_s is the Stieltjes transform of μ_s, existence and uniqueness of which is guaranteed by Proposition 4.2.10 below. Then define

$$\mu = \sum_{s \in \mathscr{S}} \lambda_s \mu_s, \tag{4.39}$$

$$m(z) = m(z; \vec{x}) = \sum_{s \in \mathscr{S}} \lambda_s m_s(z). \tag{4.40}$$

We will sometimes write $\mu = \mu_{\xi,\vec{\lambda}}(\vec{x})$ to emphasize the dependence on $\xi, \vec{\lambda}, \vec{x}$ (or include some arguments but not others). The next proposition details useful properties of m_s and μ_s. Note that $\mu_{\xi,\vec{\lambda}}(\vec{x})$ depends only on $(\xi'', \vec{\lambda}, \vec{x})$. In particular this is a finite-dimensional vector (while ξ is in principle infinite-dimensional). We also let $\vec{\lambda}^{\circ}_N = (\lambda^{\circ}_{N,s})_{s\in\mathscr{S}} \in \mathbb{R}^r$, where

$$\lambda^{\circ}_{N,s} = \frac{N_s - 1}{N - r} = \frac{|\mathcal{I}_s| - 1}{N - r}. \tag{4.41}$$

These slightly modified values of $\vec{\lambda}_N$ will be useful because they are the exact relative sizes of the species blocks in M_N (see also [McK24, Eq. (2.1)]). Of course $\vec{\lambda}^{\circ}_N \to \vec{\lambda}$ as $N \to \infty$ since we assume $\vec{\lambda}_N \to \vec{\lambda}$.

Proposition 4.2.10. *For each $\vec{x} \in \mathbb{R}^r$, there exists a unique solution $(m_1, \dots, m_r)$ to (4.38) consisting of holomorphic functions $m_s : \mathbb{H} \to \mathbb{H}$, each given by the Stieltjes transform of some $\mu_s \in \mathcal{P}(\mathbb{R})$. Moreover for any compact set $\mathcal{K} \subseteq (0,1)^r \times (0,\infty)^{r\times r} \times \mathbb{R}^r$ there exists $C = C(\mathcal{K})$ such that the following hold whenever $(\vec{\lambda}_N^\circ, \xi'', \vec{x}) \in \mathcal{K}$ (and $\sum_s \lambda_{N,s}^\circ = 1$).*

(a) The support sets $\mathsf{supp}(\mu_s) \subseteq \mathbb{R}$ are contained in $[-C, C]$ and do not depend on s.

(b) Each μ_s is absolutely continuous, with density ρ_s having $1/3$-Hölder norm at most C, piece-wise smooth on at most C intervals with disjoint interiors, and otherwise zero.

(c) Each $m_s(\,\cdot\,;\cdot)$ extends to a jointly continuous function $\overline{\mathbb{H}} \times \mathbb{R}^r \mapsto \overline{\mathbb{H}}$ solving (4.38), with $1/3$-Hölder norm at most C.

(d) $|m_s(z;\vec{x})| \geq 1/C$ for all $z \in \overline{\mathbb{H}}$ and $\vec{x} \in \mathbb{R}^r$.

(e) For any $\varepsilon > 0$ and any fixed $\boldsymbol{x} \in \mathcal{S}_N$, conditionally on $\nabla_{\mathsf{rad}} H_N(\boldsymbol{x}) = \vec{x}$, we have the bulk-typicality (recall (4.10)):

$$\begin{aligned} \mathbb{W}_2\big(\widehat{\mu}_{H_N}(\boldsymbol{x}), \mu_{\xi,\vec{\lambda}_N^\circ}(\vec{x})\big) &\leq \varepsilon, \\ d_{\mathcal{H}}\big(\mathsf{supp}(\widehat{\mu}_{H_N}(\boldsymbol{x})), \mathsf{supp}(\mu_{\xi,\vec{\lambda}_N^\circ}(\vec{x}))\big) &\leq \varepsilon. \end{aligned} \tag{4.42}$$

with probability $1 - e^{-cN}$ for $c = c(\varepsilon, \mathcal{K}) > 0$ and N large enough. (Recall that $\mu_{\vec{\lambda}_N^\circ}(\vec{x})$ is defined by (4.38), (4.39) with $\vec{\lambda}_N^\circ$ in place of $\vec{\lambda}$.)

Proof. The first three statements follow by [AEK17a, Proposition 2.1, Theorem 2.6, Corollary 2.7] (see the last sentence of [AEK19a, Theorem 2.4] for relevant local uniformity statements), except for the continuity in $\vec{x}$ in part (c). This is proved in the Appendix as Theorem 4.A.2 (which also allows $\vec{v} \in \overline{\mathbb{H}}^r$). Part (d) follows since $m_s \approx 0$ is impossible in (4.38).

We now explain part (e), which requires a bit more work. Throughout, we argue conditionally on $\nabla_{\mathsf{rad}} H_N(\boldsymbol{x}) = \vec{x}$. To start, the random matrix $\nabla^2_{\mathsf{sp}} H_N(\boldsymbol{x})$ obeys the general conditions of [AEKN19, Theorem 4.7(i)]; in particular its conditional mean (recall Fact 4.1.4) $\mathbb{E}[\nabla^2_{\mathsf{sp}} H_N(\boldsymbol{x}) | \nabla_{\mathsf{rad}} H_N(\boldsymbol{x}) = \vec{x}] = -\mathrm{diag}(\Lambda^{-1/2}\vec{x} \diamond \mathbf{1}_{\mathcal{T}})$ is bounded in operator norm by a constant (denoted κ_4 in [AEKN19]), uniformly for all $\vec{x}$ in any given compact set (with the precise value κ_4 depending on the compact set). This result implies[1] that with probability at least $1 - O(1/N)$ (with implicit constant uniform over compact sets of $\vec{x}$), the set $\mathsf{supp}(\widehat{\mu}_{H_N}(\boldsymbol{x}))$ is contained within an $\varepsilon/2$-neighborhood of $\mathsf{supp}(\mu_{\xi,\vec{\lambda}_N^\circ}(\vec{x}))$. We first improve this probability to be exponentially close to 1. By the Hoffman–Wielandt lemma (see e.g. [AGZ10, Lemma 2.1.19]), the k-th eigenvalue of any symmetric matrix is an 1-Lipschitz function of its entries. In our setting, Lemma 4.2.2 implies that conditionally on $\vec{x}$, the entries of $\nabla^2_{\mathsf{sp}} H_N(\boldsymbol{x})$ are independent Gaussians up to symmetry, each with variance $O(1/N)$ (indeed Lemma 4.2.2 shows that this variance is exactly determined by ξ and does not depend on $\vec{x}$ in any way). By concentration of Lipschitz functions of Gaussians, we find that $\lambda_k = \lambda_k(\nabla^2_{\mathsf{sp}} H_N(\boldsymbol{x}))$ satisfies for any $\vec{x}$:

$$\mathbb{P}[|\lambda_k - \mathbb{E}[\lambda_k]| \geq \varepsilon/4] \leq e^{-c(\varepsilon)N}.$$

In particular, if λ_k, λ_k' are IID copies, then $\mathbb{P}[|\lambda_k - \lambda_k'| \geq \varepsilon/2] \leq 2e^{-c(\varepsilon)N}$ by the triangle inequality. With $E_k(\varepsilon)$ the event that $d(\lambda_k, \mathsf{supp}(\mu_{\xi,\vec{\lambda}_N^\circ}(\vec{x}))) \geq \varepsilon$, we thus see that

$$\mathbb{P}[E_k(\varepsilon)] \cdot (1 - O(1/N)) \leq 2e^{-c(\varepsilon)N}.$$

This is because if $E_k(\varepsilon)$ holds for λ_k but λ_k' obeys the conclusion above from [AEKN19, Theorem 4.7(i)], then $|\lambda_k - \lambda_k'| \geq \varepsilon/2$ must hold. We thus find that $\mathbb{P}[E_k(\varepsilon)] \leq e^{-c'(\varepsilon)N}$ for N large.

Next, we employ [AEK17b, Corollary 1.10], which shows that the bounded Lipschitz distance d_{BL} between $\widehat{\mu}_{H_N}(\boldsymbol{x})$ and $\mu_{\xi,\vec{\lambda}_N^\circ}(\vec{x})$ tends to 0, with probability $1 - O(1/N)$ and uniform implicit constant over compact sets

[1] In translating [AEKN19, Theorem 4.7(i)], we use the exact equivalence between size $N - r$ Dyson equations with constant entries on the partitions $(\mathcal{I}_s - 1) \times (\mathcal{I}_{s'} - 1)$, and size r Dyson equations with weights $\lambda_{N,s}^\circ$. See [AEK19a, Section 11.5] for more details.

of $\vec{x}$. (Here we again use the equivalence between size $N-r$ Dyson equations with block sizes $(\mathcal{I}_s-1)\times(\mathcal{I}_{s'}-1)$ and size r Dyson equations with weights $\lambda^\circ_{N,s}$.) We have seen that both probability measures are supported in a fixed compact set of $\mathbb{R}$ (with probability $1-e^{-cN}$ in the former case; this compact set can be taken uniform over $\vec{x}$ in a compact set). This immediately upgrades convergence in probability within d_{BL} to $\mathbb{W}_2$. Finally, another application of Hoffman–Wielandt shows that the spectral distribution of a symmetric matrix $M \in \mathbb{R}^{d\times d}$ is a jointly 1-Lipschitz function of the entries, as a map from $\mathbb{R}^{d\times d} \to \mathbb{W}_2(\mathbb{R})$. In particular, it follows that

$$D \equiv \mathbb{W}_2(\widehat{\mu}_{H_N}(\boldsymbol{x}), \mu_{\xi,\vec{\lambda}^\circ_N}(\vec{x}))$$

is a 1-Lipschitz function of the entries of $\nabla^2_{\mathsf{sp}} H_N(\boldsymbol{x})$, and thus concentrates exponentially. We have seen that D converges in distribution to 0, hence its median $m(D)$ satisfies $|m(D)| \le \varepsilon/2$ for large N. Therefore $\mathbb{P}[|D| \le \varepsilon] \ge 1 - e^{-c(\varepsilon)N}$ for large N, yielding the desired $\mathbb{W}_2$ convergence claim.

Finally, we deduce convergence in $d_{\mathcal{H}}$. By adjusting ε, it remains to argue that with probability $1-e^{-c(\varepsilon)N}$, each $y \in \mathsf{supp}(\mu_{\xi,\vec{\lambda}^\circ_N}(\vec{x}))$ satisfies $d(y, \mathsf{supp}(\widehat{\mu}_{H_N}(\boldsymbol{x}))) \le 2\varepsilon$, as the opposite direction was shown earlier. We claim that $\mu_{\xi,\vec{\lambda}^\circ_N}(\vec{x})$ is "locally dense" in that for any $\varepsilon > 0$ there is $\delta > 0$ (independent of $\vec{x}$ within any given compact set) such that for all $y \in \mathsf{supp}(\mu_{\xi,\vec{\lambda}^\circ_N}(\vec{x}))$, we have

$$\mu_{\xi,\vec{\lambda}^\circ_N}(\vec{x})([y-\varepsilon, y+\varepsilon]) \ge \delta.$$

Indeed this assertion follows by [AEK19a, Theorem 2.6], which gives a local description of how the density for $\mu_{\xi,\vec{\lambda}^\circ_N}(\vec{x})$ behaves near its singularities. (In particular, the local scaling factor h_x therein is stated to be of constant order $h_x \sim 1$, with implicit constants depending only on norms of model parameters.) This completes the proof: if $y \in \mathsf{supp}(\mu_{\xi,\vec{\lambda}^\circ_N}(\vec{x}))$ satisfied $d(y, \mathsf{supp}(\widehat{\mu}_{H_N}(\boldsymbol{x}))) \ge 2\varepsilon$, we would directly obtain $\mathbb{W}_2\big(\widehat{\mu}_{H_N}(\boldsymbol{x}), \mu_{\xi,\vec{\lambda}^\circ_N}(\vec{x})\big) \ge \varepsilon\delta > 0$, but this $\mathbb{W}_2$ distance has been shown to tend to 0 with exponentially good probability. □

We also have continuity of vector Dyson equation solutions in the various parameters, which is needed to apply the results of [BBM23]. Sophisticated stability results for the Dyson equation were established for universality for random matrices in [AEK17a, AEK17b, AEK19a, AEK19b].

Proposition 4.2.11. *The map*

$$(\xi'', \vec{\lambda}, \vec{x}) \mapsto (\vec{m}(z), \mu(z))$$

is uniformly continuous on compact subsets of its domain. (That is, for a general symmetric matrix $\xi'' \in (0,\infty)^{r\times r}$, vector $\vec{\lambda} \in (0,1)^r$ with $\sum_s \lambda_s = 1$, and vector $\vec{x} \in \mathbb{R}^r$. We equip $\vec{m}$ with the compact-open topology and μ with the $\mathbb{W}_1$ distance.)

Proof. Suppose $(\xi''_n, \vec{\lambda}_n, \vec{x}_n)$ converge to $(\xi'', \vec{\lambda}, \vec{x})$ as $n \to \infty$. Let $\vec{m}$ be a subsequential limit of the corresponding Dyson equation solutions $\vec{m}_n$. Then $\vec{m}$ solves the limiting Dyson equation for $(\xi'', \vec{\lambda}, \vec{x})$ by continuity of the coefficients. Since the coefficients $(\xi''_n, \vec{\lambda}_n, \vec{x}_n)$ are uniformly bounded above and below, the supports of the corresponding spectral measures $\mu_{n,s}$ are uniformly bounded by [AEK19a]. Hence the imaginary parts $\Im(m_{n,s}(z))$ of their Stieltjes transforms are bounded below by $\Omega(\Im(z))$, uniformly on compact sets of $(\xi'', \vec{\lambda}, \vec{x}, z)$. In particular, the limit $\vec{m}$ is still a function from the strict upper half-plane $\mathbb{H}$ to itself. By uniqueness in Proposition 4.2.10, we find that $\vec{m}$ is *the* solution to the limiting Dyson equation for $(\xi'', \vec{\lambda}, \vec{x})$. Since $\vec{m}$ was an arbitrary subsequential limit, and the $\vec{m}_n$ are clearly tight, we find that $\lim_n \vec{m}_n = \vec{m}$, say uniformly on compact subsets of $\mathbb{H}$. Continuity of $\vec{m}$ follows; this is equivalent to continuity of μ_s and thus yields continuity of μ. □

In light of Proposition 4.2.10(a) and recalling (4.12), for any $\vec{\Delta} \in \{-1,1\}^r$ we define

$$S(\vec{\Delta}) = \mathsf{supp}(\mu(\vec{x}(\vec{\Delta}))). \tag{4.43}$$

Next we define

$$\Psi(\vec{x}) = \int \log|\gamma| \; [\mu(\vec{x})](\mathsf{d}\gamma)\,. \tag{4.44}$$

This will capture the exponential growth rate of

$$\mathbb{E}\left[\left|\det\left(\nabla^2_{\mathsf{sp}} H_N(\boldsymbol{\sigma})\right)\right| \,\Big|\, \nabla_{\mathsf{rad}} H_N(\boldsymbol{\sigma}) = \vec{x}\right]$$

which is the main term appearing in the Kac–Rice formula. We show its continuity in Proposition 4.2.15 below, using the following lemmas which will also be useful later.

Lemma 4.2.12. *For any* $(\xi, \vec{x})$ *and* $(\widetilde{\xi}, \widetilde{\vec{x}})$*, and some* $C = C(\vec{\lambda}) > 0$*,*

$$\mathbb{W}_\infty(\mu_\xi(\vec{x}), \mu_{\widetilde{\xi}}(\widetilde{\vec{x}})) \le C\big(\|\vec{x} - \widetilde{\vec{x}}\|_\infty + \|\xi'' - \widetilde{\xi}''\|_\infty^{1/2}\big). \tag{4.45}$$

Moreover for $C = C(\vec{\lambda}, \xi) > 0$ *independent of* $\vec{x}$*,*

$$\mathbb{W}_\infty\big(\mu(\vec{x}), \sum_{s \in \mathscr{S}} \lambda_s \delta_{-x_s/\sqrt{\lambda_s}}\big) \le C. \tag{4.46}$$

Proof. Let $\boldsymbol{G} = (g_{i,j})_{i,j \in \mathcal{T}} \in \mathbb{R}^{\mathcal{T} \times \mathcal{T}}$ be a GOE matrix with $\mathbb{E}[g_{i,j}^2] = (1 + \delta_{i,j})/N$. Let $\boldsymbol{W}, \widetilde{\boldsymbol{W}} \in \mathbb{R}^{\mathcal{T} \times \mathcal{T}}$ be defined by

$$W_{i,j} = \sqrt{\frac{\xi''_{s(i),s(j)}}{\lambda_{s(i)}\lambda_{s(j)}}}\, g_{i,j}, \qquad\qquad \widetilde{W}_{i,j} = \sqrt{\frac{\widetilde{\xi}''_{s(i),s(j)}}{\lambda_{s(i)}\lambda_{s(j)}}}\, g_{i,j},$$

and $\boldsymbol{M}, \widetilde{\boldsymbol{M}} \in \mathbb{R}^{\mathcal{T} \times \mathcal{T}}$ by $\boldsymbol{M} = \boldsymbol{W} - \mathrm{diag}(\Lambda^{-1/2}\vec{x} \diamond \mathbf{1}_{\mathcal{T}})$, $\widetilde{\boldsymbol{M}} = \widetilde{\boldsymbol{W}} - \mathrm{diag}(\Lambda^{-1/2}\widetilde{\vec{x}} \diamond \mathbf{1}_{\mathcal{T}})$. Then, by Proposition 4.2.5,

$$\mathbb{W}_\infty(\widehat{\mu}(\boldsymbol{M}), \widehat{\mu}(\widetilde{\boldsymbol{M}})) \le \|\boldsymbol{M} - \widetilde{\boldsymbol{M}}\|_{\mathsf{op}} \le \|\boldsymbol{W} - \widetilde{\boldsymbol{W}}\|_{\mathsf{op}} + \frac{\|\vec{x} - \widetilde{\vec{x}}\|_\infty}{\sqrt{\min \vec{\lambda}}}.$$

It is classical that $\|\boldsymbol{G}\|_{\mathsf{op}} \le 3$ with probability $1 - e^{-cN}$. By Slepian's lemma $\|\boldsymbol{W} - \widetilde{\boldsymbol{W}}\|_{\mathsf{op}}$ is stochastically dominated by

$$\frac{\|\xi'' - \widetilde{\xi}''\|_\infty^{1/2}}{\min \vec{\lambda}} \|\boldsymbol{G}\|_{\mathsf{op}},$$

so with probability $1 - e^{-cN}$, for suitable C,

$$\begin{aligned}\mathbb{W}_\infty(\widehat{\mu}(\boldsymbol{M}), \widehat{\mu}(\widetilde{\boldsymbol{M}})) \le \frac{3\|\xi'' - \widetilde{\xi}''\|_\infty^{1/2}}{\min \vec{\lambda}} + \frac{\|\vec{x} - \widetilde{\vec{x}}\|_\infty}{\sqrt{\min \vec{\lambda}}} &\le C\left(\|\vec{x} - \widetilde{\vec{x}}\|_\infty + \|\xi'' - \widetilde{\xi}''\|_\infty^{1/2}\right) \equiv \overline{C} \\ \implies \widehat{\mu}(\boldsymbol{M})([t + \overline{C}, \infty)) &\ge \widehat{\mu}(\widetilde{\boldsymbol{M}})([t + 2\overline{C}, \infty)), \quad \forall t \in \mathbb{R}.\end{aligned}$$

By Propositions 4.2.10(e) and 4.2.11, for any $\varepsilon > 0$, with probability $1 - e^{-cN}$

$$\mathbb{W}_2(\widehat{\mu}(\boldsymbol{M}), \mu_\xi(\vec{x})), \mathbb{W}_2(\widehat{\mu}(\widetilde{\boldsymbol{M}}), \mu_{\widetilde{\xi}}(\widetilde{\vec{x}})) \le \varepsilon \overline{C}.$$

In particular for any ε (depending on $\xi, \vec{x}, \tilde{\xi}, \tilde{\vec{x}}$) and $t \in \mathbb{R}$ we have

$$\mu_\xi(\vec{x})([t, \infty)) \ge \widehat{\mu}(\boldsymbol{M})([t + \overline{C}, \infty)) - \varepsilon, \qquad \widehat{\mu}(\widetilde{\boldsymbol{M}})([t + 2\overline{C}, \infty)) \ge \mu_{\widetilde{\xi}}(\widetilde{\vec{x}})([t + 3\overline{C}, \infty)) - \varepsilon.$$

Combining the above displays gives

$$\mu_\xi(\vec{x})([t, \infty)) \ge \mu_{\widetilde{\xi}}(\widetilde{\vec{x}})([t + 3\overline{C}, \infty)) - 2\varepsilon, \quad \forall \varepsilon > 0.$$

By similar reasoning the same inequality holds with $(\xi, \vec{x})$ and $(\widetilde{\xi}, \widetilde{\vec{x}})$ interchanged. This completes the proof (with $\overline{C}$ replaced by $3\overline{C}$) since ε is arbitrary. The second part (4.46) follows by similar reasoning since in the corresponding matrix model, the centered Gaussian contribution has spectral norm at most $C(\xi)$ with probability $1 - e^{-cN}$. $\square$

The following definition of distributions with bounded density and support will be convenient to ensure continuity of integrals against singular log potentials; it also reappears in Section 4.5.

Definition 4.2.13. The probability distribution $\mu \in \mathcal{P}(\mathbb{R})$ is C**-regular** if $\mathsf{supp}(\mu) \subseteq [-C, C]$ and μ has density at most C with respect to Lebesgue measure.

Lemma 4.2.14. *For any $C, \varepsilon > 0$ there exists $\delta > 0$ such that if $\mu, \widetilde{\mu}$ are C-regular and $\mathbb{W}_1(\mu, \widetilde{\mu}) \le \delta$ then*

$$\left| \int \log|\lambda| \mathsf{d}\mu(\lambda) - \int \log|\lambda| \mathsf{d}\widetilde{\mu}(\lambda) \right| \le \varepsilon.$$

Proof. Define the truncation $\log_K(x) = \min(K, \max(-K, \log x))$. It is easy to see that $\log_K(x)$ is L_K-Lipschitz for some constant L_K, so for $\delta \le \frac{\varepsilon}{2L_K}$ we have

$$\left| \int \log_K |\lambda| \mathsf{d}\mu(\lambda) - \int \log_K |\lambda| \mathsf{d}\widetilde{\mu}(\lambda) \right| \le L_K \cdot \mathbb{W}_1(\mu, \widetilde{\mu}) \le \varepsilon/2.$$

For $K \ge \log(C)$ and $|x| \le C$, we have

$$f_K(x) \equiv \log(|x|) - \log_K(|x|) = (K + \log(|x|)) \cdot 1_{|x| \le e^{-K}}.$$

C-regularity implies

$$\begin{aligned} \left| \int f_K(\lambda) \mathsf{d}\mu(\lambda) - \int f_K(\lambda) \mathsf{d}\widetilde{\mu}(\lambda) \right| &\le 2C \left| \int_{-e^{-K}}^{e^{-K}} K + \log|x| \ \mathsf{d}x \right| \\ &= -4C(x\log x - x + Kx)|_{x=0}^{e^{-K}} \\ &= 4Ce^{-K}. \end{aligned}$$

It remains to choose K so $4Ce^{-K} \le \varepsilon/2$ and then take $\delta \le \varepsilon/2L_K$ as above. □

Proposition 4.2.15. *$\Psi(\vec{x})$ is continuous in $(\xi'', \vec{\lambda}, \vec{x})$, uniformly on compact sets of $(\xi', \xi'', \vec{\lambda}, \vec{x})$ with ξ non-degenerate.*

Proof. This is immediate from Propositions 4.2.10(b) and 4.2.11, and Lemmas 4.2.12, 4.2.14. □

Organization The remainder of the paper is structured as follows. In Section 4.3 we determine the annealed complexity of critical points (Theorem 4.3.2). In Section 4.4 we solve the resulting variational problem, identifying the 2^r potential *types* of critical points for super-solvable ξ and showing that no others occur (Proposition 4.4.1). In Section 4.5 we connect Kac–Rice estimates to non-existence of approximate critical points (Theorem 4.5.2). In Section 4.6 we complete the proof of strong topological trivialization (Theorem 4.1.13) through the shrinking bands recursion explained in the introduction. In Section 4.7 we present further implications of Theorem 4.5.2 to approximate local maxima and marginal states in the single-species case (e.g. Corollary 4.7.6). Finally in Appendix 4.A we study solutions to the vector Dyson equation, obtaining joint 1/3-Hölder continuity (Theorem 4.A.2), a detailed characterization of the boundary behavior (e.g. Theorem 4.A.5 and Proposition 4.A.7), and an explicit formula for the main determinant term appearing in the Kac–Rice computation (Theorem 4.A.9).

4.3 Expected critical point counts

In this section we determine the annealed critical point statistics of H_N to leading exponential order.

4.3.1 Formula for the complexity functional

It will be crucial that only an exponentially small fraction of critical points in the annealed sense are atypical in the sense below, which closely resembles the definition of $\mathsf{Crt}_N^{\mathsf{good},\varepsilon}$.

Definition 4.3.1. We say $\boldsymbol{x} \in \mathcal{S}_N$ is respectively ε-energy-typical, ε-overlap-typical, and ε-bulk-typical if with $\vec{x} = \nabla_{\mathsf{rad}} H_N(\boldsymbol{x})$ it satisfies the three conditions (recall (4.11)):

$$\left|\frac{1}{N} H_N(\boldsymbol{x}) - (\xi')^\top A^{-1}\Lambda^{1/2}\vec{x}\right| \le \varepsilon,$$

$$\left\|\vec{R}(\boldsymbol{G}^{(1)},\boldsymbol{x}) - \Lambda^{-1/2}\mathrm{diag}(\Gamma^{(1)})A^{-1}\Lambda^{1/2}\vec{x}\right\|_\infty \le \varepsilon\,,$$

$$\mathbb{W}_2\left(\widehat{\mu}_{H_N}(\boldsymbol{x}), \mu(\vec{x})\right) \le \varepsilon \quad \text{and} \quad d_{\mathcal{H}}\left(\underset{H_N}{\mathsf{spec}}(\boldsymbol{x}), \mathsf{supp}(\mu(\vec{x}))\right) \le \varepsilon.$$

If these conditions are not satisfied, $\boldsymbol{x}$ is respectively ε-energy-atypical, ε-overlap-atypical, and ε-bulk-atypical. We say $\boldsymbol{x}$ is ε-typical if all three typicality conditions hold, and ε-atypical otherwise.

Given a set $\mathcal{D} \subseteq \mathbb{R}^r \times \mathbb{R}$, let $\mathsf{Crt}_N(\mathcal{D})$ denote the set of critical points $\boldsymbol{\sigma}$ for H_N with $(\nabla_{\mathsf{rad}} H_N(\boldsymbol{\sigma}), H_N(\boldsymbol{\sigma})/N) \in \mathcal{D}$. Also, for $\mathcal{D} \subseteq \mathbb{R}^r$, let $\mathsf{Crt}_N^{(\varepsilon)}(\mathcal{D})$ the set of critical points with $\nabla_{\mathsf{rad}} H_N(\boldsymbol{\sigma}) \in \mathcal{D}$ which are ε-atypical. Recalling (4.44), define the complexity functionals $F : \mathbb{R}^r \to \mathbb{R}$ and $F : \mathbb{R}^r \times \mathbb{R} \to \mathbb{R}$ by

$$F(\vec{x}) = \frac{1}{2}\Big(1 - \sum_{s\in\mathscr{S}} \lambda_s \log \xi^s(\vec{1}) - \|A^{-1/2}\Lambda^{1/2}\vec{x}\|_2^2\Big) + \Psi(\vec{x})\,, \tag{4.47}$$

$$F(\vec{x}, E) = F(\vec{x}) - \frac{(E - (\xi')^\top A^{-1}\Lambda^{1/2}\vec{x})^2}{2(\xi(\vec{1}) - (\xi')^\top A^{-1}\xi')}\,. \tag{4.48}$$

The following main result of this section characterizes the annealed critical point complexity of H_N in terms of these functionals.

Proposition 4.3.2. *Fix $\vec{\lambda}$ and non-degenerate ξ. Let $\mathcal{D} \subseteq \mathbb{R}^r \times \mathbb{R}$ be the closure of its non-empty interior. Then,*

$$\lim_{N\to\infty} \frac{1}{N} \log \mathbb{E}|\mathsf{Crt}_N(\mathcal{D})| = \sup_{(\vec{x},E)\in\mathcal{D}} F(\vec{x}, E). \tag{4.49}$$

Moreover, for $\mathcal{D} \subseteq \mathbb{R}^r$ equal to the closure of its non-empty interior and $\varepsilon > 0$, there exists $c = c(\xi, \varepsilon) > 0$ such that

$$\limsup_{N\to\infty} \frac{1}{N} \log \mathbb{E}|\mathsf{Crt}_N^{(\varepsilon)}(\mathcal{D})| \le \sup_{\vec{x}\in\mathcal{D}} F(\vec{x}) - c. \tag{4.50}$$

4.3.2 Proof of Proposition 4.3.2

Below, we focus on proving (4.49) and then explain the necessary changes to reach (4.50). For fixed $\vec{x} \in \mathbb{R}^r$, let $M_N = M_N(\vec{x}) \in \mathbb{R}^{\mathcal{T}\times\mathcal{T}}$ be a Gaussian matrix with distribution

$$M_N \sim \mathcal{L}(\nabla^2_{\mathsf{sp}} H_N(\boldsymbol{\sigma}) \,|\, \nabla_{\mathsf{rad}} H_N(\boldsymbol{\sigma}) = \vec{x}).$$

(Here $\mathcal{L}(\cdot|\cdot)$ denotes a conditional law; since $\nabla_{\mathsf{rad}} H_N(\boldsymbol{\sigma})$ is a linear function of H_N, there is no difficulty in defining regular conditional laws.) This can be written explicitly as follows. Let $\boldsymbol{W}$ be the random matrix with law given by (4.26). By Fact 4.1.4,

$$M_N \overset{d}{=} \boldsymbol{W} - \mathrm{diag}(\Lambda^{-1/2}\vec{x} \diamond \mathbf{1}_{\mathcal{T}})\,. \tag{4.51}$$

We let $\widehat{\mu}_{M_N} = \widehat{\mu}(M_N)$ denote the (random) spectral measure of this matrix. We further define the finite-N vector Dyson equation

$$1 + \left(z + \frac{x_s}{\sqrt{\lambda_s}} + \sum_{s'\in\mathscr{S}} \frac{\lambda^\circ_{N,s}\xi''_{s,s'}}{\lambda_s^2} m_{N,s'}(z)\right) m_{N,s}(z) = 0\,, \qquad z \in \mathbb{H}\,. \tag{4.52}$$

with unique solution $\vec{m}_N(z) = (m_{N,s}(z))_{s\in\mathscr{S}}$. (This is the matrix Dyson equation from e.g. [BBM23, Section 1.10] with $M(z) = \mathrm{diag}(\vec{m}_N(z) \diamond \mathbf{1}_{\mathcal{T}})$.) We let $\mu_{M_N,s}$ be the measure with Stieltjes transform $m_{N,s}$ and

$$\mu_{M_N} = \sum_{s\in\mathscr{S}} \lambda^\circ_{N,s}\mu_{M_N,s}.$$

This $\vec{m}_N(z)$ and μ_{M_N} exist and are unique by [EKS19, Proposition 5.1 (i),(ii)]. In the next proposition, we give the required conditions to apply [BBM23, Theorem 4.1].

Proposition 4.3.3. *Given $\vec{\lambda}, \delta$, a bounded family of $\vec{x} \in \mathbb{R}^r$ and a uniformly non-degenerate family of ξ, the following hold uniformly over the families for some $C, c > 0$:*

(a) $\mathbb{P}[\mathsf{supp}(\widehat{\mu}_{M_N}) \subseteq [-C, C]] \geq 1 - e^{-cN}$.

(b) $\mathbb{W}_1\big(\mathbb{E}[\widehat{\mu}_{M_N}], \mu_{M_N}\big) \leq N^{-c}$.

(c) $\mathbb{P}\left[\mathbb{W}_1\big(\widehat{\mu}_{M_N}, \mathbb{E}\widehat{\mu}_{M_N}\big) \geq \delta\right] \leq e^{-cN}$.

(d) $\lim_{N\to\infty} \mathbb{P}[\widehat{\mu}_{M_N}([-N^{-5}, N^{-5}]) > 0] = 0$.

(e) There exists an entrywise continuous-in-$\vec{x}$ coupling of the matrices $M_N(\vec{x})$.

(f) For all $\vec{x} \in \mathbb{R}^r$,

$$\mathbb{E}\left[|\det(M_N(\vec{x}))|\right] \leq C^N (\|\vec{x}\|_\infty + 1)^N.$$

Proof. Point (a) follows by Proposition 4.2.10(a)(e). Point (b) is a standard result on stability of the vector Dyson equation, see [BBM23, Proof of Corollary 1.9.B], except that (b) is usually shown for a variant of μ_{M_N} solving a Dyson equation with additional $O(1/N)$ terms on the diagonal entries $\xi''_{i,i}$. This discrepancy causes negligible error N^{-c} as shown in [BBM23, Proposition 3.1] and [BBM24, Lemma 3.1], so the claim does follow (see also [McK24, Section 3.1] for further discussion of this purely technical issue). By [GZ00, Lemma 1.2(b)], for any 1-Lipschitz test function f the map

$$(W_{i,j})_{1\leq i<j\leq k} \mapsto \int f(\lambda)\widehat{\mu}_{M_N}(\mathsf{d}\lambda)$$

is $O(1)$-Lipschitz. In particular, writing $W_{i,j} = \sqrt{(1+\delta_{i,j})\xi''_{s(i),s(j)}/\lambda_{s(i)}\lambda_{s(j)}}g_{i,j}$ for i.i.d. gaussians $g_{i,j}$, and applying gaussian concentration of measure,

$$\mathbb{P}\left[\left|\int f(\lambda)\widehat{\mu}_{M_N}(\mathsf{d}\lambda) - \int f(\lambda)\mathbb{E}\widehat{\mu}_{M_N}(\mathsf{d}\lambda)\right| \geq \delta'\right] \leq e^{-c(\delta')N}.$$

for any $\delta' > 0$. Point (c) then follows by union bounding over $O(1)$ test functions f. Point (d) follows by averaging over a small global shift of M_N by the identity matrix. Indeed since ξ is non-degenerate, we can express the law of M_N as the sum of two independent matrices, one of which is $N^{-2}gI_{\mathcal{T}}$ for a scalar Gaussian $g \sim \mathcal{N}(0,1)$. Then point (d) holds even after conditioning on the other summand, since each of the N eigenvalues has conditional probability $O(N^{-2})$ to lie in $[-N^{-5}, N^{-5}]$. Point (e) is clear. Finally point (f) easily follows from the deterministic inequality

$$|\det(M_N)| \leq \left(\frac{\|M_N\|_F^2}{N-r}\right)^{(N-r)/2}$$

which is a consequence of the arithmetic mean-geometric mean inequality. □

Proof of Proposition 4.3.2. Define

$$\begin{aligned}\Psi_N(\vec{x}) &= \frac{1}{N}\log\mathbb{E}|\det(M_N(\vec{x}))|,\\ F_N(\vec{x}) &= \frac{1}{2}\Big(1 - \sum_{s\in\mathscr{S}}\lambda_s\xi^s(\vec{1}) - \|A^{-1/2}\Lambda^{1/2}\vec{x}\|_2^2\Big) + \Psi_N(\vec{x})\,.\end{aligned}$$

Let φ_X be the density of random variable X w.r.t. Lebesgue measure. By the Kac–Rice formula (see e.g. [AT09, Chapter 11]),

$$\begin{aligned}\mathbb{E}|\mathsf{Crt}_N(\mathcal{D})| = \int_{\mathcal{S}_N}\int_{\mathbb{R}^r}\Bigg(\mathbb{E}\Big[|\det\nabla^2_{\mathsf{sp}}H_N(\boldsymbol{\sigma})|\mathbf{1}\{(\vec{x},H_N(\boldsymbol{\sigma})/N)\in\mathcal{D}\}\\ \Big|\ \nabla_{\mathsf{sp}}H_N(\boldsymbol{\sigma})=\mathbf{0},\nabla_{\mathsf{rad}}H_N(\boldsymbol{\sigma})=\vec{x}\Big]\\ \times\varphi_{\nabla_{\mathsf{sp}}H_N(\boldsymbol{\sigma})}(\mathbf{0})\varphi_{\nabla_{\mathsf{rad}}H_N(\boldsymbol{\sigma})}(\vec{x})\Bigg)\,\mathsf{d}\vec{x}\;\mathsf{d}\mathcal{H}^{N-r}(\boldsymbol{\sigma})\,,\end{aligned}$$

where $\mathcal{H}^{N-r}$ denotes the $(N-r)$-dimensional Hausdorff measure on $\mathcal{S}_N$. By spherical invariance, the integrand does not depend on $\boldsymbol{\sigma}$, so the integral over $\mathcal{H}^{N-r}$ simply contributes a volume factor given by Fact 4.2.3. By Fact 4.1.4 and Lemma 4.2.2,

$$\begin{aligned}&\mathbb{E}\Big[|\det\nabla^2_{\mathsf{sp}}H_N(\boldsymbol{\sigma})|\mathbf{1}\{(\vec{x},H_N(\boldsymbol{\sigma})/N)\in\mathcal{D}\}\ \Big|\ \nabla_{\mathsf{sp}}H_N(\boldsymbol{\sigma})=\mathbf{0},\nabla_{\mathsf{rad}}H_N(\boldsymbol{\sigma})=\vec{x}\Big]\\ &=\mathbb{E}\left[|\det M_N(\vec{x})|\right]\mathbb{P}\Big[(\vec{x},H_N(\boldsymbol{\sigma})/N)\in\mathcal{D}\ \Big|\ \nabla_{\mathsf{rad}}H_N(\boldsymbol{\sigma})=\vec{x}\Big] \qquad (4.53)\\ &=e^{o(N)}\mathbb{E}\left[|\det M_N(\vec{x})|\right]\int_{\mathbb{R}}\mathbf{1}\{(\vec{x},E)\in\mathcal{D}\}\\ &\qquad\times\exp\left(-\frac{N(E-(\xi')^\top A^{-1}\Lambda^{1/2}\vec{x})^2}{2(\xi(\vec{1})-(\xi')^\top A^{-1}\xi')}\right)\,\mathsf{d}E.\end{aligned}$$

We further have

$$\begin{aligned}&\varphi_{\nabla_{\mathsf{sp}}H_N(\boldsymbol{\sigma})}(\mathbf{0})=\prod_{s\in\mathscr{S}}(2\pi\xi^s(\vec{1}))^{-(|\mathcal{I}_s|-1)/2}\\ \implies\ &\frac{1}{N}\log\varphi_{\nabla_{\mathsf{sp}}H_N(\boldsymbol{\sigma})}(\mathbf{0})=-\frac{\log(2\pi)+\sum_{s\in\mathscr{S}}\lambda_s\log\xi^s(\vec{1})}{2}+o_N(1)\,. \qquad (4.54)\end{aligned}$$

and

$$\varphi_{\nabla_{\mathsf{rad}}H_N(\boldsymbol{\sigma})}(\vec{x})=(2\pi)^{-r/2}\sqrt{\frac{\det\Lambda}{\det A}}\exp\left(-\frac{N}{2}\|A^{-1/2}\Lambda^{1/2}\vec{x}\|_2^2\right).$$

Thus, up to additive $o_N(1)$ error, using Fact 4.2.3 in the second step gives

$$\begin{aligned}\frac{1}{N}\log\mathbb{E}|\mathsf{Crt}_N(\mathcal{D})|&\approx\frac{1}{N}\log\mathcal{H}^{N-r}(\mathcal{S}_N)+\frac{1}{N}\log\varphi_{\nabla_{\mathsf{sp}}H_N(\boldsymbol{\sigma})}(\mathbf{0})+\frac{1}{N}\log\int_{\mathcal{D}}\mathbb{E}|\det M_N(\vec{x})|\\ &\qquad\times\exp\left(-\frac{N}{2}\|A^{-1/2}\Lambda^{1/2}\vec{x}\|_2^2-\frac{N(E-(\xi')^\top A^{-1}\Lambda^{1/2}\vec{x})^2}{2(\xi(\vec{1})-(\xi')^\top A^{-1}\xi')}\right)\mathsf{d}(\vec{x},E)\\ &\approx\frac{1}{2}\Big(1-\sum_{s\in\mathscr{S}}\lambda_s\log\xi^s(\vec{1})\Big)+\frac{1}{N}\log\int_{\mathcal{D}}\exp(N\Psi_N(\vec{x})) \qquad (4.55)\\ &\qquad\times\exp\left(-\frac{N}{2}\|A^{-1/2}\Lambda^{1/2}\vec{x}\|_2^2-\frac{N(E-(\xi')^\top A^{-1}\Lambda^{1/2}\vec{x})^2}{2(\xi(\vec{1})-(\xi')^\top A^{-1}\xi')}\right)\mathsf{d}(\vec{x},E). \qquad (4.56)\end{aligned}$$

We next analyze the behavior of Ψ_N via [BBM23, Corollary 1.9.A], where the necessary conditions hold by Proposition 4.3.3. This result expresses the asymptotic value of Ψ_N based on the solution to the finite-N vector Dyson equation (4.52). In fact (4.52) is exactly equivalent to the limiting Dyson equation (4.38) with $\vec{\lambda}$ replaced by $\vec{\lambda}_N^\circ$ from (4.41) (see e.g. the discussion in [AEK19a, Section 11.5]). Therefore [BBM23, Corollary 1.9.A] shows that uniformly on compact sets, up to $o_N(1)$ error:

$$\Psi_N(\vec{x})\approx\int\log|\gamma|[\mu_{\vec{\lambda}_N^\circ}(\vec{x})](\mathsf{d}\gamma).$$

Recalling Propositions 4.2.11 and 4.2.15 shows that, again uniformly on compact sets:

$$\int\log|\gamma|[\mu_{\vec{\lambda}_N^\circ}(\vec{x})](\mathsf{d}\gamma)\approx\int\log|\gamma|[\mu(\vec{x})](\mathsf{d}\gamma)=\Psi(\vec{x}).$$

We next deduce (4.49) via Laplace's method similarly to [BBM23, Theorem 4.1]. (The latter result is not directly applicable as it is stated with additional technical requirements, but the proof can be routinely adapted.) For compact $\mathcal{D}$, local uniformity of the approximations just above allows replacement of $\Psi_N(\vec{x})$ by its limit $\Psi(\vec{x})$ in (4.55) up to $o_N(1)$ error. Since $\Psi(\vec{x})$ is continuous by Proposition 4.2.15, Laplace's method then immediately gives (4.49) for compact $\mathcal{D}$. It remains to show that (4.49) respects exhaustion by compact sets, both for finite N and after passing to the limit.

The needed statement at finite N is that

$$\lim_{R\to\infty} \limsup_{N\to\infty} \frac{1}{N} \log \mathbb{E}[|\mathsf{Crt}_N(\mathbb{R}^{r+1} \backslash \mathcal{D}_R)|] = -\infty \tag{4.57}$$

where $\mathcal{D}_R = (-R,R)^{r+1}$. Similarly to [BBM23, Lemma 4.3], this follows from the non-asymptotic bound $\mathbb{E}[|\det(M_N(\vec{x}))|] \le (C\max(\|\vec{x}\|_\infty,1))^N$, since when either $\|\vec{x}\|_\infty \ge R$ or $|E| \ge R$ the quadratic term (4.56) contributes an overwhelming $e^{-\Omega(NR^2)}$ factor. Said bound is shown exactly as in [McK24, Lemma 3.7], by writing $M_N(\vec{x}) = W_N + A_N(\vec{x})$ for deterministic A_N and centered Gaussian W_N. Namely one can separate $M_N(\vec{x})$ with the deterministic estimate $|\det(M_N(\vec{x})|^N \le 2^N(\|W_N\|_{\mathsf{op}}^N + \|A_N(\vec{x})\|_{\mathsf{op}}^N)$, and use the simple bound $\mathbb{P}[\|W_N\|_{\mathsf{op}} \ge t] \le e^{-cN(t-C)_+}$ (which follows because $\|W_N\|_{\mathsf{op}}$ is typically $O(1)$ and is $O(1)$-Lipschitz in its independent Gaussian entries) to control the random part.

We also need to show that $F(\vec{x},E)$ tends to $-\infty$ as $\max(\|\vec{x}\|_\infty,|E|)\to\infty$. This follows because $\Psi(\vec{x}) \le C(\max(\|\vec{x}\|_\infty,1))$ due to (4.46), which is dominated by the quadratic terms of $F(\vec{x},E)$. Together with (4.57), this allows us to deduce (4.49) for general $\mathcal{D}$ from the compact $\mathcal{D}$ case via exhaustion. Namely one restricts to the compact set $\mathcal{D}\cap[-R,R]^{r+1}$ and sends $R\to\infty$ after $N\to\infty$ (exactly as in e.g. [BBM23, Proof of Theorem 4.1]). This completes the first part of the proof.

Moving onto (4.50), we separately address the cases of energy, overlap, and bulk-atypicality. Energy-atypicality follows directly from (4.49), as the term involving E in (4.48) is nonzero. For overlap-atypicality, let $E^{\mathrm{o}}_\varepsilon(\boldsymbol{\sigma})$ denote the event that $\boldsymbol{\sigma}$ is ε-overlap-atypical, and $\mathsf{Crt}_N^{(\varepsilon,\mathrm{o})}(\mathcal{D})$ be the set of critical points with $\nabla_{\mathsf{rad}} H_N(\boldsymbol{\sigma}) \in \mathcal{D}$ which are ε-overlap-atypical. (Recall that for (4.50), $\mathcal{D}$ is a subset of $\mathbb{R}^r$ rather than $\mathbb{R}^r\times\mathbb{R}$.) By the Kac–Rice formula,

$$\mathbb{E}|\mathsf{Crt}_N^{(\varepsilon,\mathrm{o})}(\mathcal{D})| = \int_{\mathcal{S}_N}\int_{\mathcal{D}} \Bigg(\mathbb{E}\Big[|\det \nabla^2_{\mathsf{sp}} H_N(\boldsymbol{\sigma})| \mathbf{1}\{E^{\mathrm{o}}_\varepsilon(\boldsymbol{\sigma})\} \;\Big|\; \nabla_{\mathsf{sp}} H_N(\boldsymbol{\sigma}) = \mathbf{0}, \nabla_{\mathsf{rad}} H_N(\boldsymbol{\sigma}) = \vec{x}\Big] \\ \varphi_{\nabla_{\mathsf{sp}} H_N(\boldsymbol{\sigma})}(\mathbf{0}) \varphi_{\nabla_{\mathsf{rad}} H_N(\boldsymbol{\sigma})}(\vec{x}) \Bigg) \, \mathsf{d}\vec{x} \; \mathsf{d}\mathcal{H}^{N-r}(\boldsymbol{\sigma}).$$

By calculations similar to above, up to additive $o_N(1)$ error

$$\begin{aligned} &\frac{1}{N}\log \mathbb{E}|\mathsf{Crt}_N^{(\varepsilon,\mathrm{o})}(\mathcal{D})| \\ &\approx \frac{1}{2}\left(1 - \sum_{s\in\mathscr{S}} \lambda_s \log \xi^s(\vec{1})\right) + \frac{1}{N}\log\int_{\mathcal{D}} \exp\left(N\Psi(\vec{x}) - \frac{N}{2}\|A^{-1/2}\Lambda^{1/2}\vec{x}\|_2^2\right) \\ &\qquad\qquad \times \mathbb{P}\left[E^{\mathrm{o}}_\varepsilon(\boldsymbol{\sigma}) \;\Big|\; \nabla_{\mathsf{rad}} H_N(\boldsymbol{\sigma}) = \vec{x}\right] \mathsf{d}\vec{x}. \end{aligned}$$

By (4.28), each entry of $\vec{R}(\boldsymbol{G}^{(1)},\boldsymbol{\sigma})$ has variance bounded above by $1/(N\min\vec{\lambda})$. Since $\vec{R}(\boldsymbol{G}^{(1)},\boldsymbol{\sigma})$ and $\nabla_{\mathsf{rad}} H_N(\boldsymbol{\sigma})$ are jointly gaussian, this remains true after conditioning on $\nabla_{\mathsf{rad}} H_N(\boldsymbol{\sigma})$. In light of (4.33), this implies

$$\mathbb{P}\left[E^{\mathrm{o}}_\varepsilon(\boldsymbol{\sigma}) \;\Big|\; \nabla_{\mathsf{rad}} H_N(\boldsymbol{\sigma}) = \vec{x}\right] \le e^{-cN}.$$

This implies (4.50) for overlap-typicality.

The main case that needs to be addressed is bulk-atypicality. Let $E^{\mathrm{b}}_\varepsilon(\boldsymbol{\sigma})$ be the event that $\boldsymbol{\sigma}$ is ε-bulk-atypical and $\mathsf{Crt}_N^{(\varepsilon,\mathrm{b})}(\mathcal{D})$ be the set of critical points with $\nabla_{\mathsf{rad}} H_N(\boldsymbol{\sigma}) \in \mathcal{D}$ which are ε-bulk-atypical. Similarly to above,

$$\begin{aligned}
&\mathbb{E}|\mathsf{Crt}_N^{(\varepsilon,\mathrm{b})}(\mathcal{D})| \\
&= \int_{\mathcal{S}_N}\int_{\mathcal{D}}\Bigg(\mathbb{E}\Big[|\det\nabla^2_{\mathsf{sp}}H_N(\boldsymbol{\sigma})|\mathbf{1}\{E^{\mathrm{b}}_\varepsilon(\boldsymbol{\sigma})\}\;\Big|\;\nabla_{\mathsf{sp}}H_N(\boldsymbol{\sigma})=\mathbf{0},\nabla_{\mathsf{rad}}H_N(\boldsymbol{\sigma})=\vec{x}\Big] \\
&\qquad\qquad \times\varphi_{\nabla_{\mathsf{sp}}H_N(\boldsymbol{\sigma})}(\mathbf{0})\varphi_{\nabla_{\mathsf{rad}}H_N(\boldsymbol{\sigma})}(\vec{x})\Bigg)\,\mathsf{d}\vec{x}\,\mathsf{d}\mathcal{H}^{N-r}(\boldsymbol{\sigma}),
\end{aligned}$$

and so up to $o_N(1)$ additive error

$$\begin{aligned}
\frac{1}{N}\log\mathbb{E}|\mathsf{Crt}_N^{(\varepsilon,\mathrm{b})}(\mathcal{D})| \approx \frac{1}{2}\left(1-\sum_{s\in\mathscr{S}}\lambda_s\log\xi^s(\vec{1})\right)+\frac{1}{N}\log\int_{\mathcal{D}}\exp\left(-\frac{N}{2}\|A^{-1/2}\Lambda^{1/2}\vec{x}\|_2^2\right) \\
\times\,\mathbb{E}\left[|\det\nabla^2_{\mathsf{sp}}H_N(\boldsymbol{\sigma})|\mathbf{1}\{E^{\mathrm{b}}_\varepsilon(\boldsymbol{\sigma})\}\;\Big|\;\nabla_{\mathsf{rad}}H_N(\boldsymbol{\sigma})=\vec{x}\right]\mathsf{d}\vec{x}.
\end{aligned}$$

Unlike above, $\mathbf{1}\{E^{\mathrm{b}}_\varepsilon(\boldsymbol{\sigma})\}$ and the Hessian determinant are not independent. Instead, by Cauchy–Schwarz, the last expectation is bounded by

$$\mathbb{E}\left[|\det\nabla^2_{\mathsf{sp}}H_N(\boldsymbol{\sigma})|^2\;\Big|\;\nabla_{\mathsf{rad}}H_N(\boldsymbol{\sigma})=\vec{x}\right]^{1/2}\mathbb{P}\left[E^{\mathrm{b}}_\varepsilon(\boldsymbol{\sigma})\;\Big|\;\nabla_{\mathsf{rad}}H_N(\boldsymbol{\sigma})=\vec{x}\right]^{1/2}.$$

By [BBM23, Theorem A.2],

$$\begin{aligned}
&\mathbb{E}\left[|\det\nabla^2_{\mathsf{sp}}H_N(\boldsymbol{\sigma})|^2\;\Big|\;\nabla_{\mathsf{rad}}H_N(\boldsymbol{\sigma})=\vec{x}\right]^{1/2} \\
&= e^{o(N)}\mathbb{E}\left[|\det\nabla^2_{\mathsf{sp}}H_N(\boldsymbol{\sigma})|\;\Big|\;\nabla_{\mathsf{rad}}H_N(\boldsymbol{\sigma})=\vec{x}\right]=e^{o(N)}\exp(N\Psi_N(\vec{x})),
\end{aligned}$$

and by Proposition 4.2.10(e) and 4.2.11,

$$\mathbb{P}\left[E^{\mathrm{b}}_\varepsilon(\boldsymbol{\sigma})\;\Big|\;\nabla_{\mathsf{rad}}H_N(\boldsymbol{\sigma})=\vec{x}\right]^{1/2}\le e^{-cN/2}.$$

Combining and arguing as above completes the proof (with $c/2$ in place of c). □

4.4 Solving the variational problem

Due to Proposition 4.3.2, in order to establish Theorem 4.1.11 it remains to maximize $F(\vec{x})$ over $\mathbb{R}^r$. In this section we prove parts (a) and (b) of this theorem, regarding super-solvable ξ. The strictly sub-solvable case (c) will be proved in Subsection 4.5.3.

Proposition 4.4.1. *Assume ξ is strictly super-solvable. Then $F(\vec{x})\le 0$ for all $\vec{x}\in\mathbb{R}^r$, with equality at precisely the 2^r points $\vec{x}(\vec{\Delta})$ for $\vec{\Delta}\in\{-1,1\}^r$.*

We also state the following regularity property of F, which implies that its maximum is attained at a stationary point.

Lemma 4.4.2. *The function F is continuously differentiable in $\vec{x}$, and*

$$\lim_{R\to\infty}\sup_{\|\vec{x}\|_\infty\ge R}F(\vec{x})=-\infty\,.$$

Moreover the latter limit is uniform on bounded, uniformly non-degenerate ξ.

Proof. Continuous differentiability follows from Lemma 4.4.4 below and Lemma 4.2.12. For $R=\|\vec{x}\|_\infty$, we have $\Psi(\vec{x})\lesssim\log R$ while $\langle\Lambda^{1/2}\vec{x},A^{-1}\Lambda^{1/2}\vec{x}\rangle\gtrsim R^2$, which establishes the decay at infinity. □

Proof of Theorem 4.1.11 parts (a), (b). The result is immediate from Propositions 4.3.2 and 4.4.1 for strictly super-solvable ξ. The proof of (4.17) for solvable ξ follows since $\xi \mapsto \sup_{\vec{x}\in\mathbb{R}} F_\xi(\vec{x})$ is continuous at any non-degenerate ξ. Indeed F is locally uniformly continuous in non-degenerate ξ on compact $\vec{x}$-sets by Proposition 4.2.15. $\square$

We also prove the following fact which will be useful in later sections.

Lemma 4.4.3. *If ξ is strictly super-solvable, there exists $\varepsilon > 0$ such that for any $\vec{\Delta} \in \{-1,1\}^r$, $[-\varepsilon,\varepsilon] \cap S(\vec{\Delta}) = \emptyset$.*

4.4.1 Stationarity condition

We next identify all stationary points of F. It will be convenient to perform the below derivative calculations in the variable $\vec{v} = \Lambda^{1/2}\vec{x}$ (recall (4.11)). To this end, we define:

$$\begin{aligned}
\overline{F}(\vec{v}) &= F(\vec{x}),\\
\overline{\Psi}(\vec{v}) &= \Psi(\vec{x}),\\
\overline{\mu}(\vec{v}) &= \mu(\vec{x}),\\
u_s(z;\vec{v}) &= m_s(z;\vec{x}),\\
\vec{u}(z;\vec{v}) &= \vec{m}(z;\vec{x}).
\end{aligned}$$

Here we recall $m_s(z;\vec{x})$ is defined above (4.38) and define $\vec{m}(z;\vec{x}) = (m_s(z;\vec{x}))_{s\in\mathscr{S}}$. Thus,

$$\overline{F}(\vec{v}) = \frac{1}{2}\Big(1 - \sum_{s\in\mathscr{S}} \lambda_s \log \xi^s(\vec{1}) - \|A^{-1/2}\vec{v}\|_2^2\Big) + \overline{\Psi}(\vec{v})\,. \tag{4.58}$$

The Dyson equation (4.38) is equivalent (after some rearrangement) to

$$\lambda_s z + v_s = -\frac{\lambda_s}{u_s(z;\vec{v})} - \sum_{s'\in\mathscr{S}} \xi''_{s,s'} u_{s'}(z;\vec{v})\,. \tag{4.59}$$

The next lemma gives exact formulas for $\overline{\Psi}, \overline{F}$ and their gradients. These seem to be new and extend known results in the single-species case (see e.g. (4.112) and the discussion below). We believe they are of independent interest, and might lead to more explicit thresholds in e.g. [McK24, Theorem 2.5]. (This would still require optimization over the complicated set of vectors $\vec{u}$ corresponding to some $\vec{v} \in \mathbb{R}^r$; see Lemma 4.4.8 below.) The majority of the proof is carried out in Appendix 4.A.

Note that below and throughout, we always use $\langle \vec{a}, \vec{b}\rangle = \sum_{s=1}^r a_s b_s$ to denote a bilinear form rather than a complex inner product, even when $\vec{a}, \vec{b}$ are complex vectors. Also recall that $\Re(\cdot)$ denotes the real part of a complex number or vector.

Lemma 4.4.4. *The functions $\overline{\Psi}, \overline{F}$ are C^1 and satisfy, with $\vec{u} = \vec{u}(0;\vec{v})$,*

$$\overline{\Psi}(\vec{v}) = \frac{1}{2}\Re(\langle \vec{u}, \xi''\vec{u}\rangle) - \sum_{s\in\mathscr{S}} \lambda_s \log|u_s|, \tag{4.60}$$

$$\overline{F}(\vec{v}) = \frac{1}{2}\left(1 - \sum_{s\in\mathscr{S}} \lambda_s \log \xi^s(\vec{1}) - \langle \vec{v}, A^{-1}\vec{v}\rangle + \Re(\langle \vec{u}, \xi''\vec{u}\rangle)\right) - \sum_{s\in\mathscr{S}} \lambda_s \log|u_s|, \tag{4.61}$$

$$\nabla\overline{\Psi}(\vec{v}) = -\Re(\vec{u}), \tag{4.62}$$

$$\nabla\overline{F}(\vec{v}) = -A^{-1}\vec{v} - \Re(\vec{u}). \tag{4.63}$$

Proof. The formulas (4.60) and (4.62) follow from Theorem 4.A.9 and Lemma 4.A.28. Then (4.61) and (4.63) follow as straightforward consequences. $\square$

For the rest of this section, we let $\vec{u} = \vec{u}(0;\vec{v}) \in \overline{\mathbb{H}}^r$. Note that (4.59), specialized to $z = 0$, gives

$$v_s = -\frac{\lambda_s}{u_s} - \sum_{s' \in \mathscr{S}} \xi''_{s,s'} u_{s'} \,. \tag{4.64}$$

We next describe the condition for $\vec{v}$ to be a stationary point of $\overline{F}$.

Lemma 4.4.5. *If $\nabla \overline{F}(\vec{v}) = \vec{0}$, then for all $s \in \mathscr{S}$, either $\Re(u_s) = 0$ or $|u_s| = \sqrt{\lambda_s/\xi'_s} = 1/\sqrt{\xi^s(\vec{1})}$ (recall (4.5)).*

Proof. We have

$$\frac{\lambda_s}{u_s} + \sum_{s' \in \mathscr{S}} \xi''_{s,s'} u_{s'} = -v_s = \Re(A\vec{u})_s = \Re\left(\xi'_s u_s + \sum_{s' \in \mathscr{S}} \xi''_{s,s'} u_{s'}\right),$$

where the first equality is (4.64), the second is (4.63), and the third is the definition (4.11) of A. Taking real parts of both sides implies $\Re(\lambda_s/u_s) = \Re(\xi'_s u_s)$, which implies the conclusion. □

We will see that the maximizers of F described in Proposition 4.4.1 correspond to $u_s = \pm 1/\sqrt{\xi^s(\vec{1})}$. The primary remaining difficulty is to show all *other* remaining stationary points are *not* local maxima.

4.4.2 Non-maximality of stationary points with pure-imaginary u_s

The following main result of this subsection rules out the first case identified in Lemma 4.4.5 for strictly super-solvable ξ. From it, we will easily conclude (in Corollary 4.4.11 below) that all maximizers of F are as in Proposition 4.4.1.

Proposition 4.4.6. *Suppose ξ is strictly super-solvable. If $\nabla \overline{F}(\vec{v}) = \vec{0}$ and $\Re(u_s) = 0$ for some $s \in \mathscr{S}$, then $\vec{v}$ is not a local maximum of $\overline{F}$.*

We will need as input from Appendix 4.A the following two lemmas. For $\vec{u} \in \overline{\mathbb{H}}^r$ define the matrices

$$M(\vec{u}) = \operatorname{diag}\left(\frac{\lambda_s}{u_s^2}\right)_{s \in \mathscr{S}} - \xi'', \qquad \overline{M}(\vec{u}) = \operatorname{diag}\left(\frac{\lambda_s}{|u_s|^2}\right)_{s \in \mathscr{S}} - \xi''. \tag{4.65}$$

Lemma 4.4.7. *At all $\vec{v} \in \mathbb{R}^r$ such that $M(\vec{u})$ is invertible, the function $\vec{v} \mapsto \vec{u}(0;\vec{v})$ is differentiable and $\nabla_{\vec{v}} \vec{u}(0;\vec{v}) = M(\vec{u})^{-1}$.*

Proof. Follows from Lemma 4.A.4. □

The equation (4.64) relates $\vec{v}$ to its associated $\vec{u}$, which is well-defined by Proposition 4.2.10. Because $\vec{v} \in \mathbb{R}^r$ while $\vec{u} \in \overline{\mathbb{H}}^r$, one roughly expects that those $\vec{u}$ corresponding to some $\vec{v} \in \mathbb{R}^r$ lie within an r-dimensional real submanifold of $\overline{\mathbb{H}}^r$. The next lemma describes this set of $\vec{u}$.

Lemma 4.4.8. *Let $\vec{u}^* \in \overline{\mathbb{H}}^r$. There exists $\vec{v} \in \mathbb{R}^r$ such that $\vec{u}^* = \vec{u}(0;\vec{v})$ if and only if one of the following conditions holds.*

(i) $\vec{u}^* \in \mathbb{R}^r$ and $M(\vec{u}^*) \succeq 0$.

(ii) $\vec{u}^* \in \mathbb{H}^r$, $\overline{M}(\vec{u}^*) \succeq 0$, and $\overline{M}(\vec{u}^*)\Im(\vec{u}^*) = 0$.

Moreover, in case (ii), $M(\vec{u}^*)$ is invertible.

Proof. Follows from Theorem 4.A.5(b) and Corollary 4.A.6. □

Lemma 4.4.9. *Suppose ξ is strictly super-solvable. If $\overline{M}(\vec{u})$ is singular, then $|u_s| > 1/\sqrt{\xi^s(\vec{1})}$ for some $s \in \mathscr{S}$.*

Proof. Suppose otherwise; then $\operatorname{diag}(\lambda_s/|u_s|^2)_{s\in\mathscr{S}} \succeq \operatorname{diag}(\xi')$, so

$$\overline{M}(\vec{u}) \succeq \operatorname{diag}(\xi') - \xi'' \succ 0\,.$$

However, $\overline{M}(\vec{u})$ is singular, contradiction. $\square$

Further define

$$\widehat{M}(\vec{u}) = \operatorname{diag}\left(\frac{\lambda_s}{|u_s|^2}\right)_{s\in\mathscr{S}} + \xi''.$$

Lemma 4.4.10. *If $\overline{M}(\vec{u}) \succeq 0$, then $\widehat{M}(\vec{u}) \succ 0$.*

Proof. Let $D_{\xi''}$ be the diagonal matrix with (s,s) entry $\xi''_{s,s}$. We will apply Lemma 4.2.8 with matrices $\widehat{M}(\vec{u}) - 2D_{\xi''}$ and $\overline{M}(\vec{u})$. Note that the diagonal entries of both matrices coincide, as

$$(\widehat{M}(\vec{u}) - 2D_{\xi''})_{s,s} = \overline{M}(\vec{u})_{s,s} = \frac{\lambda_s}{|u_s|^2} - \xi''_{s,s}$$

and the off-diagonal entries are related by

$$(\widehat{M}(\vec{u}) - 2D_{\xi''})_{s,s'} = \xi''_{s,s'} = |\overline{M}(\vec{u})_{s,s'}|.$$

Lemma 4.2.8 thus implies $\widehat{M}(\vec{u}) - 2D_{\xi''} \succeq 0$, which by Assumption 4.1.1 implies the result. $\square$

Proof of Proposition 4.4.6. The hypothesis $\Re(u_s) = 0$ for some s implies that $\vec{u} \notin \mathbb{R}^r$, since $u_s \neq 0$ by Proposition 4.2.10. Therefore Lemma 4.4.8 case (ii) applies, so $\overline{M}(\vec{u}) \succeq 0$ is singular and $M(\vec{u})$ is invertible. Differentiation of (4.63) using Lemma 4.4.7 then gives

$$\nabla^2 \overline{F}(\vec{v}) = -A^{-1} - \Re(M(\vec{u})^{-1})\,. \tag{4.66}$$

Let $I \subseteq [r]$ be the set of indices s with $|u_s| \neq 1/\sqrt{\xi^s(\vec{1})}$, which is nonempty by Lemma 4.4.9. By Lemma 4.4.5, we have $\Re(u_s) = 0$ for all $s \in I$. Moreover, Lemma 4.4.10 implies $\widehat{M}(\vec{u}) \succ 0$. We will construct a vector $\vec{w} \in \mathbb{R}^r$ such that $\vec{w}^\top(\nabla^2\overline{F}(\vec{v}))\vec{w} > 0$, which implies $\vec{v}$ is not a local maximum. We work in the subspace $\mathbb{R}^I \subseteq \mathbb{R}^r$, consider $\vec{a} \in \mathbb{R}^I$ to be chosen later, and set

$$\vec{w} = -M(\vec{u})\vec{a} = \widehat{M}(\vec{u})\vec{a}.$$

Here the second equality uses that $\vec{a} \in \mathbb{R}^I$ and that u_s is pure imaginary for $s \in I$. Importantly, all entries of $\vec{w}$ are real.

Abbreviate $M = M(\vec{u})$, $\overline{M} = \overline{M}(\vec{u})$, $\widehat{M} = \widehat{M}(\vec{u})$ and let $D = A - \widehat{M}$. Note that D is diagonal and its entry $D_{s,s} = \xi'_s - \frac{\lambda_s}{|u_s|^2}$ is nonzero if and only if $s \in I$. Let $D^\dagger$ denote the Moore–Penrose inverse of M and $P_I = \sum_{s\in I} \vec{e}_s\vec{e}_s^\top$ be the projection onto I. Then

$$\begin{aligned}
\vec{w}^\top(\nabla^2\overline{F}(\vec{v}))\vec{w} &= -\vec{w}^\top A^{-1}\vec{w} - \Re(\vec{w}^\top M^{-1}\vec{w}) \\
&= -\vec{a}^\top \widehat{M} A^{-1} \widehat{M}\vec{a} - \Re(\vec{a}^\top M \vec{a}) \\
&= -\vec{a}^\top \widehat{M} A^{-1} \widehat{M}\vec{a} + \vec{a}^\top \widehat{M}\vec{a} \\
&= -\vec{a}^\top (A-D) A^{-1} (A-D)\vec{a} + \vec{a}^\top (A-D)\vec{a} \\
&= \vec{a}^\top D(D^\dagger - A^{-1})D\vec{a} \\
&= \vec{a}^\top DA^{-1}\left((\widehat{M}+D)D^\dagger(\widehat{M}+D) - (\widehat{M}+D)\right)A^{-1}D\vec{a} \\
&= \vec{a}^\top DA^{-1}\left(\widehat{M}D^\dagger\widehat{M} + \widehat{M}P_I + P_I\widehat{M} - \widehat{M}\right)A^{-1}D\vec{a}\,.
\end{aligned} \tag{4.67}$$

Recall that $A, \widehat{M}$ agree on rows indexed by $[r]\setminus I$. So, if $\vec{y} \in \mathbb{R}^I$ and $\vec{z} = A^{-1}\vec{y}$, then for any $s \in \mathscr{S}\setminus I$,

$$(\widehat{M}\vec{z})_s = (A\vec{z})_s = y_s = 0.$$

Thus $\mathbb{R}^I$ is an invariant subspace for $\widehat{M}A^{-1}$, i.e. $\widehat{M}A^{-1}\mathbb{R}^I \subseteq \mathbb{R}^I$. Since A and $\widehat{M}$ are both full rank (by Corollary 4.2.9 and Lemma 4.4.10), in fact $\widehat{M}A^{-1}\mathbb{R}^I = \mathbb{R}^I$ is a bijection on $\mathbb{R}^I$, and the same holds for $\widehat{M}A^{-1}D$.

Since we showed earlier in this proof that $\overline{M}$ is singular, Lemma 4.4.9 implies that there exists $s \in I$ such that $D_{s,s} > 0$. Using the bijectivity just established, we choose $\vec{a}$ such that $\widehat{M}A^{-1}D\vec{a} = \vec{e}_s$. Then

$$\vec{a}^\top DA^{-1}\widehat{M}D^\dagger\widehat{M}A^{-1}D\vec{a} = \vec{e}_s^\top D^\dagger \vec{e}_s = D^\dagger_{s,s} > 0\,.$$

Since $s \in I$ we further have

$$\begin{aligned}\vec{a}^\top DA^{-1}\left(\widehat{M}P_I + P_I\widehat{M} - \widehat{M}\right)A^{-1}D\vec{a} &= \vec{e}_s^\top P_I\widehat{M}^{-1}\vec{e}_s + \vec{e}_s^\top\widehat{M}^{-1}P_I\vec{e}_s - \vec{e}_s^\top\widehat{M}^{-1}\vec{e}_s \\ &= \vec{e}_s^\top\widehat{M}^{-1}\vec{e}_s > 0\,.\end{aligned}$$

Summing and recalling (4.67), we conclude that $\vec{w}^\top(\nabla^2\overline{F}(\vec{v}))\vec{w} > 0$ as desired. □

Corollary 4.4.11. *If $\vec{v} \in \mathbb{R}^r$ maximizes $\overline{F}$, then $\vec{v} = \vec{v}(\vec{\Delta}) \equiv \Lambda^{1/2}\vec{x}(\vec{\Delta})$ for some $\vec{\Delta} \in \{-1,1\}^r$. Conversely, each $\vec{v}(\vec{\Delta})$ is a stationary point of $\overline{F}$ with $\overline{F}(\vec{v}(\vec{\Delta})) = 0$.*

Proof. Lemma 4.4.2 implies that if $\vec{v}$ maximizes $\overline{F}$ then it is a stationary point. Lemma 4.4.5 and Proposition 4.4.6 imply that $|u_s| = 1/\sqrt{\xi^s(\vec{1})}$ for all $s \in \mathscr{S}$. Thus $\overline{M}(\vec{u}) = \mathrm{diag}(\xi') - \xi'' \succ 0$ is not singular. By Lemma 4.4.8, we have $\vec{u} \in \mathbb{R}^r$, so $u_s = \pm 1/\sqrt{\xi^s(\vec{1})}$. The 2^r possible choices of $\vec{u}$ are indexed by $\vec{\Delta} \in \{-1,1\}^r$ and given by

$$u(\vec{\Delta})_s = -\Delta_s/\sqrt{\xi^s(\vec{1})}\,. \tag{4.68}$$

Substituting $\vec{u}(\vec{\Delta})$ into (4.64) shows that $\vec{v} = \vec{v}(\vec{\Delta})$. For the converse, note that $M(\vec{u}(\vec{\Delta})) = \mathrm{diag}(\xi') - \xi'' \succ 0$, so Lemma 4.4.8 case (i) implies that $\vec{u}(\vec{\Delta}) = \vec{u}(0;\vec{v}(\vec{\Delta}))$. We can verify from the formulas for $\vec{u}(\vec{\Delta})$ and $\vec{v}(\vec{\Delta})$ that

$$\vec{v}(\vec{\Delta}) = -A\vec{u}(\vec{\Delta})\,. \tag{4.69}$$

Thus, by Lemma 4.4.5,

$$\nabla\overline{F}(\vec{v}(\vec{\Delta})) = -A^{-1}\vec{v}(\vec{\Delta}) - \Re(\vec{u}(\vec{\Delta})) = 0\,,$$

so $\vec{v}(\vec{\Delta})$ is a stationary point.

Finally, we verify that $\overline{F}(\vec{v}(\vec{\Delta})) = 0$ for all $\vec{\Delta}$ by directly using (4.61). Since $\vec{u}(\vec{\Delta}) \in \mathbb{R}^r$, the quadratic terms combine to give:

$$\begin{aligned}\frac{1}{2}\left(\langle\vec{u}(\vec{\Delta}), \xi''\vec{u}(\vec{\Delta})\rangle - \langle\vec{v}(\vec{\Delta}), A^{-1}\vec{v}(\vec{\Delta})\rangle\right) &\overset{(4.69)}{=} \langle\vec{u}(\vec{\Delta}), (\xi'' - A)\vec{u}(\vec{\Delta})\rangle/2 \\ &\overset{(4.11)}{=} -\langle\vec{u}(\vec{\Delta}), \mathrm{diag}(\xi')\vec{u}(\vec{\Delta})\rangle/2 \\ &\overset{(4.68)}{=} -1/2.\end{aligned}$$

This cancels the first term in (4.61). Meanwhile recalling (4.68), the logarithmic terms give

$$-\sum_{s\in\mathscr{S}}\lambda_s\log\left(|u_s|\sqrt{\xi^s(\vec{1})}\right) = 0.$$

Combining completes the proof. □

Proof of Proposition 4.4.1. Lemma 4.4.2 implies that $\overline{F}$ possesses at least one global maximizer. The preceding results imply that the only possibilities are the 2^r points $\vec{v}(\vec{\Delta})$, and we have just computed $\overline{F}(\vec{v}(\vec{\Delta})) = 0$ for all $\vec{\Delta}$. This completes the proof. □

Proof of Lemma 4.4.3. By Proposition 4.2.10 and e.g. [AGZ10, Chapter 2.4], $\mu(\vec{x}(\vec{\Delta}))$ has piecewise smooth density given by

$$\rho(\gamma) = \frac{1}{\pi}\Im(m(\gamma;\vec{x}(\vec{\Delta}))) = \frac{1}{\pi}\Im(u(\gamma;\vec{v}(\vec{\Delta}))),$$

where $u(\gamma;\vec{v}(\vec{\Delta})) = \sum_s \lambda_s u_s(\gamma;\vec{v}(\vec{\Delta}))$. So, it suffices to show $\vec{u}(\gamma;\vec{v}(\vec{\Delta}))$ is real for all $|\gamma| \le \varepsilon$. It is clear from (4.59) that

$$\vec{u}(\gamma;\vec{v}(\vec{\Delta})) = \vec{u}(0;\vec{v}(\vec{\Delta}) + \gamma\vec{\lambda}).$$

Recall $M(\vec{u}(\vec{\Delta})) = \operatorname{diag}(\xi') - \xi'' \succ 0$. Thus, $M(\vec{u}) \succ 0$ for $\vec{u}$ in an open neighborhood $\mathcal{N} \subseteq \mathbb{R}^r$ of $\vec{u}(\vec{\Delta})$. Hence solving (4.64) near $\vec{u}(\vec{\Delta})$ via inverse function theorem bijectively maps $\mathcal{N}$ to an open neighborhood $\mathcal{N}' \subseteq \mathbb{R}^r$ of $\vec{v}(\vec{\Delta})$. By Lemma 4.4.8 case (i), if $\vec{u} \in \mathcal{N}$ maps to $\vec{v} \in \mathcal{N}'$ under (4.64), then $\vec{u} = \vec{u}(0;\vec{v})$. In particular, for suitably small $\varepsilon > 0$, we have $\vec{v}(\vec{\Delta}) + \gamma\vec{\lambda} \in \mathcal{N}'$ for all $|\gamma| \le \varepsilon$. Thus $\vec{u}(0;\vec{v}(\vec{\Delta}) + \gamma\vec{\lambda})$ is real. ☐

4.4.3 Discussion of proof technique

Lemma 4.4.5 identifies approximately 3^r stationary points of F: for each species s, we may choose whether $\Re(u_s) = 0$, $|u_s| = 1/\sqrt{\xi^s(\vec{1})}$ has positive real part, or $|u_s| = 1/\sqrt{\xi^s(\vec{1})}$ has negative real part (though these do not always all exist, see Figure 4.4.1c). As we saw in the above proof, the 2^r stationary points where $u_s = \pm 1/\sqrt{\xi^s(\vec{1})}$ are global maximizers and the rest are saddle points of index at least 1. The main task in the proof of Proposition 4.4.1 was to rule out the extraneous critical points; in this subsection we motivate our method for doing so.

In the single-species case $r = 1$, in the topologically trivial regime F is convex on the interval $[-2\sqrt{\xi''}, 2\sqrt{\xi''}]$ and concave on its complement, and the maximum 0 is attained in the latter set; see Figure 4.4.1a. The boundary points $\pm 2\sqrt{\xi''}$ correspond to the radial derivative values where 0 enters or exits the limiting bulk spectrum of $\nabla^2_{\mathsf{sp}} H(\boldsymbol{x})$, and can be detected by the Stieltjes transform $m(0;x)$ becoming non-real. (In fact, F'' is discontinuous at these points.) This characterization of the convexity and concavity of F allows us to easily identify which of the critical points given by Lemma 4.4.5 are local maximizers.

However, a similar "region by region" convexity analysis will not work with multiple species. Similarly to the one-species case, $\nabla^2 F$ is discontinuous on a surface of radial derivative vectors $\vec{x}$ where 0 enters or exits the limiting bulk spectrum of $\nabla^2_{\mathsf{sp}} H(\boldsymbol{x})$, and this can be detected by $\vec{m}(0;\vec{x})$ (or equivalently $\vec{u}(0;\vec{v})$) becoming non-real. This boundary divides $\mathbb{R}^r$ into several regions and is depicted in Figures 4.4.1b and 4.4.1c as the blue curve. A natural approach to ruling out the extraneous critical points would be to show that, analogously to above, F is locally concave outside this boundary (in the regions containing the 2^r true maxima) and locally nonconcave inside it. However, this characterization is surprisingly not true. While F is indeed locally concave outside the boundary — if $M(\vec{u}) \succ 0$, then (4.66) implies $\nabla^2 \overline{F}(\vec{v}) \preceq 0$ — it is possible for F to also be locally concave inside it, for example in the purple regions in Figure 4.4.1c.

This counterexample rules out attempts to argue globally about convexity. This led us to the more direct approach of finding, at each extraneous critical point, a direction along which $\nabla^2 F$ is positive.

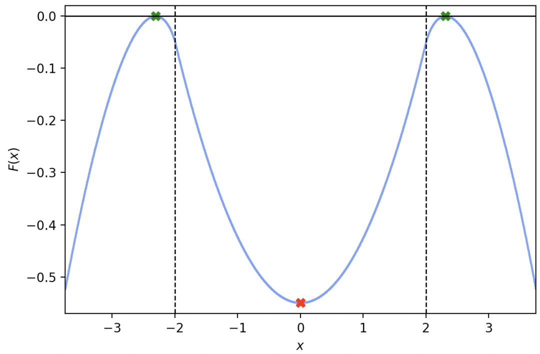
4.5 Approximate critical point control from Kac–Rice estimates

By Markov's inequality, negativity of the annealed Kac–Rice estimate (4.50) implies that H_N has no ε-atypical critical points (with high probability). The following main result of this section shows this implication is robust in some sense: the same Kac–Rice estimate also implies non-existence of certain *approximate* critical points.

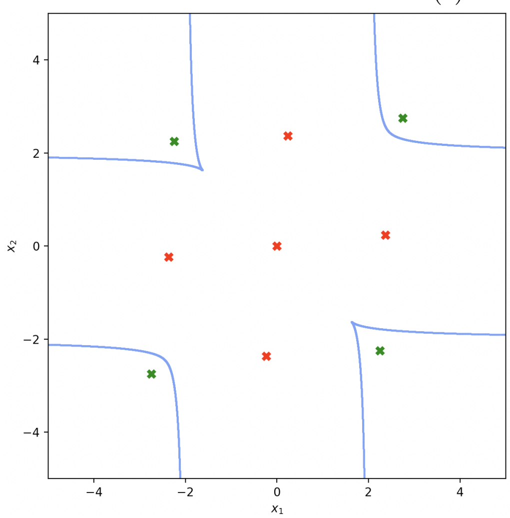
Proposition 4.5.1. *For strictly super-solvable ξ and any $\upsilon > 0$ there exists $\varepsilon = \varepsilon(\xi,\upsilon)$ such that with probability $1 - e^{-cN}$, all ε-approximate critical points are υ-good (recall Definition 4.1.10).*

Our approach proceeds as follows. Given H_N and $\delta > 0$, define the rerandomization

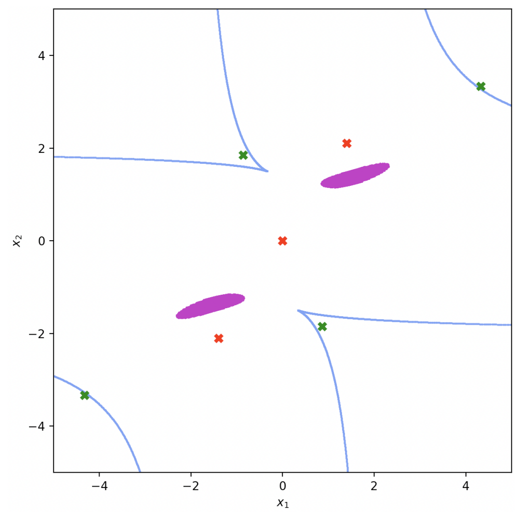
$$H_{N,\delta}(\boldsymbol{x}) = \sqrt{1-\delta}\, H_N(\boldsymbol{x}) + \sqrt{\delta}\, H'_N(\boldsymbol{x}) \tag{4.70}$$



(a) One species, $\xi' = 3$, $\xi'' = 1$



(b) Two species, $\vec{\lambda} = (0.5, 0.5)$, $\xi' = (4, 4)$, $\xi'' = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}$



(c) Two species, $\vec{\lambda} = (0.3, 0.7)$, $\xi' = (4.5, 4.5)$, $\xi'' = \begin{pmatrix} 1 & 2.4 \\ 2.4 & 1 \end{pmatrix}$

Figure 4.4.1: Figure 4.4.1a: the complexity functional F of a 1-species model is shown. F is tangent to the x-axis at two global maxima marked by green X's. The red X is a local minimum. The two dashed vertical lines mark the transition from local convexity to concavity, and F'' is discontinuous at these points.

Figures 4.4.1b and 4.4.1c: points of interest are shown in the domain $\mathbb{R}^2$ of the complexity functionals F for two different 2-species models. The green X's are global maxima where F equals 0, while the red X's are stationary points that are not local maxima.
The blue boundary is analogous to the dashed vertical lines in Figure 4.4.1a, and is where $\vec{m}(0; \vec{x})$ transitions from real and nonreal. In the four regions outside this boundary, $\vec{m}(0; \vec{x})$ is real, and in the region inside it $\vec{m}(0; \vec{x})$ is non-real. By Lemma 4.4.8 and continuity of $\vec{x} \mapsto \vec{m}(0; \vec{x})$ (see Theorem 4.A.2), this boundary is also the set of $\vec{x}$ for which $\vec{m}(0; \vec{x})$ is real and $M(\vec{m}(0; \vec{x}))$ is singular.
In Figure 4.4.1b, F is locally non-concave inside this boundary, but in Figure 4.4.1c F is also locally concave in the shaded purple regions. Note also that in Figure 4.4.1c, there are only three red X's instead of five; the 3^r stationary points identified by Lemma 4.4.5 do not necessarily all exist.

for H_N' an independent copy of H_N. Let $\mathsf{Crt}_{N,\delta}$ be the set of critical points for $H_{N,\delta}$. Our goal will be to show that if $\|\nabla_{\mathsf{sp}} H_N(\boldsymbol{x})\|_2 \le \varepsilon\sqrt{N}$ and H_N lies in a typical set (of probability $1 - e^{-cN}$), then for suitable

δ, ι tending to 0 with ε,

$$\mathbb{E}\left[|\mathsf{Crt}_{N,\delta} \cap B_{\iota\sqrt{N}}(\boldsymbol{x})| \;\middle|\; H_N\right] \geq e^{-o_\delta(N)}$$

This implies that if H_N has an ε-critical point that is not υ-good (for some υ also tending to 0 with ε), then the rerandomized Hamiltonian $H_{N,\delta}$ has on average at least $e^{-o_\delta(N)}$ critical points which are not $\upsilon/2$-good. Combining with Theorem 4.1.11, which shows the number of such critical points is exponentially small, will then yield Proposition 4.5.1.

In fact, this argument proves the following much more general result, which we believe is of significant independent interest. Let $\mathcal{J}$ consist of all compact subsets of $\mathbb{R}$, equipped with the Hausdorff metric. We consider a non-empty subset

$$\overline{\mathcal{D}} \subseteq \mathbb{R}^r \times \mathbb{R}^r \times \mathbb{R} \times \mathcal{J} \times \mathbb{W}_1(\mathbb{R}),$$

where the right-hand product is equipped with the supremum metric over its five factors.

Given $\iota, \varepsilon \geq 0$, we define the set $\mathsf{Crt}_N^{\overline{\mathcal{D}},\varepsilon,\upsilon}(H_N) \subseteq \mathcal{S}_N$ of ε-approximate critical points $\boldsymbol{x} \in \mathcal{S}_N$ for H_N which are υ-far from being described by an element of $\overline{\mathcal{D}}$, i.e. which satisfy

$$d\Big(\Big(\nabla_{\mathsf{rad}} H_N(\boldsymbol{x}), \vec{R}(\boldsymbol{x}, \boldsymbol{G}^{(1)}), \frac{H_N(\boldsymbol{x})}{N}, \underset{H_N}{\mathsf{spec}}(\boldsymbol{x}), \widehat{\mu}_{H_N}(\boldsymbol{x})\Big), \overline{\mathcal{D}}\Big) \geq \upsilon. \tag{4.71}$$

Recall (4.10) for definitions of $\mathsf{spec}_{H_N}(\cdot)$ and $\widehat{\mu}_{H_N}(\cdot)$. As usual, the distance from a point to a set is the infimal point-to-point distance; recall also the definition of ∇_{rad} near Fact 4.1.4. Note that $\mathsf{Crt}_N^{\overline{\mathcal{D}},\varepsilon,\iota}(H_N)$ is an infinite set with positive probability unless $\varepsilon = 0$.

Theorem 4.5.2. *Suppose $\xi, \overline{\mathcal{D}}, \varepsilon, \upsilon, c_0$ are such that for N large enough,*

$$\mathbb{E}\big|\mathsf{Crt}_N^{\overline{\mathcal{D}},0,\upsilon/2}(H_N)\big| \leq e^{-c_0 N}. \tag{4.72}$$

Then for some small $\varepsilon > 0$ depending only on (ξ, υ, c_0), for some $c > 0$ and all N large enough:

$$\mathbb{P}\big[\big|\mathsf{Crt}_N^{\overline{\mathcal{D}},\varepsilon,\upsilon}(H_N)\big| \geq 1\big] \leq e^{-cN}.$$

Proof of Proposition 4.5.1 from Theorem 4.5.2: Let $\overline{\mathcal{D}}$ be the size 2^r set:

$$\overline{\mathcal{D}} = \overline{\mathcal{D}}(\xi) = \left\{ \Big(\vec{x}(\vec{\Delta}), \vec{R}(\vec{\Delta}), E(\vec{\Delta}), S(\vec{\Delta}), \mu(\vec{x}(\vec{\Delta}))\Big) \;:\; \vec{\Delta} \in \{-1,1\}^r \right\}.$$

Then Theorem 4.1.11(b) implies the condition (4.72) for any $\upsilon > 0$, for some correspondingly small $c > 0$. Therefore we may apply Theorem 4.5.2 which completes the proof. ☐

We note that although $\overline{\mathcal{D}}$ is defined in near-maximal generality above, examining just one of its 5 components yields interesting consequences. For instance Section 4.7 considers only the radial derivative.

4.5.1 Technical properties of the conditional vector Dyson equation

In this somewhat technical subsection, we study the vector Dyson equation corresponding to (4.70). Note that when both H_N and H_N' are treated as random, the relevant Dyson equation is exactly as in our usual setting explained in Subsection 4.2.3. However if one first conditions on H_N, then (4.70) yields a different vector Dyson equation with an apriori somewhat different solution that depends on H_N.

We show below that with extremely high probability over H_N, the latter "conditional" solution is uniformly close to that of the unconditional equation. This is crucial because, since we do wish to condition on H_N in our main argument, we need to apply the asymptotic determinant results from [BBM23] to this conditioned vector Dyson equation. The main idea is that with or without conditioning, concentration of the empirical spectrum of $H_{N,\delta}$ implies it is well-described by the Dyson equation's solution with very high probability. Hence these solutions must in fact be similar, with high probability over H_N.

Recall the definition of C-regular probability measure from Definition 4.2.13. We next define a class of C-regular random matrix models, which have C'-regular spectral measure in some sense; see Lemma 4.5.4 just below.

Definition 4.5.3. The (law of the) random symmetric matrix $M_N \in \mathbb{R}^{N\times N}$ is C**-regular** if it is given by $M_N = W_N + A_N$ where:

(a) A_N is deterministic and $\|A_N\|_{\mathsf{op}} \le C$.

(b) W_N is a centered Gaussian matrix with independent entries on and above the diagonal.

(c) Each entry of W_N has variance in $\left[\frac{1}{CN}, \frac{C}{N}\right]$.

Given a C-regular random matrix M_N, for each $z \in \mathbb{H}$ we let $G_N(z) \in \mathbb{C}^{N\times N}$ be the unique solution to the equation

$$I_N + (zI_N - A_N + \mathbb{E}[W_N G_N(z) W_N])G_N(z) = 0 \tag{4.73}$$

with the constraint that the imaginary part $\Im(G_N(z))$ is a strictly positive definite matrix. We let $\mu_{M_N} \in \mathcal{P}(\mathbb{R})$ be the (unique) probability measure with Stieltjes transform $\mathsf{Tr}(G_N(z))/N$. Such $G_N(z)$ and μ_{M_N} exist and are unique by e.g. [EKS19, Proposition 5.1 (i),(ii)].

Lemma 4.5.4. *If M_N is C_1-regular, then μ_{M_N} is C_2-regular for C_2 depending only on C_1. Further, for any event E with $\mathbb{P}[E] \ge 1/2$,*

$$\mathbb{W}_1\big(\mu_{M_N}, \mathbb{E}[\widehat{\mu}_{M_N}]\big) \le \delta_N, \tag{4.74}$$

$$\left|\frac{1}{N}\log \mathbb{E}\left[1_E \cdot |\det M_N|\right] - \int \log|\lambda| \mathsf{d}\mu_{M_N}(\lambda)\right| \le \delta_N \tag{4.75}$$

for a sequence $\delta_N \to 0$ depending only on C_1.

Proof. The first assertion follows from [AEK17a, Theorem 2.6]. Next, (4.74) follows from [EKS19, Theorem 2.1(4b)] and [BBM23, Proposition 3.1] as in [BBM23, Proof of Corollary 1.9.B]. The second part (4.75) is a rewriting of [BBM23, Corollary 1.9.A] except for the presence of the event E. This additional ingredient follows by the same proof since e.g. at the end of [BBM23, Proof of Theorem 1.2], the probabilities of all good events $\mathcal{E}_{\mathrm{Lip}}, \mathcal{E}_{\mathrm{gap}}, \mathcal{E}_{\mathrm{b}}$ are shown to tend to 1. Indeed as $\mathbb{P}[E] \ge 1/2$, the factor 1_E only affects said lower bound by an additive $O(1/N)$. □

Lemma 4.5.5. *Suppose $M_N = W_N + A_N$ and $M_N' = W_N' + A_N'$ are C_1-regular and $W_N \overset{d}{=} W_N'$. Then*

$$\mathbb{W}_1\left(\mathbb{E}[\widehat{\mu}_{M_N}], \mathbb{E}[\widehat{\mu}_{M_N'}]\right) \le \frac{1}{\sqrt{N}}\|A_N - A_N'\|_F.$$

Proof. By coupling $W_N = W_N'$ and then using the Hoffman–Wielandt lemma (see e.g. [AGZ10, Lemma 2.1.19]), one finds:

$$\mathbb{W}_2\left(\mathbb{E}[\widehat{\mu}_{M_N}], \mathbb{E}[\widehat{\mu}_{M_N'}]\right)^2 \le \frac{1}{N}\|A_N - A_N'\|_F^2.$$

The Cauchy–Schwarz inequality implies that $\mathbb{W}_1$ distance is smaller than $\mathbb{W}_2$ distance, completing the proof. □

Given $\mu \in \mathcal{P}(\mathbb{R})$, let $\mu^{(C)}$ be the pushforward of μ under $x \mapsto \min(C, \max(-C, x))$.

Lemma 4.5.6. *Suppose $M_N = W_N + A_N$ is C_1-regular. Then for some $\varepsilon_0 > 0$ and any $C_2, \delta > 0$, there exists $C_3 = C_3(C_1, C_2, \delta)$ such that with N sufficiently large:*

$$\mathbb{P}\left[\mathbb{W}_1\big(\widehat{\mu}_{M_N}^{(C_2)}, \mathbb{E}[\widehat{\mu}_{M_N}^{(C_2)}]\big) \ge \delta\right] \le C_3 e^{-c(C_1,C_2,\delta)N^{1+\varepsilon_0}}.$$

Proof. For any 1-Lipschitz test function f, it is shown in condition (L) in [BBM23, Equation (1.11)] that the concentration

$$\mathbb{P}\left[\left|\mathbb{E}^{\widehat{\mu}_{M_N}}[f] - \mathbb{E}\big[\mathbb{E}^{\widehat{\mu}_{M_N}}[f]\big]\right| \ge \delta/10\right] \le C_3 e^{-c(C_1,C_2,\delta)N^{1+\varepsilon_0}}. \tag{4.76}$$

Here our ε_0 is their $\varepsilon_0 - \zeta$; in our setting this statement follows by a Herbst argument as in [BBM23, Proof of Corollary 1.9.B]. Since $x \mapsto \min(C_2, \max(-C_2, x))$ is 1-Lipschitz, the composition $f^{(C_2)}(x) \equiv$

$f\big(\min(C_2,\max(-C_2,x))\big)$ is as well. To obtain the claimed Wasserstein bound, recall that the $\mathbb{W}_1$ distance makes $\mathcal{P}([-C_2,C_2])$ a compact metric space, and coincides with the bounded Lipschitz metric:

$$\mathbb{W}_1(\mu,\widetilde{\mu}) = \sup_{\substack{f:\mathbb{R}\to\mathbb{R}\\ Lip(f)\leq 1}} \big(\mathbb{E}^{\mu}[f] - \mathbb{E}^{\widetilde{\mu}}[f]\big). \tag{4.77}$$

Hence for any $\delta>0$ and C_2, we may choose a finite $\delta/10$-net $\mathcal{N}\subseteq\mathcal{P}([-C_2,C_2])$ with respect to $\mathbb{W}_1$. Then for each distinct pair $\mu_i,\mu_j\in\mathcal{N}$, we may choose a 1-Lipschitz $f_{i,j}:\mathbb{R}\to\mathbb{R}$ such that $|\mathbb{E}^{\mu_i}[f_{i,j}]-\mathbb{E}^{\mu_j}[f_{i,j}]| = \mathbb{W}_1(\mu_i,\mu_j)$. Since $|\mathcal{N}|$ is independent of N, the event in (4.76) holds for all $f_{i,j}$ simultaneously with probability $1-C_3'e^{-c(C_1,C_2,\delta)N^{1+\varepsilon_0}}$. Finally, we choose i,j so that $\mathbb{W}_1(\mu_i,\widehat{\mu}^{(C_2)}_{M_N})\leq\delta/10$ and $\mathbb{W}_1(\mu_j,\mathbb{E}\widehat{\mu}^{(C_2)}_{M_N})\leq\delta/10$. On the event that (4.76) applies to $f_{i,j}$, we thus find:

$$\begin{aligned}
\mathbb{W}_1\big(\widehat{\mu}^{(C_2)}_{M_N},\mathbb{E}[\widehat{\mu}^{(C_2)}_{M_N}]\big) &\leq \mathbb{W}_1(\widehat{\mu}^{(C_2)}_{M_N},\mu_i) + \mathbb{W}_1\big(\mu_i,\mu_j\big) + \mathbb{W}_1(\mu_j,\mathbb{E}\widehat{\mu}^{(C_2)}_{M_N})\\
&\leq |\mathbb{E}^{\mu_i}[f_{i,j}]-\mathbb{E}^{\mu_j}[f_{i,j}]| + \frac{\delta}{5}\\
&\overset{(4.77)}{\leq} \big|\mathbb{E}^{\widehat{\mu}^{(C_2)}_{M_N}}[f_{i,j}] - \mathbb{E}\big[\mathbb{E}^{\widehat{\mu}^{(C_2)}_{M_N}}[f_{i,j}]\big]\big| + \frac{2\delta}{5}\\
&\overset{(4.76)}{\leq} \delta/2.
\end{aligned}$$

This completes the proof, since the value $|\mathcal{N}|$ depends only on C_2 and δ, hence can be absorbed into the value C_3. $\square$

For our Kac–Rice application, we will fix some $H_N\in K_N$ (recall Proposition 4.2.4) and condition also on $\vec{a}(\boldsymbol{x}) = \nabla_{\mathsf{rad}}H_{N,\delta}(\boldsymbol{x})$, where $H_{N,\delta}(\boldsymbol{x})$ is as in equation (4.70). Let us assume $\|\vec{a}(\boldsymbol{x})\|_\infty\leq C$, which holds with probability $1-e^{-cN}$ for some constant C by Proposition 4.2.4. Then the law of $\nabla^2_{\mathsf{sp}}H_{N,\delta}(\boldsymbol{x})$ conditionally on $\vec{a}(\boldsymbol{x})$ is a C_1-regular matrix

$$\nabla^2_{\mathsf{sp}}H_{N,\delta}(\boldsymbol{x}) = \underbrace{\sqrt{\delta}\,\nabla^2_{\mathcal{T}\times\mathcal{T}}H_N'(\boldsymbol{x})}_{W_N} + \underbrace{\sqrt{1-\delta}\,\nabla^2_{\mathcal{T}\times\mathcal{T}}H_N(\boldsymbol{x}) - \mathrm{diag}(\Lambda^{-1/2}\vec{a}(\boldsymbol{x})\diamond\mathbf{1}_{\mathcal{T}})}_{A_N}.$$

We write $\mu^{\delta}_{H_N,\vec{a},\boldsymbol{x}}\in\mathcal{P}(\mathbb{R})$ for the probability measure with Stieltjes transform the corresponding solution to (4.73) for $\nabla^2_{\mathsf{sp}}H_{N,\delta}(\boldsymbol{x})$. Also let $\mu_{\xi,\vec{a}} = \mu^{1}_{H_N,\vec{a},\boldsymbol{x}}$. However, we emphasize that $\mu^{\delta}_{H_N,\vec{a},\boldsymbol{x}}$ makes sense even for $\vec{a}\neq\nabla_{\mathsf{rad}}H_{N,\delta}(\boldsymbol{x})$. In fact in the arguments below, we will obtain control on $\mu^{\delta}_{H_N,\vec{a},\boldsymbol{x}}$ by estimating it locally uniformly in $\vec{a}$, and only substituting $\vec{a}=\vec{a}(\boldsymbol{x})$ at the end. First we show that for **fixed** $\vec{a}$, the measure $\mu^{\delta}_{H_N,\vec{a},\boldsymbol{x}}$ concentrates sharply around $\mu_{\xi,\vec{a}}$. The idea is to apply Lemma 4.5.6 both before and after conditioning on H_N: since it yields concentration of the spectral measure in both cases, they must concentrate around approximately the same measure.

Lemma 4.5.7. *There is some $\varepsilon_0>0$ such that for any $C,\delta>0$ there is $C',c>0$ such that the following holds. For each fixed $\boldsymbol{x}\in\mathcal{S}_N$ and fixed C-bounded $\vec{a}\in\mathbb{R}^r$,*

$$\mathbb{P}\left[H_N\in K_N \ \ \textit{and}\ \ \mathbb{W}_1\big(\mu^{\delta}_{H_N,\vec{a},\boldsymbol{x}},\mu_{\xi,\vec{a}}\big)\geq\delta\right]\leq C'e^{-cN^{1+\varepsilon_0}}$$

for N sufficiently large.

Proof. Lemma 4.5.4 shows $\mu_{\xi,\vec{a}}$ is C_2-regular. Moreover if $H_N\in K_N$ and $\|\vec{a}\|_\infty\leq C$, it implies that $\mu^{\delta}_{H_N,\vec{a},\boldsymbol{x}}$ is C_2-regular. Let $M_N^\delta\equiv\nabla^2_{\mathsf{sp}}H_{N,\delta}(\boldsymbol{x})$; then applying (4.74) conditionally on H_N shows

$$\mathbb{W}_1\big(\mathbb{E}[\widehat{\mu}_{M_N^\delta}\,|\,H_N],\mu^{\delta}_{H_N,\vec{a},\boldsymbol{x}}\big)\leq\delta/4.$$

Applying Lemma 4.5.6 also conditionally on H_N gives:

$$\mathbb{P}\left[H_N,\frac{H_{N,\delta}}{2}\in K_N\ \ \text{and}\ \ \mathbb{W}_1\big(\widehat{\mu}_{M_N^\delta},\mathbb{E}[\widehat{\mu}_{M_N^\delta}\,|\,H_N]\big)\geq\delta/4\right]\leq C'e^{-cN^{1+\varepsilon_0}}/4.$$

Here $\frac{H_{N,\delta}}{2} \in K_N$ was used to imply $\widehat{\mu}_{M_N^\delta} = \widehat{\mu}^{(C_2)}_{M_N^\delta}$. Lemma 4.5.6 applied without conditioning on H_N similarly yields

$$\mathbb{P}\left[H_N, \frac{H_{N,\delta}}{2} \in K_N \text{ and } \mathbb{W}_1\big(\widehat{\mu}_{M_N^\delta}, \mathbb{E}[\widehat{\mu}_{M_N^\delta}]\big) \geq \delta/4\right] \leq C' e^{-cN^{1+\varepsilon_0}}/4.$$

Finally since $\mathbb{E}[\widehat{\mu}_{M_N^\delta}] = \mathbb{E}[\widehat{\mu}_{M_N}]$, it follows from (4.74) (now applied unconditionally) that

$$\mathbb{W}_1\big(\mathbb{E}[\widehat{\mu}_{M_N}], \mu_{\xi,\vec{a}}\big) \leq \delta/4.$$

Combining using the triangle inequality yields

$$\mathbb{P}\left[H_N, \frac{H_{N,\delta}}{2} \in K_N \text{ and } \mathbb{W}_1\big(\mu^\delta_{H_N,\vec{a},\boldsymbol{x}}, \mu_{\xi,\vec{a}}\big) \geq \delta\right] \leq C' e^{-cN^{1+\varepsilon_0}}/2.$$

It remains to observe that the event $\mathbb{W}_1\big(\mu^\delta_{H_N,\vec{a},\boldsymbol{x}}, \mu_{\xi,\vec{a}}\big) \geq \delta$ is determined by H_N, and that

$$\mathbb{P}\left[\frac{H_{N,\delta}}{2} \in K_N \mid H_N\right] \geq \mathbb{P}[H_N' \in K_N] \geq 1/2 \tag{4.78}$$

for any $H_N \in K_N$. The former inequality in (4.78) follows by writing $H_{N,\delta}/2 = \frac{\sqrt{1-\delta}}{2} H_N + \frac{\sqrt{\delta}}{2} H_N'$ because $\frac{\sqrt{1-\delta}}{2} + \frac{\sqrt{\delta}}{2} \leq 1$. Indeed since K_N is a symmetric convex set it contains the origin, so if $H_N, H_N' \in K_N$ then also $H_{N,\delta}/2 \in K_N$. The latter inequality in (4.78) is one of the defining properties of K_N in Proposition 4.2.4, since H_N' is an i.i.d. copy of H_N. □

Next in Lemma 4.5.8 and Proposition 4.5.10, we exhaust all bounded $\vec{a}$ in Lemma 4.5.7. Combined with the continuity shown in Lemma 4.5.5, this shows validity of Lemma 4.5.7 uniformly in bounded $\vec{a}$.

Lemma 4.5.8. *For any $C, \varepsilon > 0$, with probability $1 - e^{-cN}$, $H_N \in K_N$ and*

$$\sup_{\boldsymbol{x} \in \mathcal{S}_N, \|\vec{a}\|_\infty \leq C} \mathbb{W}_1\big(\mu^\delta_{H_N,\boldsymbol{x},\vec{a}}, \mu_{\xi,\vec{a}}\big) \leq \varepsilon.$$

Proof. Let $\mathcal{N}_N$ be an N^{-10}-net for $\mathcal{S}_N$ and $\mathcal{A}_N$ an N^{-10}-net for the C-bounded vectors $\vec{a}$. Note that $|\mathcal{N}_N \times \mathcal{A}_N| \leq N^{O(CN)}$. Since $\mathbb{P}[H_N \in K_N] \geq 1 - e^{-cN}$, union-bounding over the events in Lemma 4.5.7 over $\mathcal{N}_N \times \mathcal{A}_N$ implies that with probability $1 - e^{-c'N}$,

$$\sup_{\boldsymbol{x} \in \mathcal{N}_N, \vec{a} \in \mathcal{A}_N} \mathbb{W}_1\big(\mu^\delta_{H_N,\boldsymbol{x},\vec{a}}, \mu_{\xi,\vec{a}}\big) \leq \varepsilon/2. \tag{4.79}$$

Next assuming again that $H_N \in K_N$, for any $\hat{\boldsymbol{x}} \in \mathcal{S}_N$ and C-bounded $\hat{\vec{a}}$, let $\boldsymbol{x}, \vec{a}$ be the nearest points in $\mathcal{N}_N, \mathcal{A}_N$. Let $M_N^\delta = M_N^\delta(\boldsymbol{x}), \widehat{M}_N^\delta = M_N^\delta(\hat{\boldsymbol{x}})$ be the associated random Riemannian Hessians of $H_{N,\delta}$ given H_N. Then Lemma 4.5.5 implies

$$\mathbb{W}_1\big(\mathbb{E}[\widehat{\mu}_{M_N^\delta}], \mathbb{E}[\widehat{\mu}_{\widehat{M}_N^\delta}]\big) \leq N^{-3}.$$

Recalling the deterministic bound (4.74), we have

$$\mathbb{W}_1\big(\mu_{M_N^\delta}, \mu_{\widehat{M}_N^\delta}\big) \leq 2\delta_N + N^{-3}$$

whenever $H_N \in K_N$. Combining with (4.79) completes the proof, as $\mu_{M_N^\delta} = \mu^\delta_{H_N,\boldsymbol{x},\vec{a}}$. □

For $\boldsymbol{x} \in \mathcal{S}_N$, define the distorted Hessian

$$J_N(\boldsymbol{x}) \equiv \Xi \diamond \nabla^2_{\mathsf{sp}} H_N(\boldsymbol{x}), \qquad \Xi = (\xi')^{-1/2} \otimes (\xi')^{-1/2} \in \mathbb{R}^{r\times r}. \tag{4.80}$$

This is again C-regular conditionally on $\vec{a} = \nabla_{\mathsf{rad}} H_N(\boldsymbol{x})$, and we let $\widetilde{\mu}_{\xi,\vec{a}}$ be the corresponding solution to (4.73).

Definition 4.5.9. The set $K_N(\varepsilon) \subseteq K_N$ consists of all $H_N \in \mathscr{H}_N$ satisfying:

(a) $H_N \in K_N$.

(b) For all $\boldsymbol{x} \in \mathcal{S}_N$, with $M_N = \nabla^2_{\mathsf{sp}} H_N(\boldsymbol{x})$,

$$\mathbb{W}_1(\widehat{\mu}_{M_N}, \mu_{\xi, \nabla_{\mathsf{rad}} H_N(\boldsymbol{x})}) \le \varepsilon.$$

(c) For all $\boldsymbol{x} \in \mathcal{S}_N$, with $\widetilde{M}_N = J_N(\boldsymbol{x})$,

$$\mathbb{W}_1(\widehat{\mu}_{\widetilde{M}_N}, \widetilde{\mu}_{\xi, \nabla_{\mathsf{rad}} H_N(\boldsymbol{x})}) \le \varepsilon.$$

(d) For all $\boldsymbol{x} \in \mathcal{S}_N$ and $\|\vec{a}\|_\infty \le C$ we have

$$\mathbb{W}_1\big(\mu^{\delta}_{H_N, \boldsymbol{x}, \vec{a}}, \mu_{\xi, \vec{a}}\big) \le \varepsilon.$$

Proposition 4.5.10. *For any $\varepsilon > 0$, we have $\mathbb{P}[H_N \in K_N(\varepsilon)] \ge 1 - e^{-c(\varepsilon)N}$.*

Proof. Part (a) follows from Proposition 4.2.4. Part (d) follows by Lemma 4.5.8. A similar argument implies parts (b) and (c). ☐

4.5.2 Main argument

We fix a small constant δ and set

$$(\varepsilon, \alpha, \eta, \iota) = \big(\delta^{10}, \delta^{1/3}, \delta^{1/10}, \delta^{1/100}\big). \tag{4.81}$$

Lemma 4.5.11. *Fix $H_N \in K_N(\varepsilon)$. For each $\boldsymbol{y} \in \mathcal{S}_N$ and C-bounded $\vec{a}$, with N sufficiently large, let $E_{\boldsymbol{y}}$ be an event satisfying*

$$\inf_{\boldsymbol{y} \in \mathcal{S}_N} \mathbb{P}\Big[E_{\boldsymbol{y}} \;\Big|\; \big(\nabla_{\mathsf{sp}} H_{N,\delta}(\boldsymbol{y}), \nabla_{\mathsf{rad}} H_{N,\delta}(\boldsymbol{y}), H_N\big)\Big] \ge 1/2.$$

Then including $1_{E_{\boldsymbol{y}}}$ within the determinant expectation for $\nabla^2_{\mathsf{sp}} H_{N,\delta}(\boldsymbol{y})$ has a negligible effect, in the sense that uniformly in $\boldsymbol{y}$:

$$\begin{aligned} &\Bigg| \frac{1}{N} \log \mathbb{E}\Big[1_{E_{\boldsymbol{y}}} \cdot \big|\det \nabla^2_{\mathsf{sp}} H_{N,\delta}(\boldsymbol{y})\big| \;\Big|\; \big(\nabla_{\mathsf{sp}} H_{N,\delta}(\boldsymbol{y}) = 0, \nabla_{\mathsf{rad}} H_{N,\delta}(\boldsymbol{y}) = \vec{a}, H_N\big)\Big] \\ &\quad - \int \log|\lambda| \mathsf{d}\mu_{\xi,\vec{a}}(\lambda) \Bigg| \le o_\delta(1). \end{aligned} \tag{4.82}$$

Proof. By Lemma 4.2.2, $\nabla^2_{\mathcal{T}\times\mathcal{T}} H_{N,\delta}(\boldsymbol{y})$ and $\nabla H_{N,\delta}(\boldsymbol{y})$ are independent. Recalling (4.70) the conditional law of $\nabla^2_{\mathsf{sp}} H_{N,\delta}(\boldsymbol{y})$ agrees with that of

$$\sqrt{1-\delta}\nabla^2_{\mathcal{T}\times\mathcal{T}} H_N(\boldsymbol{y}) + \sqrt{\delta}\nabla^2_{\mathcal{T}\times\mathcal{T}} H'_N(\boldsymbol{y}) - \mathrm{diag}(\Lambda^{-1/2}\vec{a} \diamond \mathbf{1}_{\mathcal{T}})$$

which is $C(\delta)$-regular for $H_N \in K_N$. Upper-bounding the left-hand side of (4.82) by

$$\begin{aligned} &\Bigg| \frac{1}{N} \log \mathbb{E}\Big[1_{E_{\boldsymbol{y}}} \cdot \big|\det \nabla^2_{\mathsf{sp}} H_{N,\delta}(\boldsymbol{y})\big| \;\Big|\; \big(\nabla_{\mathsf{sp}} H_{N,\delta}(\boldsymbol{y}) = 0, \nabla_{\mathsf{rad}} H_{N,\delta}(\boldsymbol{y}) = \vec{a}, H_N\big)\Big] \\ &\quad - \int \log|\lambda| \mathsf{d}\mu^{\delta}_{H_N,\boldsymbol{x},\vec{a}}(\lambda) \Bigg| + \left| \int \log|\lambda| \mathsf{d}\mu^{\delta}_{H_N,\boldsymbol{x},\vec{a}}(\lambda) - \int \log|\lambda| \mathsf{d}\mu_{\xi,\vec{a}}(\lambda) \right|, \end{aligned}$$

we can apply Lemma 4.5.4 to bound the first term and Lemma 4.2.14 (using part (d) of Definition 4.5.9) to the second. ☐

The next lemma gives a Taylor expansion estimate for $\|\nabla_{\mathsf{sp}} H_N(\boldsymbol{y})\|_2$. In it, we let $\gamma_{\boldsymbol{x}\to\boldsymbol{y}} : [0,1] \to \mathcal{S}_N$ be the shortest path geodesic with $\gamma_{\boldsymbol{x}\to\boldsymbol{y}}(0) = \boldsymbol{x}$ and $\gamma_{\boldsymbol{x}\to\boldsymbol{y}}(1) = \boldsymbol{y}$ (say, whenever $\boldsymbol{y} \in B_{\alpha\sqrt{N}}(\boldsymbol{x})$ so the shortest path is clear). Note that $\gamma'_{\boldsymbol{x}\to\boldsymbol{y}}(0)$ is approximately $\boldsymbol{y} - \boldsymbol{x}$; the result below holds with this replacement as well, but for the application $\gamma'_{\boldsymbol{x}\to\boldsymbol{y}}(0)$ will be more convenient since it is in the tangent space $T_{\boldsymbol{x}}\mathcal{S}_N$.

Lemma 4.5.12. *Let $H_N \in K_N$. For $\boldsymbol{y} \in B_{\alpha\sqrt{N}}(\boldsymbol{x}) \cap \mathcal{S}_N$ and bounded $\vec{v} \in \mathbb{R}^r$, we have*

$$\|\vec{v} \diamond \nabla_{\mathsf{sp}} H_N(\boldsymbol{y})\|_2 \le \|\vec{v} \diamond \nabla_{\mathsf{sp}} H_N(\boldsymbol{x}) + \vec{v} \diamond \nabla^2_{\mathsf{sp}} H_N(\boldsymbol{x}) \cdot \gamma'_{\boldsymbol{x}\to\boldsymbol{y}}(0)\|_2 + C\alpha^2\sqrt{N}\,. \tag{4.83}$$

Proof. Let $P_t : T_{\gamma_{\boldsymbol{x}\to\boldsymbol{y}}(t)}\mathcal{S}_N \to T_{\boldsymbol{x}}\mathcal{S}_N$ be the associated parallel transport map on tangent spaces. This is a product of isometries on subspaces of each $\mathbb{R}^{\mathcal{I}_s}$, so

$$\|\vec{v} \diamond \nabla_{\mathsf{sp}} H_N(\gamma(t))\|_2 = \|\vec{v} \diamond P_t\big(\nabla_{\mathsf{sp}} H_N(\gamma(t))\big)\|_2.$$

Moreover,

$$\frac{\mathsf{d}}{\mathsf{d}t}\left[\vec{v} \diamond P_t\big(\nabla_{\mathsf{sp}} H_N(\gamma_{\boldsymbol{x}\to\boldsymbol{y}}(t))\big)\right]\Big|_{t=0} = \vec{v} \diamond \nabla^2_{\mathsf{sp}} H_N(\gamma_{\boldsymbol{x}\to\boldsymbol{y}}(t))\gamma'_{\boldsymbol{x}\to\boldsymbol{y}}(t)\,.$$

It remains to apply Taylor's theorem to $\vec{v} \diamond P_t\big(\nabla_{\mathsf{sp}} H_N(\gamma(0))\big)$ with derivative bounds from Proposition 4.2.4 as $H_N \in K_N$. ☐

Lemma 4.5.13. *Fix $H_N \in K_N(\varepsilon)$ and $\boldsymbol{x}$ such that*

$$\|\nabla_{\mathsf{sp}} H_N(\boldsymbol{x})\|_2 \le \varepsilon\sqrt{N}\,. \tag{4.84}$$

Then in expectation over H'_N, there are at least $e^{-o_\delta(N)}$ points $\boldsymbol{y} \in B_{\alpha\sqrt{N}}(\boldsymbol{x}) \cap \mathcal{S}_N$ such that

$$|H_N(\boldsymbol{x}) - H_{N,\delta}(\boldsymbol{y})| \le \iota^2 N, \tag{4.85}$$

$$\nabla_{\mathsf{sp}} H_{N,\delta}(\boldsymbol{y}) = 0, \tag{4.86}$$

$$\|\nabla_{\mathsf{rad}} H_N(\boldsymbol{x}) - \nabla_{\mathsf{rad}} H_{N,\delta}(\boldsymbol{y})\|_\infty \le \iota^2, \tag{4.87}$$

$$\|\nabla^2_{\mathsf{sp}} H_N(\boldsymbol{x}) - \nabla^2_{\mathsf{sp}} H_{N,\delta}(\boldsymbol{y})\|_{\mathsf{op}} \le \iota. \tag{4.88}$$

Proof. Throughout the proof we fix and condition on $H_N \in K_N(\varepsilon)$. Then with high conditional probability, the events (4.85), (4.87), and (4.88) all occur (recall from (4.81) that ι is much larger than α and δ). Let E_1 be the event that (4.85) and (4.87) hold, and $E_{2,\boldsymbol{y}}$ the event that (4.88) holds.

By considering only the contribution from $E_1 \cap E_{2,\boldsymbol{y}}$ and recalling Lemma 4.2.2 part (a), we find that the expected number of such critical points is at least

$$\begin{aligned}(2\pi\delta)^{-\frac{N-r}{2}} \prod_{s\in\mathscr{S}} (\xi^s)^{\frac{N_s-1}{2}} \int_{\boldsymbol{y}\in B_{\alpha\sqrt{N}}(\boldsymbol{x})\cap\mathcal{S}_N} \Bigg(\exp\left(\frac{-(1-\delta)}{2\delta}\|(\xi')^{-1/2} \diamond \nabla_{\mathsf{sp}} H_N(\boldsymbol{y})\|_2^2\right) 1_{E_1} \\ \times \min_{\|\vec{a}-\nabla_{\mathsf{rad}} H_N(\boldsymbol{x})\|_\infty \le \iota^2} \mathbb{E}\Big[1_{E_{2,\boldsymbol{y}}} \cdot \big|\det \nabla^2_{\mathsf{sp}} H_{N,\delta}(\boldsymbol{y})\big| \\ \Big| \big(\nabla_{\mathsf{sp}} H_{N,\delta}(\boldsymbol{y}) = 0, \nabla_{\mathsf{rad}} H_{N,\delta}(\boldsymbol{y}) = \vec{a}, H_N\big)\Big]\Bigg)\,\mathsf{d}\boldsymbol{y}\end{aligned} \tag{4.89}$$

If in addition to H_N one conditions on any $(H_{N,\delta}(\boldsymbol{y}), \nabla_{\mathsf{rad}} H_{N,\delta}(\boldsymbol{y}))$ satisfying E_1, we claim the conditional probability of $E_{2,\boldsymbol{y}}$ (i.e. (4.88)) is at least 1/2. Indeed, the conditional mean of $\nabla^2_{\mathsf{sp}} H_{N,\delta}(\boldsymbol{y})$ is given by $\sqrt{1-\delta}\nabla^2_{\mathsf{sp}} H_{N,\delta}$, plus an additive shift of operator norm $O(\iota^2) \ll \iota$ (coming from linear regression via Lemma 4.2.2). The conditionally random part of $\nabla^2_{\mathsf{sp}} H_{N,\delta}(\boldsymbol{y})$ consists of an additive $\sqrt{\delta}\nabla^2_{\mathcal{T}\times\mathcal{T}} H'_N(\boldsymbol{y})$, which has operator norm $O(\sqrt{\delta}) \ll \iota$ with high probability.

Hence Lemma 4.5.11 (which holds uniformly in $\boldsymbol{y}$) implies that the latter part of the integrand is, for each $\vec{a}$ such that $\|\vec{a} - \nabla_{\mathsf{rad}} H_N(\boldsymbol{x})\|_\infty \le \iota^2$:

$$\exp\left(N \int \log|\lambda|\,[\mu_{\xi,\nabla_{\mathsf{rad}} H_N(\boldsymbol{x})}](\mathsf{d}\lambda) \pm o_\iota(N)\right). \tag{4.90}$$

Indeed, (4.75) shows it is suitably close to $\exp\big(N\int \log|\lambda|\,[\mu_{\xi,\vec{a}}](\mathsf{d}\lambda) \pm o_\iota(N)\big)$ and combining Lemmas 4.5.5 and 4.2.14 allows us to replace $\vec{a}$ by $\nabla_{\mathsf{rad}} H_N(\boldsymbol{x})$. It remains to integrate the other term, namely the origin density of $\nabla_{\mathsf{sp}} H_{N,\delta}(\boldsymbol{y})$ conditional on H_N:

$$(2\pi\delta)^{-\frac{N-r}{2}} \prod_{s\in\mathscr{S}} (\xi^s)^{\frac{N_s-1}{2}} \exp\left(\frac{-(1-\delta)}{2\delta}\|(\xi')^{-1/2} \diamond \nabla_{\mathsf{sp}} H_N(\boldsymbol{y})\|_2^2\right). \tag{4.91}$$

Using (4.83) with $\vec{v} = (\xi')^{-1/2}$ and $\|\boldsymbol{u}+\boldsymbol{v}\|_2^2 \le (1+\iota)\|\boldsymbol{u}\|_2^2 + (C/\iota)\|\boldsymbol{v}\|_2^2$ we get:

$$\|(\xi')^{-1/2} \diamond \nabla_{\mathsf{sp}} H_N(\boldsymbol{y})\|_2^2 \le (1+\iota)\|(\xi')^{-1/2} \diamond \nabla_{\mathsf{sp}}^2 H_N(\boldsymbol{x}) \cdot \gamma'_{\boldsymbol{x}\to\boldsymbol{y}}(0)\|_2^2 + \frac{C(\varepsilon^2+\alpha^4)}{\iota} N\,.$$

Recalling (4.91), we integrate over a subset of $\boldsymbol{y} \in B_{\alpha\sqrt{N}}(\boldsymbol{x})$ chosen as follows. Let $T_1(\boldsymbol{x})$ be the span of the eigenvectors of $J_N(\boldsymbol{x})$ (recall (4.80)) with eigenvalues inside $[-\eta,\eta]$, and $T_2(\boldsymbol{x})$ the orthogonal complement in the tangent space $T_x\mathcal{S}_N$. Since $H_N \in K_N(\varepsilon)$ and $\mu_{\xi,\vec{a}}$ is C_1-regular uniformly in C-bounded $\vec{a}$, we have $\dim(T_1(\boldsymbol{x})) \le O(N\eta)$.

Write $\widetilde{\gamma}'_{\boldsymbol{x}\to\boldsymbol{y}}(0) = (\xi')^{1/2} \diamond \gamma'_{\boldsymbol{x}\to\boldsymbol{y}}(0)$. Let $S(\boldsymbol{x})$ be the Minkowski sum of a radius $\alpha^2\sqrt{N/2}$ ball $S_1(\boldsymbol{x}) \subseteq T_1(\boldsymbol{x})$ and a radius $\alpha\sqrt{N/2}$ ball $S_2(\boldsymbol{x}) \subseteq T_2(\boldsymbol{x})$. We consider $\boldsymbol{y}$ for which $\widetilde{\gamma}'_{\boldsymbol{x}\to\boldsymbol{y}}(0) \in S(\boldsymbol{x})$ and accordingly write $\widetilde{\gamma}'_{\boldsymbol{x}\to\boldsymbol{y}}(0) = \boldsymbol{s}_1 + \boldsymbol{s}_2$ for $\boldsymbol{s}_i \in S_i(\boldsymbol{x})$. It is easy to see that the map $\boldsymbol{y} \to \gamma'_{\boldsymbol{x}\to\boldsymbol{y}}(0)$ is a diffeomorphism on $\boldsymbol{y} \in B_{\alpha\sqrt{N}}(\boldsymbol{x})$ with Jacobian determinant $1 \pm o_\alpha(1)$ uniformly. We will thus freely switch to integration over $\gamma'_{\boldsymbol{x}\to\boldsymbol{y}}(0)$, which is equivalent to integration over $\widetilde{\gamma}'_{\boldsymbol{x}\to\boldsymbol{y}}(0)$ after picking up a factor

$$\prod_{s\in\mathscr{S}} (\xi^s)^{(|\mathcal{I}_s|-1)/2} = e^{o(N)} \cdot \prod_{s\in\mathscr{S}} (\xi^s)^{\lambda_s N/2}.$$

Continuing,

$$\begin{aligned}\|J_N(\boldsymbol{x}) \cdot \widetilde{\gamma}'_{\boldsymbol{x}\to\boldsymbol{y}}(0)\|_2^2 &= \|J_N(\boldsymbol{x}) \cdot (\boldsymbol{s}_1 + \boldsymbol{s}_2)\|_2^2 \\ &= \|J_N(\boldsymbol{x}) \cdot \boldsymbol{s}_1\|_2^2 + \|J_N(\boldsymbol{x}) \cdot \boldsymbol{s}_2\|_2^2 \\ &\le C\eta^2\alpha^4 N + \|J_N(\boldsymbol{x}) \cdot \boldsymbol{s}_2\|_2^2.\end{aligned}$$

From (4.81), we have $\eta^2\alpha^4 \ll \delta$ so the first term will be negligible below. By definition of $S_2(\boldsymbol{x})$, the vector $\boldsymbol{v}_2 = J_N(\boldsymbol{x}) \cdot \boldsymbol{s}_2$ ranges over a superset of the ball of radius $\alpha\eta\sqrt{N}/C$ in $S_2(\boldsymbol{x})$. Since $\alpha\eta \gg \sqrt{\delta}$ from (4.81), this captures at least $1/2$ of the Gaussian integral mass for (4.91). Writing $S_{1,2}$ for the product $S_1(\boldsymbol{x}) \times S_2(\boldsymbol{x})$, we find that

$$
\begin{aligned}
&(2\pi\delta)^{-\frac{N-r}{2}} \prod_{s\in\mathscr{S}} (\xi^s)^{-\frac{N_s-1}{2}} \int\limits_{B_{\alpha\sqrt{N}}(\boldsymbol{x})} \exp\left(\frac{-(1-\delta)}{2\delta}\|(\xi')^{-1/2}\diamond\nabla_{\mathsf{sp}}H_N(\boldsymbol{y})\|_2^2\right)\mathsf{d}\boldsymbol{y} \\
&\geq (2\pi\delta)^{-\frac{N-r}{2}} \prod_{s\in\mathscr{S}} (\xi^s)^{-\lambda_s N} \\
&\quad\times \int\limits_{S_{1,2}} \exp\left(\frac{-(1+\iota)\|J_N(\boldsymbol{x})\cdot(\boldsymbol{s}_1+\boldsymbol{s}_2)\|_2^2 - C(\varepsilon^2+\alpha^4)N/\iota}{2\delta}\right)\mathsf{d}\boldsymbol{s}_2\mathsf{d}\boldsymbol{s}_1 \\
&= (2\pi\delta)^{-\frac{N-r}{2}} \prod_{s\in\mathscr{S}} (\xi^s)^{-\lambda_s N} \exp\left(-\frac{C(\varepsilon^2+\alpha^4)N}{2\delta\iota}\right) \\
&\quad\times \int\limits_{S_{1,2}} \exp\left(\frac{-\|(1+\iota)J_N(\boldsymbol{x})\cdot(\boldsymbol{s}_1+\boldsymbol{s}_2)\|_2^2}{2\delta}\right)\mathsf{d}\boldsymbol{s}_2\mathsf{d}\boldsymbol{s}_1 \\
&\geq (2\pi\delta)^{-\frac{N-r}{2}} \prod_{s\in\mathscr{S}} (\xi^s)^{-\lambda_s N} \exp\left(-\frac{C(\varepsilon^2+\alpha^4)N}{2\delta\iota} - \frac{(1+\iota)C\eta^2\alpha^4 N}{\delta}\right) \\
&\quad\times \int\limits_{S_{1,2}} \exp\left(\frac{-\|(1+\iota)J_N(\boldsymbol{x})\cdot\boldsymbol{s}_2\|_2^2}{2\delta}\right)\mathsf{d}\boldsymbol{s}_2\mathsf{d}\boldsymbol{s}_1 \\
&\geq \prod_{s\in\mathscr{S}} (\xi^s)^{-\lambda_s N}\exp(-\iota N)\mathsf{Vol}(S_1(\boldsymbol{x}))\left|\det\left(J_N(\boldsymbol{x})\big|_{S_2(\boldsymbol{x})}\right)^{-1}\right| \\
&\quad\times \int\limits_{\substack{\boldsymbol{v}_2\in S_2(\boldsymbol{x}),\\ \|\boldsymbol{v}_2\|_2\leq\alpha\eta\sqrt{N}/C}} (2\pi\delta)^{-\frac{N-r}{2}}\exp\left(\frac{-\|(1+\iota)\boldsymbol{v}_2\|_2^2}{2\delta}\right)\,\mathsf{d}\boldsymbol{v}_2 \\
&\geq \prod_{s\in\mathscr{S}} (\xi^s)^{-\lambda_s N}\exp(-\iota N)\mathsf{Vol}(S_1(\boldsymbol{x}))\left|\det\left(J_N(\boldsymbol{x})\big|_{S_2(\boldsymbol{x})}\right)^{-1}\right|(1+\iota)^{-\frac{N-r}{2}}/2.
\end{aligned}
\tag{4.92}
$$

Recalling (4.81) and that $\dim(T_1(\boldsymbol{x}))\leq O(N\eta)$, we find $\mathsf{Vol}(S_1(\boldsymbol{x}))\geq\alpha^{O(N\eta)}\geq e^{-\iota N}$. Uniformly over $H_N\in K_N(\varepsilon)$ and $\boldsymbol{x}\in\mathcal{S}_N$, from Definition 4.5.9 part (c) we have

$$
\lim_{\delta\to 0}\lim_{N\to\infty}\left|\frac{1}{N}\log\det\left(J_N(\boldsymbol{x})\big|_{S_2(\boldsymbol{x})}\right) - \left(\int\log|\lambda|\,[\mu_{\xi,\nabla_{\mathsf{rad}}H_N(\boldsymbol{x})}](\mathsf{d}\lambda) - \sum_{s\in\mathscr{S}}\lambda_s\log(\xi^s)\right)\right| = 0.
$$

Indeed up to factors of $e^{o_\delta(N)}$, uniformly over these sets we have

$$
\begin{aligned}
\det\left(J_N(\boldsymbol{x})\big|_{S_2(\boldsymbol{x})}\right) &\approx \mathbb{E}\left[\det\left(J_N(\boldsymbol{x})\right)\;\middle|\;\nabla_{\mathsf{rad}}H_N(\boldsymbol{x})\right] \\
&= \mathbb{E}\left[\det\left(\nabla^2_{\mathsf{sp}}(H_N(\boldsymbol{x}))\right)\;\middle|\;\nabla_{\mathsf{rad}}H_N(\boldsymbol{x})\right]\prod_{s\in\mathscr{S}}(\xi^s)^{-\lambda_s N} \\
&\approx \exp\left(N\int\log|\lambda|\,[\mu_{\xi,\nabla_{\mathsf{rad}}H_N(\boldsymbol{x})}](\mathsf{d}\lambda)\right)\prod_{s\in\mathscr{S}}(\xi^s)^{-\lambda_s N}.
\end{aligned}
$$

Thus (4.92) equals

$$
e^{-o_\delta(N)}\exp\left(-N\int\log|\lambda|\,[\mu_{\xi,\nabla_{\mathsf{rad}}H_N(\boldsymbol{x})}](\mathsf{d}\lambda)\right).
$$

This cancels the expected determinant given approximately by (4.90), leaving $e^{-o_\delta(N)}$ and completing the proof. □

Remark 4.5.14. We restricted attention to small $\|\boldsymbol{s}_1\|_2$ above because this causes $\|\nabla^2_{\mathsf{sp}}H_N(\boldsymbol{x})\cdot\boldsymbol{s}_1\|_2^2$ to be negligible. The trade-off is that $\mathsf{Vol}(S_1(\boldsymbol{x}))$ becomes smaller. However since $\dim(T_1(\boldsymbol{x}))\leq O(N\eta)$, this volumetric factor is also irrelevant because all small parameters were polynomially related.

Using Lemma 4.5.13, we now deduce Theorem 4.5.2 from the start of this section.

Proof of Theorem 4.5.2. Let $K_N(\varepsilon, \upsilon)$ consist of those $H_N \in K_N(\varepsilon)$ such that $|\mathsf{Crt}_N^{\overline{\mathcal{D}},\varepsilon,\upsilon}(H_N)| \geq 1$. Let $H_N \in K_N(\varepsilon,\upsilon)$ and let $\boldsymbol{x} \in \mathsf{Crt}_N^{\overline{\mathcal{D}},\varepsilon,\upsilon}(H_N)$ be an ε-critical point satisfying (4.71). For ε small enough that (recalling (4.81)) $\iota \leq \upsilon/C$, we claim that Lemma 4.5.13 implies

$$\mathbb{E}\left[\left|\mathsf{Crt}_N^{\overline{\mathcal{D}},0,\upsilon/2}(H_N)\right| \mid H_N\right] \geq e^{-o_\varepsilon(N)} \cdot \mathbf{1}\{H_N \in K_N(\varepsilon,\upsilon)\}. \tag{4.93}$$

Indeed (4.87) and (4.85) handle the radial derivative and energy errors between $\boldsymbol{x}$ and $\boldsymbol{y}$. Proposition 4.2.5 is used exploit the hypothesis (4.88); the resulting $\mathbb{W}_\infty$ estimate controls both the distances in $\mathcal{J}$ (Hausdorff distance between the spectral supports) and in $\mathbb{W}_1(\mathbb{R})$. Finally the overlap with $\boldsymbol{G}^{(1)}$ is controlled by the simple bound

$$\boldsymbol{y} \in B_{\alpha\sqrt{N}}(\boldsymbol{x}) \cap \mathcal{S}_N \quad \Longrightarrow \quad \|\vec{R}(\boldsymbol{x},\boldsymbol{G}^{(1)}) - \vec{R}(\boldsymbol{y},\boldsymbol{G}^{(1)})\| \leq O(\alpha) \ll \upsilon,$$

where we used that $H_N \in K_N(\varepsilon) \subseteq K_N$ implies $\|\boldsymbol{G}^{(1)}\| \leq O(\|\nabla H_N(\mathbf{0})\|) \leq O(\sqrt{N})$.

Averaging (4.93) over H_N and applying the hypothesis (4.72), we find

$$e^{-c_0 N} \geq \mathbb{E}\left|\mathsf{Crt}_N^{\overline{\mathcal{D}},0,\upsilon/2}(H_N)\right| \geq e^{-o_\varepsilon(N)} \cdot \mathbb{P}[H_N \in K_N(\varepsilon,\upsilon)].$$

Choosing ε to also be sufficiently small depending on c_0 and rearranging yields $\mathbb{P}[H_N \in K_N(\varepsilon,\upsilon)] \leq e^{-cN}$. Recalling that $\mathbb{P}[H_N \in K_N(\varepsilon)] \geq 1 - e^{-cN}$ by Proposition 4.5.10 completes the proof. $\square$

4.5.3 Failure of annealed topological trivialization for sub-solvable ξ

We showed in Section 4.4 that annealed topological trivialization occurs for super-solvable ξ (recall Definition 4.1.6). In this subsection we prove the strictly sub-solvable case (c) of Theorem 4.1.11, which is equivalent by Proposition 4.3.2 to the following. Recall the reparameterized form $\overline{F}$ of the complexity functional F defined in Subsection 4.4.1.

Proposition 4.5.15. *If ξ is strictly sub-solvable, then* $\sup_{\vec{v}\in\mathbb{R}^r} \overline{F}(\vec{v}) > 0$.

We will require a computation from our concurrent paper [HS24] as well as Theorem 4.5.2. The point is that in [HS24], we gave an explicit algorithm to construct approximate critical points for H_N whenever ξ is strictly sub-solvable. Applying Theorem 4.5.2 then shows $\overline{F}$ is non-negative at the radial derivative of such points (which is computed explicitly therein). While we cannot show this input makes $\overline{F}$ strictly positive, we do show it is not a stationary point of $\overline{F}$, which suffices.

The algorithm from [HS24] relies on a certain coordinate-wise increasing C^1 path $\Phi : [0,1] \to [0,1]^r$. In short, it proceeds by outward exploration in $\mathbb{R}^N$ starting from $\mathbf{0}$, greedily optimizing H_N in each step similarly to [Sub21a] (though implemented with approximate message passing as in [Mon21, AMS21, Sel24a]). The function Φ determines the schedule at which exploration occurs in the r species (which is trivial in the single-species setting). The optimal choice of Φ obeys stationarity conditions, which were established in [HS23a] and exploited in [HS24]. In particular the optimal Φ exhibits a phase transition when ξ shifts from super-solvable to sub-solvable (which in fact motivated the present paper).

Φ does not seem to have an explicit formula in the sub-solvable case, but is given by any maximizer of a $(\xi,\vec{\lambda})$-dependent functional $\mathbb{A}$ defined in [HS23a, Equation (1.6)]. (As explained in Remark 4.1.9, the fact that these companion results technically use deterministic external field instead of Gaussian $\boldsymbol{G}^{(1)}$ is inconsequential.) Φ satisfies the normalization $\langle\vec{\lambda},\Phi'(q)\rangle = 1$ for all $q \in [0,1]$, and moreover $\Phi_s'(1) > 0$ for all $s \in \mathscr{S}$. The input from [HS24] is as follows.

Proposition 4.5.16 ([HS24, Proposition 3.3])**.** *For non-degenerate and strictly super-solvable ξ, and Φ as above, and any $\varepsilon > 0$, with probability $1 - e^{-cN}$ there exists an ε-approximate critical point $\boldsymbol{x}_* \in \mathcal{S}_N$ such that*

$$\|\Lambda^{1/2}\nabla_{\mathsf{rad}} H_N(\boldsymbol{x}_*) - \vec{v}_*(\Phi)\|_2 \leq \varepsilon, \tag{4.94}$$

where $\vec{v}_*(\Phi) = (v_{*,s}(\Phi))_{s\in\mathscr{S}}$ *is given by*

$$v_{*,s}(\Phi) = \lambda_s f_s^{-1} + \sum_{s'\in\mathscr{S}} \xi''_{s,s'} f_{s'}\,; \tag{4.95}$$

$$f_s = \sqrt{\frac{\Phi'_s(1)}{(\xi^s\circ\Phi)'(1)}}. \tag{4.96}$$

Proposition 4.5.15 is a direct consequence of the following result.

Theorem 4.5.17. *If ξ is non-degenerate and strictly sub-solvable, then for any Φ as above, we have $\overline{F}(\vec{v}_*) \geq 0$ and $\nabla\overline{F}(\vec{v}_*) \neq \vec{0}$. Hence $\sup_{\vec{v}\in\mathbb{R}^r}\overline{F}(\vec{v}) > 0$, i.e. the annealed complexity is strictly positive.*

Proof. We apply Theorem 4.5.2 with

$$\overline{\mathcal{D}} = \{\Lambda^{-1/2}\vec{v}_*(\Phi)\} \times \mathbb{R}^r \times \mathbb{R} \times \mathcal{J} \times \mathbb{W}_1(\mathbb{R}).$$

(I.e. we consider only the radial derivative and ignore the remaining components of $\overline{\mathcal{D}}$.) The high-probability existence of $\boldsymbol{x}_*$ obeying (4.94) for arbitrarily small ε, combined with continuity of $\overline{F}$, yields $\overline{F}(\vec{v}_*) \geq 0$.

Let $\vec{u}_* = -\vec{f}$. We claim that $\vec{u}(0;\vec{v}_*) = \vec{u}_*$. Due to the formula (4.95), $\vec{u}_*$ satisfies (4.64), so by Lemma 4.4.8 case (i) it suffices to check $M(\vec{u}_*) \succeq 0$. This follows by Lemma 4.2.7, as $\Phi'(1) \succ 0$ and $M(\vec{u}_*)\Phi'(1) = \vec{0}$ by inspection.

Next, Lemma 4.4.5 implies that stationary points with $\vec{u}\in\mathbb{R}^r$ satisfy $u_s^2 = \frac{1}{\xi^s(\vec{1})}$. Hence if $\vec{v}_*$ were stationary, rearranging using the definition (4.96) of f_s would directly yield

$$\begin{aligned} f_s^2 = \frac{\Phi'_s(1)}{(\xi^s\circ\Phi)'(1)} &= \frac{1}{\xi^s(\vec{1})} \quad \forall s\in\mathscr{S} \\ \implies \big(\mathrm{diag}(\xi') - \xi''\big)\Phi'(1) &= 0. \end{aligned}$$

Since $\Phi'(1) \succ 0$, Lemma 4.2.7 implies $\mathrm{diag}(\xi') - \xi'' \succeq 0$, contradicting that ξ is strictly sub-solvable.

We conclude that $\nabla\overline{F}(\vec{v}_*) \neq \vec{0}$. Combined with the fact that $\overline{F}(\vec{v}_*) \geq 0$ immediately yields $\sup_{\vec{v}\in\mathbb{R}^r}\overline{F}(\vec{v}) > 0$ as desired. □

Remark 4.5.18. We expect that for all (or at least almost all) strictly sub-solvable ξ one has

$$\overline{F}(\vec{v}_*) > \overline{F}(\vec{v}_*, E_*) > 0,$$

where $E_* = \mathbb{A}(\Phi) \approx H_N(\boldsymbol{x}_*)/N$ is the associated energy of Φ as described in [HS23a, Equation (1.6)].[2] However both inequalities seem much more involved to prove. Given the branching tree construction of algorithmic maximizers in [HS24], it is natural to speculate that $\overline{F}(\vec{v}_*, E_*)$ strictly increases along the tree-descending part $q\in[q_0,1]$ of any Φ satisfying the conclusions of [HS23a, Theorem 3].

Remark 4.5.19. As we recalled in Proposition 4.1.14, our work [HS24] actually constructs $\exp(\delta N)$ points $\boldsymbol{x}_*$ satisfying the conditions of Proposition 4.5.16, with all pairwise distances at least $\sqrt{N}/C(\xi)$. If one had ε sufficiently small given δ, then Theorem 4.5.2 would imply that $\overline{F}(\vec{v}_*) > 0$. However [HS24] only guarantees $\delta > 0$ is positive for each ε, which does not yield strict inequality. On the other hand as explained in the introduction, it does imply *quenched* failure of *strong* topological trivialization.

Finally we show that $\vec{v}_*$ above corresponds to the top of the bulk spectrum of $\nabla^2_{\mathsf{sp}}H_N(\boldsymbol{x}_*)$ equalling zero. Informally, this means $\boldsymbol{x}_*$ is on the verge of being a local maximum. More formally, it is an ε-marginal local maximum as defined in Section 4.7.

Proposition 4.5.20. *For $\vec{v}_*$ as above,* $\max\mathsf{supp}(\overline{\mu}_\xi(\vec{v}_*)) = 0$.

[2] As written therein $E_* = \mathbb{A}(p,\Phi;q_0)$; both $p:[0,1]\to[0,1]$ and $q_0\in[0,1]$ are implicitly determined by Φ via [HS23a, Theorem 3].

Proof. Consider a path $\vec{u}(t) = (1+t)^{-1}\vec{u}_*$ for $t \geq 0$ and with $u_{*,s} = -f_s$ as above. We verified above that $M(\vec{u}_*) \succeq 0$, and from the definition (4.65) of M it follows that $M(\vec{u}(t)) \succeq 0$ for all $t \geq 0$. By Lemma 4.4.8 this means that for all $t \geq 0$, $\vec{u}(t) = \vec{u}(0; \vec{v}(t))$ for some $\vec{v}(t) \in \mathbb{R}^r$.

For t sufficiently large, (4.59) yields $v_s(t) \leq -\lambda_s/u_s \leq -C$ for large $C = C(\xi) > 0$. At this point, (4.46) implies

$$\mathbb{W}_\infty\Big(\overline{\mu}(\vec{v}(t)), \sum_{s\in\mathscr{S}} \lambda_s \delta_{-v_s/\lambda_s}\Big) \leq C$$

whence $\max \mathsf{supp}(\overline{\mu}(\vec{v}(t)) < 0$.

Since $\vec{v}(t) \in \mathbb{R}^r$, it follows that $\overline{\mu}(\vec{v}(t))$ always has density 0 at 0. By a continuity argument via Lemma 4.2.12, if $\max \mathsf{supp}(\overline{\mu}(\vec{v}_*)) > 0$ held, then there would exist t such that $\max \mathsf{supp}(\overline{\mu}(\vec{v}(t)) = \delta$ for arbitrarily small $\delta > 0$. It follows by [AEK19a, Eq. (2.15)] that $\overline{\mu}(\vec{v}(t))$ must have positive density at 0 for such t when δ is taken sufficiently small, a contradiction.

For the opposite direction, suppose that $\max \mathsf{supp}(\overline{\mu}(\vec{v}_*)) < 0$. Then $0 \notin \mathsf{supp}(\overline{\mu}(\vec{v}_*))$, which implies that $M(\vec{u}_*)$ is invertible by Proposition 4.A.7 or Lemma 4.A.25. However $M(\vec{u}_*)\Phi'(1) = \vec{0}$ so this cannot hold. $\square$

4.6 Locating the critical points

In this section, we complete the proof of Theorem 4.1.13 by combining the description of ε-approximate critical points from Proposition 4.5.1 with a recursive argument that localizes all approximate critical points. Throughout this section we assume ξ is strictly super-solvable.

In Subsection 4.6.1, we define the *type* $\vec{\Delta} \in \{-1,1\}^r$ of an approximate critical point based on its radial derivative. It follows by the previous section that with high probability, all critical points have a well-defined type. Next in Subsection 4.6.2 we explain the conditional law of H_N on subspherical bands. This lets us analyze the recursive algorithm of Subsection 4.6.3. In Subsection 4.6.4 we deduce that all approximate critical points of each type $\vec{\Delta}$ are localized inside a single small (random) subset of $\mathcal{S}_N$. Subsection 4.6.5 uses this to deduce existence and uniqueness of type of (exact) critical point. Finally Subsection 4.6.6 determines the exact index of each critical point by gradually perturbing ξ and arguing that eigenvalues do not cross 0.

4.6.1 Critical points of type $\vec{\Delta}$

Our argument will separately localize each critical point of type $\vec{\Delta} \in \{-1,1\}^r$, defined as follows.

Definition 4.6.1. Let $\upsilon = o_\varepsilon(1)$ be given by Proposition 4.5.1. Say $\boldsymbol{x} \in \mathcal{S}_N$ is a ε**-critical point of type** $\vec{\Delta}$, or alternatively a $(\varepsilon, \vec{\Delta})$**-critical point**, if it is an ε-critical point (recall Definition 4.1.5) and

$$\|\nabla_{\mathsf{rad}} H_N(\boldsymbol{x}) - \vec{x}(\vec{\Delta})\|_\infty \leq \upsilon\,. \tag{4.97}$$

Fact 4.6.2. *There exists $\varepsilon_0 = \varepsilon_0(\xi)$ such that with probability $1 - e^{-cN}$ the following holds. For all $\varepsilon \leq \varepsilon_0$, all ε-critical points of H_N are $(\varepsilon, \vec{\Delta})$-critical points for a unique $\vec{\Delta} \in \{-1,1\}^r$.*

Proof. Immediate from Proposition 4.5.1. The signs $\vec{\Delta}$ are unique since for small ε, the υ-balls around the $\vec{x}(\vec{\Delta})$ are disjoint. $\square$

Definition 4.6.3 (Species-wise rescaling)**.** Let $\boldsymbol{v} \in \mathbb{R}^N$ such that $\boldsymbol{v}_s \neq \boldsymbol{0}$ for all $s \in \mathscr{S}$. For $\vec{q} \in [0,1]^r$, $\vec{\Delta} \in \{-1,1\}^r$, define $\mathsf{scale}(\boldsymbol{v}; \vec{\Delta}, \vec{q})$ to be the vector $\boldsymbol{u} \in \mathbb{R}^N$ with

$$\boldsymbol{u}_s = \Delta_s \sqrt{q_s \lambda_s N} \frac{\boldsymbol{v}_s}{\|\boldsymbol{v}_s\|_2}\,.$$

That is, $\boldsymbol{u}$ is the vector parallel to $\boldsymbol{v}$ in each species with $\vec{R}(\boldsymbol{u},\boldsymbol{u}) = \vec{q}$, whose species-$s$ component is correlated (resp. anti-correlated) with that of $\boldsymbol{v}$ if $\Delta_s = 1$ (resp. -1). The following corollary of Proposition 4.5.1 shows that $(\varepsilon, \vec{\Delta})$-critical points have nearly constant correlation with $\nabla H_N(\boldsymbol{0}) = \Gamma^{(1)} \diamond \boldsymbol{G}^{(1)}$, the 1-spin part of H_N.

Corollary 4.6.4. *For any $\varepsilon > 0$, there exists $\upsilon = o_\varepsilon(1)$ such that with probability $1 - e^{-cN}$ the following holds. For any $(\varepsilon, \vec{\Delta})$-critical point $\boldsymbol{x}$, let $\boldsymbol{y}$ be its species-wise projection onto $\nabla H_N(\mathbf{0})$, i.e.*

$$\boldsymbol{y}_s = \frac{\langle \boldsymbol{x}_s, (\nabla H_N(\mathbf{0}))_s \rangle}{\|(\nabla H_N(\mathbf{0}))_s\|_2^2} (\nabla H_N(\mathbf{0}))_s$$

for all $s \in \mathscr{S}$. Then

$$\left\| \boldsymbol{y} - \mathsf{scale}\left(\nabla H_N(\mathbf{0}); \vec{\Delta}, \nabla\xi(\vec{0})/\nabla\xi(\vec{1}) \right) \right\|_2 \le \upsilon\sqrt{N}\,.$$

Proof. We can write

$$\boldsymbol{y}_s = \frac{\langle \boldsymbol{x}_s, \boldsymbol{G}^{(1)}_s \rangle}{\|\boldsymbol{G}^{(1)}_s\|_2^2} \boldsymbol{G}^{(1)}_s = \frac{\langle \boldsymbol{x}_s, \boldsymbol{G}^{(1)}_s \rangle}{\|\boldsymbol{G}^{(1)}_s\|_2} \cdot \frac{\boldsymbol{G}^{(1)}_s}{\|\boldsymbol{G}^{(1)}_s\|_2}\,.$$

By Proposition 4.5.1, with probability $1-e^{-cN}$ all $(\varepsilon, \vec{\Delta})$-critical point $\boldsymbol{x}$ satisfies $\langle \boldsymbol{x}_s, \boldsymbol{G}^{(1)}_s \rangle = \frac{\Delta_s \gamma_s}{\sqrt{\xi'_s}} \cdot \lambda_s N(1+o_\varepsilon(1))$, and by a standard concentration bound $\|\boldsymbol{G}^{(1)}_s\|_2 = \sqrt{\lambda_s N}(1+o_\varepsilon(1))$. So up to $1+o_\varepsilon(1)$ multiplicative error

$$\boldsymbol{y}_s = \frac{\Delta_s \gamma_s}{\sqrt{\xi'_s}} \cdot \sqrt{\lambda_s N} \frac{\boldsymbol{G}^{(1)}_s}{\|\boldsymbol{G}^{(1)}_s\|_2} = \Delta_s \sqrt{\frac{\partial_s \xi(\vec{0})}{\partial_s \xi(\vec{1})} \lambda_s N} \frac{\boldsymbol{G}^{(1)}_s}{\|\boldsymbol{G}^{(1)}_s\|_2}\,.$$

The result follows because

$$\mathsf{scale}\left(\nabla H_N(\mathbf{0}); \vec{\Delta}, \nabla\xi(\vec{0})/\nabla\xi(\vec{1}) \right) = \mathsf{scale}\left(\boldsymbol{G}^{(1)}; \vec{\Delta}, \nabla\xi(\vec{0})/\nabla\xi(\vec{1}) \right)\,. \qquad \square$$

4.6.2 Conditional band models

Our arguments rely on a self-similarity in law obtained by restriction to a band. The point is that bands inside $\mathcal{S}_N$ are still of the same form as our original model, with N_s replaced by $N_s - 1$ and ξ replaced by a new mixture function. This lets us apply the preceding results of this paper to said bands. This idea has been used extensively in recent work by Subag, e.g. [Sub24, Sub21b, Sub23b].

For the below definitions, U is a species-aligned subspace (recall Definition 4.2.1), which in our applications will always be of dimension $O(1)$. Let

$$\mathcal{B}_N = \left\{ \boldsymbol{x} \in \mathbb{R}^N : \|\boldsymbol{x}_s\|_2^2 \le \lambda_s N \quad \forall s \in \mathscr{S} \right\}$$

be the convex hull of $\mathcal{S}_N$. Over the course of the recursive argument, we will be interested in the landscape of various Hamiltonians in the following domains where we project out the subspace U. The original model corresponds to $U = \emptyset$.

Definition 4.6.5. Let $\mathcal{S}_N^U = \mathcal{S}_N \cap U^\perp$ and $\mathcal{B}_N^U = \mathcal{B}_N \cap U^\perp$ (recalling the notation (4.25)).

Definition 4.6.6. Let $\boldsymbol{m} \in \mathcal{B}_N^U$ such that $\boldsymbol{m}_s \neq \mathbf{0}$ for all $s \in \mathscr{S}$. The band of $\mathcal{S}_N^U$ centered at $\boldsymbol{m}$ is

$$\mathsf{Band}^U(\boldsymbol{m}) = \left\{ \boldsymbol{\sigma} \in \mathcal{S}_N^U : \vec{R}(\boldsymbol{\sigma} - \boldsymbol{m}, \boldsymbol{m}) = \vec{0} \right\}\,.$$

Note that for $\vec{q} = \vec{R}(\boldsymbol{m}, \boldsymbol{m}) \in [0,1]^r$ and $U \bowtie \boldsymbol{m} = \mathsf{span}(U, \boldsymbol{m}_1, \ldots, \boldsymbol{m}_s)$,

$$\mathsf{Band}^U(\boldsymbol{m}) = (\vec{1} - \vec{q})^{1/2} \diamond \mathcal{S}_N^{U \bowtie \boldsymbol{m}} + \boldsymbol{m}\,. \tag{4.98}$$

We will be interested in the following bands, whose centers are rescalings of $\nabla H_N(\mathbf{0})$ (projected to $U^\perp$).

Definition 4.6.7. For $\vec{q} \in [0,1]^r$ and $\vec{\Delta} \in \{-1,1\}^r$, define

$$\boldsymbol{m}^U_{\vec{\Delta},\vec{q}}(H_N) = \mathsf{scale}(\mathsf{proj}_{U^\perp} \nabla H_N(\mathbf{0}); \vec{\Delta}, \vec{q})\,, \qquad \mathsf{Band}^U_{\vec{\Delta},\vec{q}}(H_N) = \mathsf{Band}^U(\boldsymbol{m}^U_{\vec{\Delta},\vec{q}}(H_N))\,.$$

We will abbreviate these $\boldsymbol{m}^U_{\vec{\Delta},\vec{q}}$ and $\mathsf{Band}^U_{\vec{\Delta},\vec{q}}$ when H_N is clear. The following corollary shows that all critical points of H_N in $\mathcal{S}_N^U$ lie near one of these bands, given by a specific $\vec{q}$.

Corollary 4.6.8. *There exists $\upsilon = o_\varepsilon(1)$ such that the following holds. Let U be a species-aligned subspace of dimension $O(1)$. With probability $1-e^{-cN}$, all $(\varepsilon,\vec{\Delta})$-critical points of the restriction of H_N to the manifold $\mathcal{S}_N^U$ lie within $\upsilon\sqrt{N}$ of $\mathsf{Band}^U_{\vec{\Delta},\vec{q}}(H_N)$, where $\vec{q} = \nabla\xi(\vec{0})/\nabla\xi(\vec{1})$.*

Proof. Immediate from Corollary 4.6.4. Since U has dimension $O(1)$, H_N restricted to $\mathcal{S}_N^U$ is a multi-species spin glass whose species dimensions $N_s' = N_s - O(1)$ still satisfy $N_s'/N \to \lambda_s$. Thus restricting the model to $\mathcal{S}_N^U$ does not affect the result. $\square$

Finally let $U' = U \bowtie \boldsymbol{m}^U_{\vec{\Delta},\vec{q}}$. We define the centered band Hamiltonian $H^U_{N,\vec{\Delta},\vec{q}}(\boldsymbol{\sigma}) : \mathcal{S}_N^{U'} \to \mathbb{R}$ by

$$H^U_{N,\vec{\Delta},\vec{q}}(\boldsymbol{\sigma}) = H_N\left((\vec{1}-\vec{q})^{1/2}\diamond\boldsymbol{\sigma} + \boldsymbol{m}^U_{\vec{\Delta},\vec{q}}\right) - \left\langle \mathsf{proj}_{U^\perp}\nabla H_N(\mathbf{0}), (\vec{1}-\vec{q})^{1/2}\diamond\boldsymbol{\sigma} + \boldsymbol{m}^U_{\vec{\Delta},\vec{q}}\right\rangle . \tag{4.99}$$

The following lemma shows that conditional on $\mathsf{proj}_{U^\perp}\nabla H_N(\mathbf{0})$, $H^U_{N,\vec{\Delta},\vec{q}}$ is itself a multi-species spin glass. Note that the last term of (4.99) is constant for $\boldsymbol{\sigma}\in\mathcal{S}_N^{U'}$, and that $(\vec{1}-\vec{q})^{1/2}\diamond\boldsymbol{\sigma}+\boldsymbol{m}^U_{\vec{\Delta},\vec{q}}$ ranges over $\mathsf{Band}^U_{\vec{\Delta},\vec{q}}$ as $\boldsymbol{\sigma}$ ranges over $\mathcal{S}_N^{U'}$. Thus $H^U_{N,\vec{\Delta},\vec{q}}$ is the remaining randomness of H_N on this band.

Lemma 4.6.9. *Conditionally on $\mathsf{proj}_{U^\perp}\nabla H_N(\mathbf{0})$, $H^U_{N,\vec{\Delta},\vec{q}}$ is a centered Gaussian process on $\mathcal{S}_N^{U'}$ with covariance*

$$\mathbb{E}[H^U_{N,\vec{\Delta},\vec{q}}(\boldsymbol{\sigma})H^U_{N,\vec{\Delta},\vec{q}}(\boldsymbol{\rho})] = N\xi_{\vec{q}}(\vec{R}(\boldsymbol{\sigma},\boldsymbol{\rho}))\,, \qquad \textit{where}$$
$$\xi_{\vec{q}}(\vec{x}) = \xi\left((\vec{1}-\vec{q})\odot\vec{x}+\vec{q}\right) - \left\langle\nabla\xi(\vec{0}), (\vec{1}-\vec{q})\odot\vec{x}+\vec{q}\right\rangle .$$

Proof. For $\boldsymbol{\sigma}\in\mathcal{S}_N^U$, we have

$$H_N(\boldsymbol{\sigma}) = \langle \mathsf{proj}_{U^\perp}\nabla H_N(\mathbf{0}),\boldsymbol{\sigma}\rangle + H_{N,\geq 2}(\boldsymbol{\sigma})\,,$$

where $H_{N,\geq 2}$ consists of the interactions of degree at least 2. These two constituent functions are, respectively, $\mathsf{proj}_{U^\perp}\nabla H_N(\mathbf{0})$-measurable and independent of $\mathsf{proj}_{U^\perp}\nabla H_N(\mathbf{0})$, while $\boldsymbol{m}^U_{\vec{\Delta},\vec{q}}$ and $\mathsf{Band}^U_{\vec{\Delta},\vec{q}}$ are both $\mathsf{proj}_{U^\perp}\nabla H_N(\mathbf{0})$-measurable. Since

$$H^U_{N,\vec{\Delta},\vec{q}}(\boldsymbol{\sigma}) = H_{N,\geq 2}\left((\vec{1}-\vec{q})^{1/2}\diamond\boldsymbol{\sigma}+\boldsymbol{m}^U_{\vec{\Delta},\vec{q}}\right),$$

this is a centered Gaussian process. Moreover, $H_{N,\geq 2}$ has covariance

$$\mathbb{E}[H_{N,\geq 2}(\boldsymbol{\sigma})H_{N,\geq 2}(\boldsymbol{\rho})] = N\xi_{\geq 2}(\vec{R}(\boldsymbol{\sigma},\boldsymbol{\rho}))\,, \qquad \text{where} \qquad \xi_{\geq 2}(\vec{x}) = \xi(\vec{x}) - \langle\nabla\xi(\vec{0}),\vec{x}\rangle\,.$$

The covariance formula for $H^U_{N,\vec{\Delta},\vec{q}}$ now follows because for $\boldsymbol{\sigma},\boldsymbol{\rho}\in\mathcal{S}_N^{U'}$, we have $\boldsymbol{\sigma},\boldsymbol{\rho}\in(\boldsymbol{m}^U_{\vec{\Delta},\vec{q}})^\perp$, and so

$$\vec{R}\left((\vec{1}-\vec{q})^{1/2}\diamond\boldsymbol{\sigma}+\boldsymbol{m}^U_{\vec{\Delta},\vec{q}}, (\vec{1}-\vec{q})^{1/2}\diamond\boldsymbol{\rho}+\boldsymbol{m}^U_{\vec{\Delta},\vec{q}}\right) = (\vec{1}-\vec{q})\diamond\vec{R}(\boldsymbol{\sigma},\boldsymbol{\rho})+\vec{q}. \qquad \square$$

4.6.3 Recursive algorithm

We now consider a recursive *critical point finding* algorithm. Roughly speaking, Corollary 4.6.8 shows that all $(\varepsilon,\vec{\Delta})$-critical points of H_N lie near a band $\mathsf{Band}_1(\vec{\Delta}) = \mathsf{Band}^{\emptyset}_{\vec{\Delta},\vec{q}}$, for a deterministic $\vec{q}$. Lemma 4.6.9 shows that H_N restricted to this band is conditionally another multi-species spin glass. So, Corollary 4.6.8 implies all $(\varepsilon,\vec{\Delta})$-critical points lie near a sub-band $\mathsf{Band}_2(\vec{\Delta})\subseteq\mathsf{Band}_1(\vec{\Delta})$. Repeating this argument, all $(\varepsilon,\vec{\Delta})$-critical points of H_N lie near a nested sequence of bands $\mathcal{S}_N\supseteq\mathsf{Band}_1(\vec{\Delta})\supseteq\mathsf{Band}_2(\vec{\Delta})\cdots$, and we will show these bands' diameters shrink to 0. After a large constant number of recursions, this shows all $(\varepsilon,\vec{\Delta})$-critical points lie in a region of diameter $o_\varepsilon(\sqrt{N})$, and the well-conditionedness of $\nabla^2_{\mathsf{sp}}H_N$ (by Proposition 4.5.1) shows there is at most one critical point in this region.

We now define the recursive bands, starting with sequence of radii of their centers. Define $\vec{R}^0 = \vec{0}$ and recursively

$$\vec{R}^{k+1} = \nabla\xi(\vec{R}^k)/\nabla\xi(\vec{1})\,.$$

Because ξ is coordinate-wise increasing, the sequence $\vec{R}^k$ is coordinate-wise increasing up to some limit in $[0,1]^r$. The following lemma, which relies on super-solvability of ξ, shows this limit is $\vec{1}$. That is, the band centers approach the surface $\mathcal{S}_N$ of $\mathcal{B}_N$ and the band diameters limit to zero.

Lemma 4.6.10 ([HS24, Lemma 2.3])**.** *We have that* $\lim_{k\to\infty} \vec{R}^k = \vec{1}$.

Fix $\vec{\Delta} \in \{-1,1\}^r$. Let $\boldsymbol{m}^0(\vec{\Delta}) = \boldsymbol{0}$ and $\mathsf{Band}_0(\vec{\Delta}) = \mathcal{S}_N$. Recursively for $k \geq 1$ define

$$
\begin{aligned}
\boldsymbol{m}^k(\vec{\Delta}) &= \boldsymbol{m}^{k-1}(\vec{\Delta}) + \mathsf{scale}(\boldsymbol{g}^{k-1}(\vec{\Delta}); \vec{\Delta}, \vec{R}^k - \vec{R}^{k-1})\,, \\
\boldsymbol{g}^k(\vec{\Delta}) &= \mathsf{proj}_{U_k(\vec{\Delta})^\perp} \nabla H_N(\boldsymbol{m}^k(\vec{\Delta}))\,, \\
U_k(\vec{\Delta}) &= \mathsf{span}(\boldsymbol{m}^j_s(\vec{\Delta}))_{1\leq j\leq k, s\in\mathscr{S}}\,, \\
\mathsf{Band}_k(\vec{\Delta}) &= \mathcal{S}_N \cap \big(\boldsymbol{m}^k(\vec{\Delta}) + U_k(\vec{\Delta})^\perp\big)\,.
\end{aligned}
\tag{4.100}
$$

Note that because $\boldsymbol{g}^k_s(\vec{\Delta}) \in U_k(\vec{\Delta})^\perp$ for all $s \in \mathscr{S}$, we have $\boldsymbol{m}^k(\vec{\Delta}) \in (\boldsymbol{m}^{k-1}(\vec{\Delta}) + U_{k-1}(\vec{\Delta})^\perp)$, so $\mathsf{Band}_k(\vec{\Delta}) \subseteq \mathsf{Band}_{k-1}(\vec{\Delta})$. Also, by induction $\vec{R}(\boldsymbol{m}^k(\vec{\Delta}), \boldsymbol{m}^k(\vec{\Delta})) = \vec{R}^k$ for each $k \geq 0$, so analogously to (4.98),

$$
\mathsf{Band}_k(\vec{\Delta}) = \phi_{k,\vec{\Delta}}(\mathcal{S}_N^{U_k(\vec{\Delta})})\,, \qquad \text{where} \qquad \phi_{k,\vec{\Delta}}(\boldsymbol{\sigma}) = (\vec{1} - \vec{R}^k)^{1/2} \diamond \boldsymbol{\sigma} + \boldsymbol{m}^k(\vec{\Delta})\,. \tag{4.101}
$$

Define the band Hamiltonian $\widetilde{H}_{N,\vec{\Delta},k} : \mathcal{S}_N^{U_k(\vec{\Delta})} \to \mathbb{R}$ by

$$
\widetilde{H}_{N,\vec{\Delta},k}(\boldsymbol{\sigma}) = H_N\left(\phi_{k,\vec{\Delta}}(\boldsymbol{\sigma})\right) - \left\langle \boldsymbol{g}^{k-1}(\vec{\Delta}), \phi_{k,\vec{\Delta}}(\boldsymbol{\sigma})\right\rangle + \sum_{i=1}^{k-1} \left\langle \boldsymbol{g}^i(\vec{\Delta}) - \boldsymbol{g}^{i-1}(\vec{\Delta}), \boldsymbol{m}^i(\vec{\Delta})\right\rangle\,, \tag{4.102}
$$

whose meaning is explained in Lemma 4.6.12 below. We first show that the bands $\mathsf{Band}_k(\vec{\Delta})$ can be constructed by applying the construction from Definition 4.6.7 recursively. The radius $\vec{q}^k$ of the band center in the next lemma is chosen to be near all $(\varepsilon, \vec{\Delta})$-critical points of the Hamiltonian restricted to $\mathsf{Band}_k(\vec{\Delta})$, as will be explained in Corollary 4.6.14 below.

Lemma 4.6.11. *Let* $\vec{q}^k = (\vec{R}^{k+1} - \vec{R}^k)/(\vec{1} - \vec{R}^k)$ *and note that* $\phi^{-1}_{k,\vec{\Delta}}$ *maps* $\mathsf{Band}_k(\vec{\Delta})$ *to* $\mathcal{S}_N^{U_k(\vec{\Delta})}$*. Then* $\phi^{-1}_{k,\vec{\Delta}}(\mathsf{Band}_{k+1}(\vec{\Delta})) = \mathsf{Band}^{U_k(\vec{\Delta})}_{\vec{\Delta},\vec{q}^k}(\widetilde{H}_{N,\vec{\Delta},k})$*, where the latter band is defined in Definition 4.6.7.*

Proof. Since $\phi^{-1}_{k,\vec{\Delta}}(\boldsymbol{\sigma}) = (1 - \vec{R}^k)^{-1/2} \diamond (\boldsymbol{\sigma} - \boldsymbol{m}_k(\vec{\Delta}))$, we have

$$
\begin{aligned}
\phi^{-1}_{k,\vec{\Delta}}(\boldsymbol{m}^{k+1}(\vec{\Delta})) &= (\vec{1} - \vec{R}^k)^{-1/2} \diamond \mathsf{scale}(\boldsymbol{g}^k(\vec{\Delta}); \vec{\Delta}, \vec{R}^{k+1} - \vec{R}^k) \\
&= \mathsf{scale}\left(\boldsymbol{g}^k(\vec{\Delta}); \vec{\Delta}, \vec{q}^k\right) = \boldsymbol{m}^{U_k(\vec{\Delta})}_{\vec{\Delta},\vec{q}^k}(\widetilde{H}_{N,\vec{\Delta},k})\,.
\end{aligned}
\tag{4.103}
$$

Let us denote this point $\boldsymbol{m}$. Moreover,

$$
\phi^{-1}_{k,\vec{\Delta}}(\mathsf{Band}_{k+1}(\vec{\Delta})) = \phi^{-1}_{k,\vec{\Delta}}(\phi_{k+1,\vec{\Delta}}(\mathcal{S}_N^{U_{k+1}(\vec{\Delta})})) = \left(\frac{\vec{1} - \vec{R}^k}{\vec{1} - \vec{R}^{k+1}}\right)^{1/2} \diamond \mathcal{S}_N^{U_{k+1}(\vec{\Delta})} + \boldsymbol{m}\,. \tag{4.104}
$$

Also,

$$
U_{k+1}(\vec{\Delta}) = U_k(\vec{\Delta}) \bowtie \boldsymbol{m}^k(\vec{\Delta}) = U_k(\vec{\Delta}) \bowtie \boldsymbol{g}^k(\vec{\Delta}) = U_k(\vec{\Delta}) \bowtie \boldsymbol{m}\,,
$$

so elements of $\mathcal{S}_N^{U_{k+1}(\vec{\Delta})}$ are orthogonal to $\boldsymbol{m}$. This implies the conclusion. □

Let $F_k(\vec{\Delta}) = (\boldsymbol{g}^0(\vec{\Delta}), \boldsymbol{g}^1(\vec{\Delta}), \ldots \boldsymbol{g}^{k-1}(\vec{\Delta}))$. The following lemma shows that conditional on $F_k(\vec{\Delta})$, $\widetilde{H}_{N,\vec{\Delta},k}$ is a multi-species spin glass. Moreover, all terms on the right-hand side of (4.102) except $H_N(\phi_{k,\vec{\Delta}}(\boldsymbol{\sigma}))$ are constant on $\mathcal{S}_N^{U_k(\vec{\Delta})}$, so $\widetilde{H}_{N,\vec{\Delta},k}$ is the remaining randomness of H_N on $\mathcal{S}_N^{U_k(\vec{\Delta})}$.

Lemma 4.6.12. *Conditional on $F_k(\vec{\Delta})$, $\widetilde{H}_{N,\vec{\Delta},k}$ is a centered Gaussian process on $\mathcal{S}_N^{U_k(\vec{\Delta})}$ with covariance*

$$\mathbb{E}\left[\widetilde{H}_{N,\vec{\Delta},k}(\boldsymbol{\sigma})\widetilde{H}_{N,\vec{\Delta},k}(\boldsymbol{\rho})\right] = N\xi_k(\vec{R}(\boldsymbol{\sigma},\boldsymbol{\rho}))\,,$$

where

$$\begin{aligned}\xi_k(\vec{x}) = \xi\left((\vec{1}-\vec{R}^k)\odot\vec{x}+\vec{R}^k\right) &- \left\langle\nabla\xi(\vec{R}^{k-1}),(\vec{1}-\vec{R}^k)\odot\vec{x}+\vec{R}^k\right\rangle \\ &+ \sum_{i=1}^{k-1}\left\langle\nabla\xi(\vec{R}^i)-\nabla\xi(\vec{R}^{i-1}),\vec{R}^i\right\rangle\,.\end{aligned}$$

Proof. We induct on k. Assume the claim holds for k and let $\boldsymbol{m} = \mathsf{scale}(\boldsymbol{g}^k(\vec{\Delta});\vec{\Delta},\vec{q}^k)$ as in (4.103). Lemma 4.6.9 implies that conditional on $(F_k(\vec{\Delta}),P_{U_k(\vec{\Delta})^\perp}\nabla\widetilde{H}_{N,\vec{\Delta},k}(\mathbf{0}))$ the function $\widetilde{H}^\circ_{N,\vec{\Delta},k}:\mathcal{S}_N^{U_k(\vec{\Delta})\bowtie\boldsymbol{m}}\to\mathbb{R}$ given by

$$\begin{aligned}\widetilde{H}^\circ_{N,\vec{\Delta},k}(\boldsymbol{\sigma}) &= \widetilde{H}_{N,\vec{\Delta},k}\left((\vec{1}-\vec{q}^k)^{1/2}\diamond\boldsymbol{\sigma}+\boldsymbol{m}\right) \\ &\quad - \left\langle P_{U_k(\vec{\Delta})^\perp}\nabla\widetilde{H}_{N,\vec{\Delta},k}(\mathbf{0}),(\vec{1}-\vec{q}^k)^{1/2}\diamond\boldsymbol{\sigma}+\boldsymbol{m}\right\rangle\end{aligned}$$

is a centered Gaussian process. We calculate that

$$\begin{aligned}\mathsf{proj}_{U_k(\vec{\Delta})^\perp}\nabla\widetilde{H}_{N,\vec{\Delta},k}(\mathbf{0}) &= \mathsf{proj}_{U_k(\vec{\Delta})^\perp}(\vec{1}-\vec{R}^k)^{1/2}\diamond\left(\nabla H_N(\boldsymbol{m}^k(\vec{\Delta}))-\boldsymbol{g}^{k-1}(\vec{\Delta})\right) \\ &= (\vec{1}-\vec{R}^k)^{1/2}\diamond(\boldsymbol{g}^k(\vec{\Delta})-\boldsymbol{g}^{k-1}(\vec{\Delta}))\,.\end{aligned}$$

Thus conditioning on $(F_k(\vec{\Delta}),P_{U_k(\vec{\Delta})^\perp}\nabla\widetilde{H}_{N,\vec{\Delta},k}(\mathbf{0}))$ is equivalent to conditioning on $F_{k+1}(\vec{\Delta})$. Also, (4.104) shows $U_k(\vec{\Delta})\bowtie\boldsymbol{m}=U_{k+1}(\vec{\Delta})$. For all $\boldsymbol{\sigma}\in\mathcal{S}_N^{U_{k+1}(\vec{\Delta})}$,

$$\phi_{k,\vec{\Delta}}\left((\vec{1}-\vec{q}^k)^{1/2}\diamond\boldsymbol{\sigma}+\boldsymbol{m}\right) = (\vec{1}-\vec{R}^{k+1})^{1/2}\diamond\boldsymbol{\sigma}+(\vec{1}-\vec{R}^k)^{1/2}\diamond\boldsymbol{m}+\boldsymbol{m}^k(\vec{\Delta})\,,$$

and this equals $\phi_{k+1,\vec{\Delta}}(\boldsymbol{\sigma}) = (\vec{1}-\vec{R}^{k+1})^{1/2}\diamond\boldsymbol{\sigma}+\boldsymbol{m}^{k+1}(\vec{\Delta})$ because

$$(\vec{1}-\vec{R}^k)^{1/2}\diamond\boldsymbol{m} = \mathsf{scale}(\boldsymbol{g}^k(\vec{\Delta});\vec{\Delta},\vec{R}^{k+1}-\vec{R}^k) = \boldsymbol{m}^{k+1}(\vec{\Delta})-\boldsymbol{m}^k(\vec{\Delta})\,.$$

Moreover

$$\begin{aligned}&\left\langle P_{U_k(\vec{\Delta})^\perp}\nabla\widetilde{H}_{N,\vec{\Delta},k}(\mathbf{0}),(\vec{1}-\vec{q}^k)^{1/2}\diamond\boldsymbol{\sigma}+\boldsymbol{m}\right\rangle \\ &= \left\langle(\vec{1}-\vec{R}^k)^{1/2}\diamond(\boldsymbol{g}^k(\vec{\Delta})-\boldsymbol{g}^{k-1}(\vec{\Delta})),(\vec{1}-\vec{q}^k)^{1/2}\diamond\boldsymbol{\sigma}+\boldsymbol{m}\right\rangle \\ &= \left\langle\boldsymbol{g}^k(\vec{\Delta})-\boldsymbol{g}^{k-1}(\vec{\Delta}),(\vec{1}-\vec{R}^{k+1})^{1/2}\diamond\boldsymbol{\sigma}+\boldsymbol{m}^{k+1}(\vec{\Delta})-\boldsymbol{m}^k(\vec{\Delta})\right\rangle \\ &= \left\langle\boldsymbol{g}^k(\vec{\Delta})-\boldsymbol{g}^{k-1}(\vec{\Delta}),\phi_{k+1,\vec{\Delta}}(\boldsymbol{\sigma})\right\rangle - \left\langle\boldsymbol{g}^k(\vec{\Delta})-\boldsymbol{g}^{k-1}(\vec{\Delta}),\boldsymbol{m}^k(\vec{\Delta})\right\rangle\,.\end{aligned}$$

Combining the above,

$$\begin{aligned}\widetilde{H}^\circ_{N,\vec{\Delta},k}(\boldsymbol{\sigma}) &= H_N(\phi_{k+1,\vec{\Delta}}(\boldsymbol{\sigma})) - \left\langle\boldsymbol{g}^{k-1}(\vec{\Delta}),\phi_{k+1,\vec{\Delta}}(\boldsymbol{\sigma})\right\rangle + \sum_{i=1}^{k-1}\left\langle\boldsymbol{g}^i(\vec{\Delta})-\boldsymbol{g}^{i-1}(\vec{\Delta}),\boldsymbol{m}^i\right\rangle \\ &\quad - \left\langle\boldsymbol{g}^k(\vec{\Delta})-\boldsymbol{g}^{k-1}(\vec{\Delta}),\phi_{k+1,\vec{\Delta}}(\boldsymbol{\sigma})\right\rangle + \left\langle\boldsymbol{g}^k(\vec{\Delta})-\boldsymbol{g}^{k-1}(\vec{\Delta}),\boldsymbol{m}^k(\vec{\Delta})\right\rangle \\ &= \widetilde{H}_{N,\vec{\Delta},k+1}(\boldsymbol{\sigma})\,.\end{aligned}$$

This proves that conditional on $F_{k+1}(\vec{\Delta})$, $\widetilde{H}_{N,\vec{\Delta},k}$ is a centered Gaussian process on $\mathcal{S}_N^{U_{k+1}(\vec{\Delta})}$. The covariance formula is shown by similarly verifying that

$$\xi_k\left((\vec{1}-\vec{q}^k)\odot\vec{x}+\vec{q}^k\right) - \left\langle\nabla\xi_k(\vec{0}),(\vec{1}-\vec{q}^k)\odot\vec{x}+\vec{q}^k\right\rangle = \xi_{k+1}\left((\vec{1}-\vec{q}^{k+1})\odot\vec{x}+\vec{q}^{k+1}\right)$$

This completes the induction. □

We next verify that if ξ is strictly super-solvable, then so is ξ_k. This fact is key for our recursion.

Proposition 4.6.13. *For each k, the model ξ_k is strictly super-solvable.*

Proof. The definition $\vec{R}^k = \nabla\xi(\vec{R}^{k-1})/\nabla\xi(\vec{1})$ rearranges to

$$\nabla\xi(\vec{1}) - \nabla\xi(\vec{R}^{k-1}) = (\vec{1} - \vec{R}^k) \odot \nabla\xi(\vec{1})\,. \tag{4.105}$$

Thus,

$$\nabla\xi_k(\vec{1}) = (\vec{1} - \vec{R}^k) \odot (\nabla\xi(\vec{1}) - \nabla\xi(\vec{R}^{k-1})) = (\vec{1} - \vec{R}^k)^2 \odot \nabla\xi(\vec{1})\,, \tag{4.106}$$

$$\nabla^2\xi_k(\vec{1}) = (\vec{1} - \vec{R}^k)^{\otimes 2} \odot \nabla^2\xi''(\vec{1})\,. \tag{4.107}$$

Combining the above gives

$$\operatorname{diag}(\nabla\xi_k(\vec{1})) = (\vec{1} - \vec{R}^k)^{\otimes 2} \odot \operatorname{diag}(\nabla\xi(\vec{1})) \succ (\vec{1} - \vec{R}^k)^{\otimes 2} \odot \nabla^2\xi''(\vec{1}) = \nabla^2\xi_k(\vec{1})\,.$$ □

Finally the next corollary explains the choice of radius $\vec{q}^k$. Combined with Lemma 4.6.11, this explains the choice of radii $\vec{R}^k$ in the construction of the bands $\mathsf{Band}_k(\vec{\Delta})$: $\vec{R}^{k+1}$ is chosen so that $\mathsf{Band}_{k+1}(\vec{\Delta})$ is the sub-band of $\mathsf{Band}_k(\vec{\Delta})$ orthogonal to $\boldsymbol{g}^k$ which lies near all approximate critical points of type $\vec{\Delta}$.

Corollary 4.6.14. *With probability $1 - e^{-cN}$, all $(\varepsilon, \vec{\Delta})$-critical points of $\widetilde{H}_{N,\vec{\Delta},k}$ on $U_k(\vec{\Delta})$ lie within $\upsilon\sqrt{N}$ of $\mathsf{Band}^{U_k(\vec{\Delta})}_{\vec{\Delta},\vec{q}^k}(\widetilde{H}_{N,k,\vec{\Delta}})$ for some $\upsilon = o_\varepsilon(1)$.*

Proof. Since ξ_k is strictly super-solvable by Proposition 4.6.13, Corollary 4.6.8 applies. Recalling (4.105), we have

$$\begin{aligned}
\nabla\xi_k(\vec{0}) &= (\vec{1} - \vec{R}^k) \odot (\nabla\xi(\vec{R}^k) - \nabla\xi(\vec{R}^{k-1})) \\
&= (\vec{1} - \vec{R}^k) \odot \Big((\nabla\xi(\vec{1}) - \nabla\xi(\vec{R}^{k-1})) - (\nabla\xi(\vec{1}) - \nabla\xi(\vec{R}^k))\Big) \\
&= (\vec{1} - \vec{R}^k) \odot (\vec{R}^{k+1} - \vec{R}^k) \odot \nabla\xi(\vec{1})\,.
\end{aligned}$$

Recalling (4.106), this implies $\nabla\xi_k(\vec{0})/\nabla\xi_k(\vec{1}) = (\vec{R}^{k+1} - \vec{R}^k)/(\vec{1} - \vec{R}^k) = \vec{q}^k$. The result follows from Corollary 4.6.8. □

4.6.4 Localization of approximate critical points

For $S \subseteq \mathbb{R}^N$ and $\iota > 0$, let $B_\iota(S) \subseteq \mathbb{R}^N$ denote the set of points whose distance to S is at most ι. The following proposition localizes all $(\varepsilon, \vec{\Delta})$-critical points of H_N. Note that by Fact 4.6.2, all ε-critical points of H_N are described by this proposition.

Proposition 4.6.15. *For any constant $k \in \mathbb{N}$, $\varepsilon > 0$, there exists $\iota_k = o_\varepsilon(1)$ (depending on k) such that with probability $1 - e^{-cN}$, all $(\varepsilon, \vec{\Delta})$-critical points of H_N lie in $B_{\iota_k\sqrt{N}}(\mathsf{Band}_k(\vec{\Delta}))$.*

We begin by relating the $(\varepsilon, \vec{\Delta})$-critical points of H_N to those of $\widetilde{H}_{N,\vec{\Delta},k}$.

Lemma 4.6.16. *For any $k \geq 1$, $\varepsilon > 0$, $\vec{\Delta} \in \{-1, 1\}^r$ the following holds. If $\boldsymbol{\sigma} \in \mathsf{Band}_k(\vec{\Delta})$ is an $(\varepsilon, \vec{\Delta})$-critical point of H_N, then $\phi^{-1}_{k,\vec{\Delta}}(\boldsymbol{\sigma}) \in \mathcal{S}_N^{U_k(\vec{\Delta})}$ is a $(\varepsilon, \vec{\Delta})$-critical point of $\widetilde{H}_{N,\vec{\Delta},k}$.*

Proof. Let $\vec{x}(\vec{\Delta}, \xi) \in \mathbb{R}^r$ be the radial derivative defined in (4.12), where we make the dependence on ξ explicit. By definition of $(\varepsilon, \vec{\Delta})$-approximate critical point,

$$\|\nabla H_N(\boldsymbol{\sigma}) - \Lambda^{-1/2}\vec{x}(\vec{\Delta}, \xi) \diamond \boldsymbol{\sigma}\|_2 \leq \upsilon\sqrt{N}\,.$$

We write $\boldsymbol{\sigma} = \phi_{k,\vec{\Delta}}(\boldsymbol{\rho})$ for $\boldsymbol{\rho} \in \mathcal{S}_N^{U_k(\vec{\Delta})}$. Let $\nabla_{U_k(\vec{\Delta})^\perp}$ denote the Euclidean gradient projected into the subspace $U_k(\vec{\Delta})^\perp$. Because this projection is 1-Lipschitz, we also have

$$\|\nabla_{U_k(\vec{\Delta})^\perp} H_N(\boldsymbol{\sigma}) - \Lambda^{-1/2}\vec{x}(\vec{\Delta}, \xi) \diamond \mathsf{proj}_{U_k(\vec{\Delta})^\perp}\boldsymbol{\sigma}\|_2 \leq \upsilon\sqrt{N}\,. \tag{4.108}$$

Taking this gradient of (4.102) yields

$$\nabla_{U_k(\vec{\Delta})^\perp} \widetilde{H}_{N,\vec{\Delta},k}(\boldsymbol{\rho}) = (\vec{1}-\vec{R}^k)^{1/2} \diamond \nabla_{U_k(\vec{\Delta})^\perp} H_N(\phi_{k,\vec{\Delta}}(\boldsymbol{\rho})) = (\vec{1}-\vec{R}^k)^{1/2} \diamond \nabla_{U_k(\vec{\Delta})^\perp} H_N(\boldsymbol{\sigma}) \,,$$

as the gradient contribution from $\boldsymbol{g}^{k-1}(\vec{\Delta})$ projects to zero. Moreover,

$$\mathsf{proj}_{U_k(\vec{\Delta})^\perp} \boldsymbol{\sigma} = \mathsf{proj}_{U_k(\vec{\Delta})^\perp} \phi_{k,\vec{\Delta}}(\boldsymbol{\rho}) = (\vec{1}-\vec{R}^k)^{1/2} \diamond \boldsymbol{\rho} \,.$$

From (4.106) and (4.107), it readily follows that

$$\vec{x}(\vec{\Delta},\xi_k) = (\vec{1}-\vec{R}^k) \odot \vec{x}(\vec{\Delta},\xi) \,.$$

Thus (4.108) implies

$$\begin{aligned}
\upsilon\sqrt{N} &\geq \|(\vec{1}-\vec{R}^k)^{-1/2} \diamond \nabla_{U_k(\vec{\Delta})^\perp} \widetilde{H}_{N,\vec{\Delta},k}(\boldsymbol{\rho}) - (\Lambda^{-1/2}\vec{x}(\vec{\Delta},\xi) \odot (\vec{1}-\vec{R}^k)^{1/2}) \diamond \boldsymbol{\rho}\|_2 \\
&= \|(\vec{1}-\vec{R}^k)^{-1/2} \diamond (\nabla_{U_k(\vec{\Delta})^\perp} \widetilde{H}_{N,\vec{\Delta},k}(\boldsymbol{\rho}) - \Lambda^{-1/2}\vec{x}(\vec{\Delta},\xi_k) \diamond \boldsymbol{\rho})\|_2 \\
&\geq \|\nabla_{U_k(\vec{\Delta})^\perp} \widetilde{H}_{N,\vec{\Delta},k}(\boldsymbol{\rho}) - \Lambda^{-1/2}\vec{x}(\vec{\Delta},\xi_k) \diamond \boldsymbol{\rho}\|_2 \,.
\end{aligned}$$

So, $\boldsymbol{\rho} = \phi_{k,\vec{\Delta}}^{-1}(\boldsymbol{\sigma})$ is an $(\varepsilon,\vec{\Delta})$-critical point of $\widetilde{H}_{N,\vec{\Delta},k}$. □

Proof of Proposition 4.6.15. Throughout we assume $H_N \in K_N$, which holds with probability $1-e^{-cN}$ by Proposition 4.2.4. We induct on k. Suppose the claim holds for k, so all $(\varepsilon,\vec{\Delta})$-critical points of H_N lie in $B_{\iota_k\sqrt{N}}(\mathsf{Band}_k(\vec{\Delta}))$. Let $\boldsymbol{\sigma}$ be one such critical point, and let $\boldsymbol{\rho} \in \mathsf{Band}_k(\vec{\Delta})$ be its projection in $\mathsf{Band}_k(\vec{\Delta})$. Because $H_N \in K_N$, $\boldsymbol{\rho}$ is a $(\varepsilon',\vec{\Delta})$-critical point of H_N for some $\varepsilon' = o_k(1)$.

By Lemma 4.6.16, $\boldsymbol{\tau} = \phi_{k,\vec{\Delta}}^{-1}(\boldsymbol{\rho}) \in \mathcal{S}_N^{U_k(\vec{\Delta})}$ is a $(\varepsilon',\vec{\Delta})$-critical point of $\widetilde{H}_{N,k,\vec{\Delta}}$. By Corollary 4.6.14, (with probability $1-e^{-cN}$)

$$\boldsymbol{\tau} \in B_{\upsilon\sqrt{N}}(\mathsf{Band}^{U_k(\vec{\Delta})}_{\vec{\Delta},\vec{q}^k}(\widetilde{H}_{N,k,\vec{\Delta}}))$$

for some $\upsilon = o_{\varepsilon'}(1) = o_\varepsilon(1)$. By Lemma 4.6.11,

$$\phi_{k,\vec{\Delta}}(\mathsf{Band}^{U_k(\vec{\Delta})}_{\vec{\Delta},\vec{q}^k}(\widetilde{H}_{N,k,\vec{\Delta}})) = \mathsf{Band}_{k+1}(\vec{\Delta}) \,,$$

so (letting C_k be the Lipschitz constant of $\phi_{k,\vec{\Delta}}$)

$$\boldsymbol{\rho} = \phi_{k,\vec{\Delta}}(\boldsymbol{\tau}) \in B_{C_k\upsilon\sqrt{N}}(\mathsf{Band}_{k+1}(\vec{\Delta})) \,.$$

It follows that $\boldsymbol{\sigma} \in B_{\iota_{k+1}\sqrt{N}}$ for $\iota_{k+1} = C_k\upsilon + \iota_k$. Over this argument we union bounded over $k+1 = O(1)$ events with probability $1-e^{-cN}$, so the conclusion holds with probability $1-e^{-cN}$. □

4.6.5 Existence and uniqueness of exact critical points

So far we have established that all $(\varepsilon,\vec{\Delta})$-critical points of H_N are close together. This easily implies that each $\vec{\Delta}$ has at most 1 associated (exact) critical point.

Definition 4.6.17. Let $\varepsilon > 0$ be a sufficiently small constant independent of N. A $\vec{\Delta}$-critical point $\boldsymbol{x}_{\vec{\Delta}}$ of H_N is a critical point that is also a $(\varepsilon,\vec{\Delta})$-critical point (i.e. whose radial derivative $\nabla_{\mathsf{rad}} H_N(\boldsymbol{x}_{\vec{\Delta}})$ satisfies (4.97)).

Proposition 4.6.18. *With probability $1-e^{-cN}$, for each $\vec{\Delta} \in \{-1,1\}^r$ there is at most one $\vec{\Delta}$-critical point of H_N.*

Proof. Let $\boldsymbol{x}_{\vec{\Delta}}, \boldsymbol{x}'_{\vec{\Delta}}$ be two such critical points of H_N. Then they are both $(\varepsilon, \vec{\Delta})$ critical points for small $\varepsilon > 0$. By Fact 4.6.2 and Proposition 4.6.15, we find that for any $\delta > 0$ independent of N,

$$\|\boldsymbol{x}_{\vec{\Delta}} - \boldsymbol{x}'_{\vec{\Delta}}\|_2 \le \delta\sqrt{N}$$

holds with probability $1 - e^{-cN}$. Moreover (4.50) and Lemma 4.4.3 together imply that with the same probability, the spherical Hessians of H_N at both points are $C(\xi)$ well-conditioned, with all eigenvalues inside $\pm[C^{-1}, C]$. For δ small enough and $H_N \in K_N$, this is impossible since C-well-conditioned critical points cannot be arbitrarily close together (as can be shown by Taylor expanding along a geodesic as in Lemma 4.5.12). □

To show existence we appeal to Morse theory, which shows the total number of critical points is almost surely at least 2^r just from the geometry of $\mathcal{S}_N$.

Proposition 4.6.19. *Almost surely, H_N has at least 2^r critical points on $\mathcal{S}_N$. Hence by Fact 4.6.2 and Proposition 4.6.18, with probability $1 - e^{-cN}$, H_N has a unique $\vec{\Delta}$-critical point of each type $\vec{\Delta} \in \{-1,1\}^r$, and no other critical points.*

Proof. It suffices to show the first claim. Recall that any sphere has two non-zero homology groups, each of dimension 1. Hence by the Künneth formula, the sum of the dimensions of the homology groups for $\mathcal{S}_N$ is 2^r. Finally by the Morse inequalities (see e.g. [Mil63]), this sum lower bounds the number of critical points of any Morse function, in particular H_N. □

Putting everything together, we obtain most of Theorem 4.1.13.

Proof of Theorem 4.1.13 except for part (b). Existence and uniqueness of each $\boldsymbol{x}_{\vec{\Delta}}$ have just been shown. As above, (4.50) and Lemma 4.4.3 imply the well-conditioning. Proposition 4.6.15 implies part (a). Part (c) is immediate from part (a) since all approximate ground states of $H_N \in K_N$ are approximate critical points. □

Remark 4.6.20. As an alternative to the Morse inequalities, we could instead use [HS24, Proposition 3.2] which, for each $\vec{\Delta} \in \{-1,1\}^r$ and $\varepsilon > 0$, explicitly constructs a $(\vec{\Delta}, \varepsilon)$-approximate critical point $\widetilde{\boldsymbol{x}}_{\vec{\Delta}} \in \mathcal{S}_N$ (with probability $1 - e^{-cN}$). Since the limiting spectral support $S(\vec{\Delta})$ is bounded away from 0 by Proposition 4.4.3, Proposition 4.5.1 implies that for ε small enough, each $\widetilde{\boldsymbol{x}}_{\vec{\Delta}}$ has well-conditioned Hessian. Then Newton's method can be used to locate a nearby exact critical point $\boldsymbol{x}_{\vec{\Delta}}$. This route is more cumbersome than the one taken above, but has a chance to work in situations where the number of critical points in the trivial regime is larger than the lower bound from the Morse inequalities.

4.6.6 The index of each critical point

Finally we compute the index of each critical point, which is the only remaining part of Theorem 4.1.13. We use a "critical point following" argument, showing that critical points move stably as H_N is gradually deformed into a linear function, while their indices remain fixed. This can easily be turned into an efficient algorithm to locate each $\boldsymbol{x}_{\vec{\Delta}}$ as mentioned in the introduction, as the proof of [MS23, Lemma 3.1] used below is via projected gradient descent on $\|\nabla H_N(\cdot)\|_2^2$ (i.e. Newton's method).

Proposition 4.6.21. *For any $(\vec{\lambda}, \xi)$ and $\iota > 0$ there is $\varepsilon > 0$ such that the following holds. Suppose $H_N \in K_N$, and $\nabla^2_{\mathsf{sp}} H_N(\boldsymbol{x})$ is an ι-well-conditioned ε-approximate critical point. Then there exists an exact critical point $\boldsymbol{y} \in \mathcal{S}_N$ such that $\|\boldsymbol{x} - \boldsymbol{y}\|_2 \le C(\vec{\lambda}, \xi, \iota)\varepsilon\sqrt{N}$.*

Proof. This follows by [MS23, Lemma 3.1] applied to ∇H_N; the constants J_n, L_n, M_n are bounded as $H_N \in K_N$. (The stated result is for a single sphere, but the extension to a finite products of spheres poses no issues.) □

Proof of Theorem 4.1.13(b). Fix $(\vec{\lambda}, \xi)$. Writing $\widetilde{H}_N$ for the degree two and higher terms in H_N, for $t \in [0,1]$ we set

$$H_{N,t} = \langle \boldsymbol{G}^{(1)}, \boldsymbol{x} \rangle + t\widetilde{H}_N(\boldsymbol{x}).$$

The marginal distribution of $H_{N,t}$ thus corresponds to the mixture $\xi^{(t)}(\vec{x}) = (1-t^2)\xi'(0) \odot \vec{x} + t^2\xi(\vec{x})$. It is easy to see that if ξ is strictly super-solvable then so is $\xi^{(t)}$ for each $t \in [0,1]$. Moreover the proof of Lemma 4.4.3 holds uniformly on $(\xi^{(t)})_{t\in[0,1]}$, implying that for some $c > 0$,

$$S_t(\vec{\Delta}) \cap [-c,c] = \emptyset$$

holds simultaneously for all $t \in [0,1]$ and $\vec{\Delta} \in \{-1,1\}^r$. In particular, our results then imply the following. Fix a small unit fraction $\delta > 0$ depending on $(\vec{\lambda}, \xi, c)$, and let $\boldsymbol{x}_{\vec{\Delta},k\delta}$ be the corresponding critical point for $H_{N,k\delta}$ (which exists with probability $1 - e^{-cN}$). Then for $k \geq 0$, if $H_N, \widetilde{H}_N \in K_N$:

$$\begin{aligned} \|\nabla_{\mathsf{sp}} H_{N,(k+1)\delta}(\boldsymbol{x}_{\vec{\Delta},k\delta})\|_2 &\leq C\delta\sqrt{N}, \\ \mathsf{spec}\left(\nabla^2_{\mathsf{sp}} H_{N,(k+1)\delta}(\boldsymbol{x}_{\vec{\Delta},k\delta})\right) \cap [-c/2, c/2] &= \emptyset. \end{aligned} \tag{4.109}$$

For $H_N, \widetilde{H}_N \in K_N$, the above two estimates imply via Proposition 4.6.21 the existence of a nearby critical point $\boldsymbol{y}_{\vec{\Delta},(k+1)\delta}$ for $H_{N,(k+1)\delta}$ such that, for a constant $C_1 = C_1(\vec{\lambda}, \xi, c)$:

$$|\langle \boldsymbol{y}_{\vec{\Delta},(k+1)\delta} - \boldsymbol{x}_{\vec{\Delta},k\delta}, \boldsymbol{G}^{(1)} \rangle| \leq C_1\delta, \tag{4.110}$$

$$\|\nabla^2_{\mathsf{sp}} H_{N,(k+1)\delta}(\boldsymbol{y}_{\vec{\Delta},(k+1)\delta}) - \nabla^2_{\mathsf{sp}} H_{N,k\delta}(\boldsymbol{x}_{\vec{\Delta},k\delta})\|_{\mathsf{op}} \leq C_1\delta. \tag{4.111}$$

Recalling (4.12) and (4.14), it follows from (4.110) that $\boldsymbol{y}_{\vec{\Delta},(k+1)\delta} = \boldsymbol{x}_{\vec{\Delta},(k+1)\delta}$ is a critical point of the same $\vec{\Delta}$. Combining (4.109) and (4.111), we see that $\nabla^2_{\mathsf{sp}} H_{N,(k+1)\delta}(\boldsymbol{x}_{\vec{\Delta},(k+1)\delta})$ and $\nabla^2_{\mathsf{sp}} H_{N,k\delta}(\boldsymbol{x}_{\vec{\Delta},k\delta})$ have the same number of positive eigenvalues for each δ. Taking $k = 0$, it is easy to see that this number is $\sum_{s:\vec{\Delta}_s=-1} N_s$. Taking $k\delta = 1$ shows that the same holds for $\nabla^2_{\mathsf{sp}} H_N(\boldsymbol{x}_{\vec{\Delta}})$ as desired. $\square$

4.7 Estimates for approximate critical points in single-species models

In this section we detail further consequences of Lemma 4.5.13 which are of independent interest, and of relevance for several concurrent works. Below, we restrict our attention to single-species models without external field (i.e. $r = 1$, $\xi'(0) = 0$) for which the relevant Kac–Rice estimates are known from previous work. In particular we consider mixture functions of the form $\xi(t) = \sum_{p=2}^{P} \gamma_p^2 t^p$ for $\gamma_2, \ldots, \gamma_P \geq 0$. We assume for sake of normalization that $\xi(1) = 1$ and similarly to Definition 4.1.6, we write $\xi' = \xi'(1), \xi'' = \xi''(1)$ and $\alpha^2 = \xi'' + \xi' - (\xi')^2$ (unrelated to (4.81)). Recall from [AB13] the thresholds:

$$E_\infty^{\pm}(\xi) \equiv \frac{2\xi'\sqrt{\xi''} \pm \sqrt{4\xi''(\xi')^2 - (\xi'' + \xi')\left(2\left(\xi'' - \xi' + (\xi')^2\right) - \alpha^2 \log \frac{\xi''}{\xi'}\right)}}{\xi' + \xi''}.$$

One always has $\alpha \geq 0$, with equality exactly in the pure case $\xi(t) = t^p$ for some p. In this case, the thresholds $E_\infty^{\pm}$ agree at the value $E_\infty(p) = 2\sqrt{\frac{p-1}{p}}$ from [ABČ13].

We give the relevant Kac–Rice result in Proposition 4.7.3 below after recalling some definitions and results from [AB13]. For open $\mathcal{D}, \mathcal{D}_{\mathsf{rad}} \subseteq \mathbb{R}$ we let $\mathsf{Crt}_N(\mathcal{D}; \mathcal{D}_{\mathsf{rad}}) \subseteq \mathcal{S}_N$ consist of all critical points with

$$H_N(\boldsymbol{x})/N \in \mathcal{D}, \qquad \nabla_{\mathsf{rad}} H_N(\boldsymbol{x}) \in \mathcal{D}_{\mathsf{rad}}.$$

As in Equation (1.21) therein, for $\gamma \in (0,1)$ define $s_\gamma \in (-\sqrt{2}, \sqrt{2})$ as the rescaled semicircular law quantile satisfying:

$$\gamma = \frac{1}{\pi} \int_{-\sqrt{2}}^{-s_\gamma} \sqrt{2 - x^2}\, \mathsf{d}x.$$

Moreover, define the function

$$\Theta(s) = \left(-\frac{|s|\sqrt{s^2-2}}{2} + \log\left(\frac{|s| + \sqrt{s^2-2}}{\sqrt{2}}\right)\right) 1_{|s| \geq \sqrt{2}} \leq 0. \tag{4.112}$$

The critical point complexity functional at $\big(H_N(\boldsymbol{x})/N, \nabla_{\mathsf{rad}} H_N(\boldsymbol{x})\big) \approx \big(y, s\sqrt{2\xi''}\big)$ and its quadratic upper bound are given by:

$$\begin{aligned} F(s,y) &= \frac{1}{2}\left(\log\frac{\xi''}{\xi'} + s^2 - y^2 - \frac{2\xi''}{\alpha^2}\left(s - \frac{y\xi'}{\sqrt{2\xi''}}\right)^2 + \Theta(s)\right), \\ \widetilde{F}(s,y) &= \frac{1}{2}\left(\log\frac{\xi''}{\xi'} + s^2 - y^2 - \frac{2\xi''}{\alpha^2}\left(s - \frac{y\xi'}{\sqrt{2\xi''}}\right)^2\right). \end{aligned} \tag{4.113}$$

(If $\alpha = 0$, we interpret $-0/0 = 0$ and $-x/0 = -\infty$ for $x > 0$.)

Indeed the following holds as a direct consequence of Proposition 4.3.2, see also the proof of [AB13, Theorem 1.3]. (In fact our scaling of $\nabla_{\mathsf{rad}} H_N(\boldsymbol{x})$ by $\sqrt{2\xi''}$ is chosen to enforce agreement with the latter formula).

Proposition 4.7.1. *For any $\gamma \in (0,1)$ and open $\mathcal{D}, \mathcal{D}_{\mathsf{rad}} \subseteq \mathbb{R}$:*

$$\lim_{N\to\infty} \frac{1}{N} \log \mathbb{E}\big|\mathsf{Crt}_N\big(\mathcal{D}\ ; \mathcal{D}_{\mathsf{rad}}\big)\big| = \sup_{\substack{y\in\mathcal{D}, \\ s\sqrt{2\xi''}\in\mathcal{D}_{\mathsf{rad}}}} F(s,y).$$

Define the set

$$S = S_\xi = \{(s,y) \in \mathbb{R}^2 \ :\ \widetilde{F}(s,y) \ge 0\}.$$

We now make an important observation on the function $\widetilde{F}$.

Proposition 4.7.2. *The function $\widetilde{F}$ is negative definite, i.e. S is a centered ellipsoid (which degenerates to a line-segment if $\alpha = 0$). Moreover the major axis of S has "positive" slope in $[0, \pi/2]$.*

Proof. We begin with negative definiteness, assuming $\alpha > 0$ as the pure case is easy. Note the first three terms of $\widetilde{F}(s,y)$ are a quadratic of type $(1,1)$, while the last term subtracts a positive-semidefinite quadratic. Hence $\widetilde{F}$ cannot be positive definite, so it suffices to prove it has positive discriminant. After some easy computation, the discriminant's positivity reduces to proving that

$$(2\xi'' - \alpha^2)((\xi')^2 + \alpha^2) \overset{?}{>} 2\xi''(\xi')^2.$$

Dividing by α^2, this reduces to showing $\alpha^2 > 2\xi'' - (\xi')^2$. This in turn rearranges to $\xi'' > \xi'$ which is clear.

The latter assertion holds as if $(s,y) \in S$ with $sy \le 0$ then also $(s,-y), (-s,y) \in S$. $\square$

In the next proposition, we use the notation $GS(\xi) = \text{p-lim}_{N\to\infty} \max_{\boldsymbol{x}\in\mathcal{S}_N} H_N(\boldsymbol{x})/N$ for the ground state energy.

Proposition 4.7.3. *For any $\upsilon > 0$, and for ι small enough depending on (ξ, υ):*

$$\lim_{N\to\infty} \frac{1}{N} \log \mathbb{E}\big|\mathsf{Crt}_N\big((-\infty, E_\infty^- - \upsilon)\ ;\ (\sqrt{2\xi''(1)} - \iota, \infty)\big)\big| < -c(\xi,\upsilon) < 0. \tag{4.114}$$

Furthermore, either $GS(\xi) \le E_\infty^+$ or

$$\lim_{N\to\infty} \frac{1}{N} \log \mathbb{E}\big|\mathsf{Crt}_N\big((E_\infty^+ + \upsilon, \infty)\ ;\ (-\infty, \sqrt{2\xi''(1)} + \iota)\big)\big| < -c(\xi,\upsilon) < 0. \tag{4.115}$$

Proof. We assume $\alpha > 0$ as otherwise the statement is easy. Note that $\sqrt{2\xi''(1)}$ in the statement corresponds to $s = \sqrt{2}$. By inspection, the line $s = \sqrt{2}$ intersects the boundary of S at $(E_\infty^-, \sqrt{2}), (E_\infty^+, \sqrt{2})$. The remainder of the proof is an elementary two-dimensional geometry argument depicted in Figure 4.7.1.[3] First, because the major axis of S has positive slope in $[0, \pi/2]$, the point in S with minimal y coordinate must have negative s coordinate, while the point in S with maximal s coordinate must have positive y

[3] In many cases $E_\infty^- \ge 0$, but the picture is drawn to emphasize that we do not require it. The red region is non-empty only when $GS(\xi) \le E_\infty^+$.

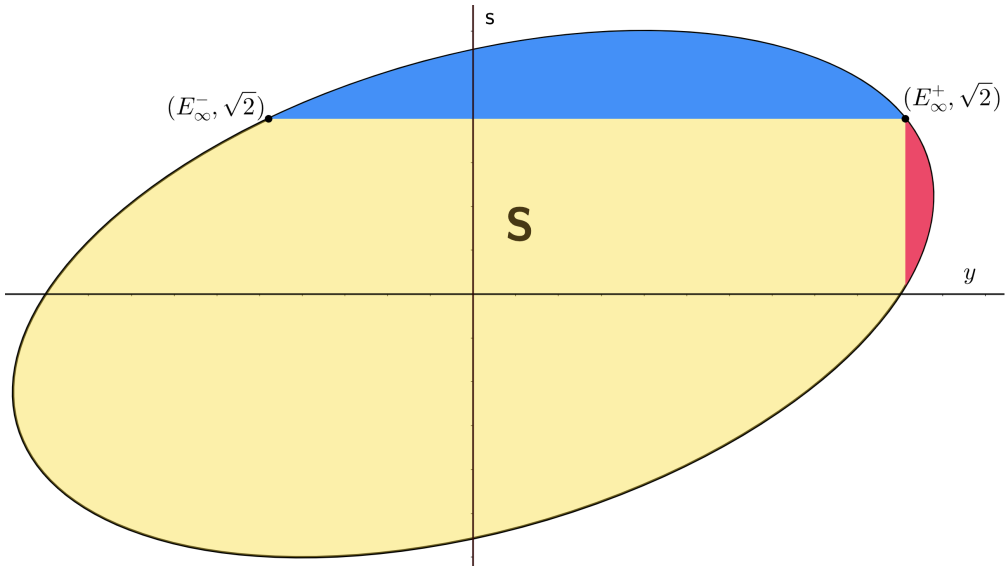


Figure 4.7.1: A diagram of the ellipsoid S, used in the proof of Proposition 4.7.3. If the tangent line to S at $(E_\infty^+,\sqrt{2})$ has positive slope, then the red region is empty and we conclude (4.115). If not, all local maxima correspond to the blue region, hence have energy at most $E_\infty^+ + o_N(1)$. This implies that $GS(\xi) \le E_\infty^+$.

coordinate. For all points on the boundary of S between these two points, in particular $(E_\infty^-,\sqrt{2})$, the tangent line to S has positive slope. Therefore

$$S \cap \Big((-\infty, E_\infty^-) \times (\sqrt{2},\infty)\Big) = \emptyset,$$

which easily implies the first claim.

For the second claim, suppose in the first case that the the tangent line to $(E_\infty^+,\sqrt{2})$ has slope in $[0,\pi/2]$. Then the result follows similarly to the first part of the proof. However as shown in the diagram, it may be that this tangent slope is strictly negative, in $(\pi/2,\pi)$. In this case, we observe (see e.g. Proposition 4.7.4 below) that with probability at least $1-e^{-cN}$, all local maxima of H_N have $s \ge \sqrt{2} - o_N(1)$. And if the tangent slope at $(E_\infty^+,\sqrt{2})$ is negative,

$$S \cap \Big(\mathbb{R} \times (\sqrt{2},\infty)\Big) \subseteq (-\infty, E_\infty^+) \times \mathbb{R}.$$

Since the global maximum of H_N is a local maximum, we conclude $GS(\xi) \le E_\infty^+$, completing the proof. $\square$

The following fact was used above.

Proposition 4.7.4 ([Sub21a, Lemma 3]). *For any $\varepsilon > 0$, there exists c,δ such that*

$$\mathbb{P}\left[\sup_{\boldsymbol{x}\in\mathcal{S}_N} \left|\boldsymbol{\lambda}_{\lfloor \delta N\rfloor}(\nabla^2_{\mathsf{sp}} H_N(\boldsymbol{x})) - \sqrt{2\xi''(1)}\nabla_{\mathsf{rad}} H_N(\boldsymbol{x})\right| \le \varepsilon\right] \ge 1 - e^{-cN}.$$

Proposition 4.7.4 also justifies the following definitions. The last is motivated in part by [FSU21], which suggests that optimization algorithms for general mean-field disordered systems ought to get stuck in ε-marginal local maxima.

Definition 4.7.5. We say the ε-critical point $\boldsymbol{x} \in \mathcal{S}_N$ is:

- An **ε-approximate local maximum** if $\nabla_{\mathsf{rad}} H_N(\boldsymbol{x}) \ge \sqrt{2\xi''(1)} - \varepsilon$.
- An **ε-approximate local non-maximum** if $\nabla_{\mathsf{rad}} H_N(\boldsymbol{x}) \le \sqrt{2\xi''(1)} + \varepsilon$.
- An **ε-marginal local maximum** if both preceding estimates hold: $|\nabla_{\mathsf{rad}} H_N(\boldsymbol{x}) - \sqrt{2\xi''(1)}| \le \varepsilon$.

We now use Lemma 4.5.13 to control the energy levels at which such ε-critical points can exist. (One could also directly apply Theorem 4.5.2, but this leads to some notational burden.)

Corollary 4.7.6. *Fix any $\upsilon > 0$. For sufficiently small ε, with probability $1 - e^{-cN}$ all ε-marginal local maxima satisfy*

$$E_\infty^- - \upsilon \le H_N(\boldsymbol{x})/N \le E_\infty^+ + \upsilon. \tag{4.116}$$

In fact the lower bound holds for all ε-approximate local maxima, while the upper bound holds for all ε-approximate local non-maxima.

Proof. We proceed in two similar cases, first showing the left-hand side of (4.116) for ε-approximate local maxima.

Case 1: lower bound Let $K_N^{\max}(\varepsilon,\upsilon)$ consist of those $H_N \in K_N(\varepsilon)$ for which there exists an ε-approximate local maximum $\boldsymbol{x} \in \mathcal{S}_N$ with

$$H_N(\boldsymbol{x})/N \le E_\infty^- - \upsilon.$$

We apply Lemma 4.5.13 as in the proof of Proposition 4.5.1. We find that for some $\iota = o_\varepsilon(1)$,

$$\mathbb{E}\big|\,\mathsf{Crt}_N\big((-\infty, E_\infty^- - \upsilon/2)\ ;\ (\sqrt{2\xi''(1)} - \iota, \infty)\big)\,\big| \ge e^{-o_\varepsilon(N)} \cdot \mathbb{P}[H_N \in K_N^{\max}(\varepsilon,\upsilon)]. \tag{4.117}$$

Again using Proposition 4.5.10, it suffices to show the left-hand side above is at most $e^{-c(\xi,\upsilon)N}$ for ε sufficiently small, which is the statement of (4.114). This proves the left-hand inequality of (4.116) in the claimed sense.

Case 2: upper bound Let $K_N^{\mathsf{nonmax}}(\varepsilon,\upsilon)$ consist of those $H_N \in K_N(\varepsilon)$ for which some ε-local non-maximum $\boldsymbol{x} \in \mathcal{S}_N$ satisfies

$$H_N(\boldsymbol{x})/N \ge E_\infty^+ + \upsilon.$$

Again using Lemma 4.5.13, we find

$$\mathbb{E}\big|\mathsf{Crt}_N\big((E_\infty^+ + \upsilon/2, \infty)\ ;\ (-\infty, \sqrt{2\xi''(1)} + \iota)\big)\big| \ge e^{-o_\varepsilon(N)} \cdot \mathbb{P}[H_N \in K_N^{\mathsf{nonmax}}(\varepsilon,\upsilon)]. \tag{4.118}$$

In the case that (4.115) holds, the proof is as in the first case. If $GS(\xi) \le E_\infty^+$, then the upper bound holds trivially. □

We briefly summarize the applications of Corollary 4.7.6 in our concurrent works. See the individual papers for more detail. As partially mentioned in Remark 4.5.18, our work [HS24] uses approximate message passing to construct ε-marginal local maxima at the algorithmic threshold energy ALG for any strictly sub-solvable ξ. For $r = 1$, Corollary 4.7.6 thus implies that $\mathsf{ALG} \in [E_\infty^-, E_\infty^+]$, generalizing the fact that $\mathsf{ALG} = E_\infty$ for pure models. Separately, [Sel24b] proves that spherical Langevin dynamics at large inverse temperature β rapidly climbs to and stays above the energy of the lowest lying ε-approximate local maximum, up to error $o_{\beta\to\infty}(1)$. Corollary 4.7.6 thus gives $E_\infty^- - o_\beta(1)$ as an explicit energy lower bound.

Finally we present a consequence for approximate critical points of finite index. Recall from [AB13] the positive thresholds $(E_k)_{k\ge 0}$, defined so that E_k is the larger of two zeros for the index k critical point complexity function $\theta_{k,\xi}$ (where we have implicitly negated H_N to make all energies positive). [AB13] deduced from Markov's inequality that H_N has no index k critical points at energies strictly above E_k, and we extend this to approximate critical points. Note that we consider positive energy values, so our signs are switched.

Corollary 4.7.7. *Let $E > E_k$ for fixed k and let $\varepsilon(E,k)$ be sufficiently small. Then with probability $1-e^{-cN}$, all ε-approximate critical points $\boldsymbol{x}$ with $H_N(\boldsymbol{x})/N \ge E$ have index at most k. Furthermore if $k = 1$, then all such $\boldsymbol{x}$ are within distance $\eta\sqrt{N}$ from a local maximum where $\eta = \eta(E,k,\varepsilon) \to 0$ as $\varepsilon \to 0$ for each fixed E, k.*

Proof. It follows from the smoothness and monotonicity properties for $\theta_{k,\xi}$ in [AB13, Proposition 1] that for some $\delta(E,k) > 0$, the expected number of critical points $\boldsymbol{x} \in \mathcal{S}_N$ with $\boldsymbol{\lambda}_k(\nabla_{\mathsf{sp}}^2 H_N(\boldsymbol{x})) \ge -2\delta$ is at most e^{-cN}. Another similar application of Lemma 4.5.13 for sufficiently small ε implies with probability $1 - e^{c'N}$, all ε-approximate critical points $\boldsymbol{x} \in \mathcal{S}_N$ satisfy $\boldsymbol{\lambda}_k(\nabla_{\mathsf{sp}}^2 H_N(\boldsymbol{x})) \le -\delta$. This completes the proof. For the second claim, note that when $k = 1$, we have just showed $\mathsf{supp}\,\big(\nabla_{\mathsf{sp}}^2 H_N(\boldsymbol{x})\big) \cap [-\delta,\delta] = \emptyset$. Hence for ε small compared to δ, Proposition 4.6.21 yields the result. □

Appendix

4.A Properties of solutions to the vector Dyson equation

In this appendix, we establish properties of the vector Dyson equation (4.59) that we use in the paper. It will be useful to rename $\vec{v} + z\vec{\lambda}$ to $\vec{v}$ and allow $\vec{v}$ to vary in all of $\overline{\mathbb{H}}^r$. Thus we study the equation

$$v_s = -\frac{\lambda_s}{u_s} - \sum_{s' \in \mathscr{S}} \xi''_{s,s'} u_s. \tag{4.119}$$

The right-hand side of (4.119) is a function of $\vec{u}$, which we will denote $\vec{v}(\vec{u})$. There will be no confusion with the notations $\vec{v}(\vec{\Delta})$, $\vec{u}(\vec{\Delta})$ (defined in Corollary 4.4.11 and equation (4.68)), which do not appear in this appendix.

Lemma 4.A.1 ([HFS07, Section 3])**.** *For any $\vec{v} \in \mathbb{H}^r$, there exists a unique solution $\vec{u} = \vec{u}(\vec{v}) \in \mathbb{H}^r$ to* (4.119)*.*

As a result, for $z \in \mathbb{H}$ the solution $\vec{u}(z; \vec{v})$ to the Dyson equation (4.38) is given by

$$\vec{u}(z; \vec{v}) = \vec{u}(\vec{v} + z\vec{\lambda}). \tag{4.120}$$

The first result of this Appendix establishes continuity of $\vec{u}(\cdot)$ in $\vec{v}$. This extends the $1/3$-Hölder continuity in z proved in [AEK17a], which corresponds for us to varying $\vec{v}$ along certain 1-dimensional subspaces. See also [AEK20, Section 10] for continuity properties in ξ. We note that, importantly, these works treat extremely general models with a continuum of "species" parametrized by a probability measure. By contrast we will not hesitate to use the assumption that r is finite (e.g. in (4.126)).

Theorem 4.A.2. *The solution $\vec{u}(\vec{v})$ to* (4.119) *identified by Lemma 4.A.1 extends to a $1/3$-Hölder continuous function $\vec{u} : \overline{\mathbb{H}}^r \to \overline{\mathbb{H}}^r$.*

Thus the identification (4.120) remains true as z tends to the real line. I.e. as $z \to \gamma \in \mathbb{R}$, the limit $\vec{u}(\gamma; \vec{v})$ of $\vec{u}(z; \vec{v})$ (well-defined by Proposition 4.2.10) equals $\vec{u}(\vec{v} + \gamma\vec{\lambda})$.

The proof of Theorem 4.A.2 consists of two steps. We first show $\Im\vec{u}$ is $1/3$-Hölder, and then we extend this to $\Re\vec{u}$. The first step is handled similarly to [AEK17a], though care must be used to handle $\vec{v}$ with imaginary parts of very different sizes. In the second step, we start by deducing via Stieltjes transforms that $\vec{u}$ can be extended in a Hölder continuous way within certain 1-dimensional subspaces. To glue these extensions together, we employ results from harmonic analysis on the boundary behavior of harmonic functions. In particular the consistency of these extensions on different lines intersecting at a common point $\vec{v} \in \mathbb{R}^r$ follows from the existence of non-tangential limits.

Remark 4.A.3. By continuity of $\vec{v}(\cdot)$, for any $\vec{v} \in \overline{\mathbb{H}}^r$ the point $\vec{u} = \vec{u}(\vec{v})$ defined by Theorem 4.A.2 is a solution to the Dyson equation $\vec{v} = \vec{v}(\vec{u})$. However, for $\vec{v} \in \overline{\mathbb{H}}^r \setminus \mathbb{H}^r$, this solution is not necessarily the unique preimage of $\vec{v}$ in $\overline{\mathbb{H}}^r$.

For $\vec{u} \in \overline{\mathbb{H}}^r$, recall from (4.65) the definitions:

$$M(\vec{u}) = \operatorname{diag}\left(\frac{\lambda_s}{u_s^2}\right) - \xi'', \qquad \overline{M}(\vec{u}) = \operatorname{diag}\left(\frac{\lambda_s}{|u_s|^2}\right) - \xi''.$$

Lemma 4.A.4. *For any $\vec{v} \in \overline{\mathbb{H}}^r$ such that $M(\vec{u}(\vec{v}))$ is invertible, $\vec{u}(\cdot)$ is differentiable at $\vec{v}$ and $\nabla \vec{u}(\vec{v}) = M(\vec{u}(\vec{v}))^{-1}$.*

Our next result determines the images $\vec{u}(\overline{\mathbb{H}}^r)$ and $\vec{u}(\mathbb{R}^r)$. In particular, this characterizes which $(\vec{v}, \vec{u})$ pairs solving (4.119) for $\vec{v} \in \overline{\mathbb{H}}^r$ are genuine solutions obtainable as limits of solutions with $\vec{v} \in \mathbb{H}^r$.

Theorem 4.A.5. *Let $\vec{u}^* \in \overline{\mathbb{H}}^r$.*

(a) There exists $\vec{v} \in \overline{\mathbb{H}}^r$ such that $\vec{u}^* = \vec{u}(\vec{v})$ if and only if $\overline{M}(\vec{u}^*) \succeq 0$ and $\overline{M}(\vec{u}^*)\Im(\vec{u}^*) \succeq \vec{0}$.

(b) There exists $\vec{v} \in \mathbb{R}^r$ such that $\vec{u}^* = \vec{u}(\vec{v})$ if and only if one of the following conditions holds.

(i) $\vec{u}^* \in \mathbb{R}^r$ and $M(\vec{u}^*) \succeq 0$.

(ii) $\vec{u}^* \in \mathbb{H}^r$, $\overline{M}(\vec{u}^*) \succeq 0$, and $\overline{M}(\vec{u}^*)\Im(\vec{u}^*) = 0$.

Corollary 4.A.6. *If $\vec{u}^* \in \mathbb{H}^r$ and there exists $\vec{v} \in \overline{\mathbb{H}}^r$ such that $\vec{u}^* = \vec{u}(\vec{v})$, then $M(\vec{u}^*)$ is invertible.*

Next in Proposition 4.A.7, we show that singularity of $M(\vec{u}(\vec{v}))$ (which by the previous corollary requires $\vec{v} \in \mathbb{R}^r$) corresponds to 0 being an edge/cusp of associated spectral measures, and give a precise description of each case. For any $\vec{\chi} \in \mathbb{R}^r_{>0}$ with $\|\vec{\chi}\|_1 = 1$, the restriction of $\vec{u}(\cdot)$ onto the line $\vec{v} + z\vec{\chi}$, $z \in \overline{\mathbb{H}}$ is a rescaled Stieltjes transform of a suitable random matrix. Indeed, this restriction solves

$$v_s + \chi_s z = -\frac{\lambda_s}{u_s(\vec{v} + z\vec{\chi})} - \sum_{s' \in \mathscr{S}} \xi''_{s,s'} u_{s'}(\vec{v} + z\vec{\chi}).$$

We set

$$\tilde{\xi}''_{s,s'} = \frac{\lambda_s \lambda_{s'} \xi''_{s,s'}}{\chi_s \chi_{s'}}, \qquad \tilde{m}_s(z; \vec{v}) = \frac{\chi_s}{\lambda_s} u_s(\vec{v} + z\vec{\chi}), \qquad \tilde{x}_s = \frac{v_s \sqrt{\lambda_s}}{\chi_s}, \tag{4.121}$$

which we note match $\xi''_{s,s'}$, m_s, x_s when $\vec{\chi} = \vec{\lambda}$. Then the Dyson equation rearranges to

$$\frac{\tilde{x}_s}{\sqrt{\lambda_s}} + z = -\frac{1}{\tilde{m}_s(z; \vec{v})} - \sum_{s' \in \mathscr{S}} \frac{\tilde{\xi}''_{s,s'}}{\lambda_s} \tilde{m}_{s'}(z; \vec{v}).$$

Comparing with (4.38), we find that $\tilde{\vec{m}}(z; \vec{v})$ is the limiting Stieltjes transform of the random matrix

$$\widetilde{M}_N(\tilde{\vec{x}}) = \widetilde{\boldsymbol{W}} - \mathrm{diag}(\Lambda^{-1/2}\tilde{\vec{x}} \diamond \mathbf{1}_{\mathcal{T}}), \tag{4.122}$$

where $\widetilde{\boldsymbol{W}}$ has law (4.26) but with $\tilde{\xi}''$ in place of ξ''. Denote the associated limiting spectral measure (cf. (4.39)) by

$$\widetilde{\mu}_{\vec{\chi}}(\vec{v}) \equiv \mu_{\vec{\chi}}(\tilde{\vec{x}}) \tag{4.123}$$

We recall from [AEK17a, Theorem 2.6] that $\widetilde{\mu}_{\vec{\chi}}(\vec{v})$ is supported on a finite union of intervals, which is the closure of $\{\gamma \in \mathbb{R} : \vec{u}(\vec{v} + \gamma\vec{\chi}) \in \mathbb{H}^r\}$, with *edges* at the boundary of its support and finitely many *cusps* within the support at which $\vec{u}(\vec{v} + \gamma\vec{\chi}) \in \mathbb{R}^r$. We say 0 is a *left edge* of the support of $\mu_{\vec{\chi}}(\tilde{\vec{x}})$ if it is an edge and $(0, c) \subseteq \mathsf{supp}\ \mu_{\vec{\chi}}(\tilde{\vec{x}})$ for small enough $c > 0$. A *right edge* is defined similarly. For $\gamma \in \mathbb{R}$ and $\vec{\chi}$ as above, we set $\vec{u}^{\gamma}_{\vec{\chi}} = \vec{u}(\vec{v} + \gamma\vec{\chi})$.

Proposition 4.A.7. *Suppose $\vec{v} \in \mathbb{R}^r$, $\vec{u} = \vec{u}(\vec{v}) \in \mathbb{R}^r$, and $M(\vec{u})$ is singular. Fix $\vec{\chi} \in \mathbb{R}^r_{>0}$ with $\|\vec{\chi}\|_1 = 1$. Then 0 is an edge or cusp of $\widetilde{\mu}_{\vec{\chi}}(\vec{v})$. In more detail, there exists $\gamma_0 > 0$ such that for each $\Delta \in \{\pm 1\}$, one of the following holds.*

(i) For all $\gamma \in (0, \gamma_0]$, $\vec{u}^{\gamma\Delta}_{\vec{\chi}} \in \mathbb{R}^r$ and $M(\vec{u}^{\gamma\Delta}_{\vec{\chi}}) \succ 0$.

(ii) For all $\gamma \in (0, \gamma_0]$, $\vec{u}^{\gamma\Delta}_{\vec{\chi}} \in \mathbb{H}^r$.

Moreover case (ii) holds for at least one $\Delta \in \{\pm 1\}$, and:

(a) If case (i) holds for $\Delta = 1$, then 0 is a right edge of $\widetilde{\mu}_{\vec{\chi}}(\vec{v})$.

(b) If case (i) holds for $\Delta = -1$, then 0 is a left edge of $\widetilde{\mu}_{\vec{\chi}}(\vec{v})$.

(c) If case (ii) holds for both $\Delta \in \{\pm 1\}$, then 0 is a cusp of $\widetilde{\mu}_{\vec{\chi}}(\vec{v})$.

Finally, which of (a),(b),(c) occurs for a given $\vec{v}$ does not depend on $\vec{\chi}$.

Remark 4.A.8. The $\vec{\chi}$-independence of being an edge or cusp is consistent with Figures 4.4.1b and 4.4.1c. Recall that in these plots, $\vec{u}(0;\vec{v})$ is real in the four regions outside the blue boundary and nonreal in the region inside it. An edge corresponds to a point on the blue boundary where a positive-slope line in direction $\vec{\chi}$ through $\vec{v}$ crosses from a real region to a nonreal region. A cusp (two in each plot) corresponds to a point where such a line remains in the nonreal region on either side of $\vec{v}$. In both pictures, this property is independent of the slope $\vec{\chi}$.

Our final result computes the annealed exponential growth rate of the determinant of the deformed Gaussian band matrix $M_N(\vec{x})$ defined in (4.51). Namely we give an explicit formula for $\Psi(\vec{x})$, defined in (4.44). Recall from the proof of Proposition 4.3.2 that this equals

$$\lim_{N\to\infty} \frac{1}{N} \log \mathbb{E}|\det M_N(\vec{x})|.$$

Recall that $\langle \vec{a}, \vec{b}\rangle = \sum_{i=1}^r a_i b_i$ denotes a bilinear form rather than a complex inner product, even when $\vec{a}, \vec{b}$ are complex vectors.

Theorem 4.A.9. *Let $\vec{x} \in \mathbb{R}^r$ and $\vec{v} = \Lambda^{1/2}\vec{x}$. Then*

$$\Psi(\vec{x}) = \frac{1}{2}\Re(\langle \vec{u}(\vec{v}), \xi''\vec{u}(\vec{v})\rangle) - \sum_{s\in\mathscr{S}} \lambda_s \log|u_s(\vec{v})|.$$

4.A.1 Preliminaries

Lemma 4.A.10. *Suppose $A \in \mathbb{R}^{r\times r}$ is diagonally signed, $A \succeq 0$, and $A' \in \mathbb{C}^{r\times r}$ satisfies:*

$$\begin{aligned} |A'_{s,s}| &\geq A_{s,s}, \quad \forall s \in \mathscr{S}; \\ |A'_{s,s'}| &\leq |A_{s,s'}|, \quad \forall s \neq s' \in \mathscr{S}. \end{aligned} \tag{4.124}$$

Then:

(a) If A is invertible, A^{-1} has only positive entries.

(b) If A is invertible, then so is A' and $\|(A')^{-1}\|_{\mathsf{op}} \leq \|A^{-1}\|_{\mathsf{op}}$.

(c) If at least one inequality in (4.124) holds strictly, then A' is invertible.

Proof. Suppose A is invertible. Then $A \succ 0$, so $A_{s,s} > 0$ for all $s \in \mathscr{S}$. Let $D = \operatorname{diag}(A)^{1/2}$, so $A = D(I-B)D$ for some $B \in \mathbb{R}^{r\times r}$ with zero diagonal and positive entries off the diagonal, and with $I - B \succ 0$. Let $t = \boldsymbol{\lambda}_{\min}(I-B) \in (0,1)$, so $(1-t)I - B \succeq 0$. By Lemma 4.2.8 (applied to $(1-t)I - B$), $(1-t)I + B \succeq 0$. Thus $(1-t)I \succeq B \succeq -(1-t)I$. So,

$$A^{-1} = (D(I-B)D)^{-1} = D^{-1}(I + B + B^2 + \cdots)D^{-1},$$

as the geometric series converges. Since D and B have positive entries, part (a) follows.

There exists diagonal $\widetilde{D} \in \mathbb{C}^{r\times r}$ such that $D^2\widetilde{D}^2 = \operatorname{diag}(A')$, and (4.124) implies $|\widetilde{D}_{s,s}| \geq 1$ for all $s \in \mathscr{S}$. Then $A' = D\widetilde{D}(I - \widetilde{B})\widetilde{D}D$. Note that for all $s, s' \in \mathscr{S}$, $|\widetilde{B}_{s,s'}| \leq B_{s,s'}$, and therefore for all $k \geq 1$, $|(\widetilde{B}^k)_{s,s'}| \leq (B^k)_{s,s'}$. Thus

$$(A')^{-1} = D^{-1}\widetilde{D}^{-1}(I + \widetilde{B} + \widetilde{B}^2 + \cdots)\widetilde{D}^{-1}D^{-1},$$

as the geometric series converges. For any $\vec{x} \in \mathbb{C}^r$, consider $\vec{y} \in \mathbb{R}^r$ defined by $y_s = |x_s|$. Then it is clear that for all $s \in \mathscr{S}$, $|((A')^{-1}\vec{x})_s| \le (A^{-1}\vec{y})_s$, so $\left\|(A')^{-1}\right\|_{\mathsf{op}} \le \left\|A^{-1}\right\|_{\mathsf{op}}$. This proves part (b).

Finally consider the setting of part (c), where $A \succeq 0$ is not necessarily invertible and at least one inequality in (4.124) is strict. Let $\widetilde{A} \in \mathbb{R}^{r\times r}$ be the diagonally signed matrix with $\widetilde{A}_{s,s} = |A'_{s,s}|$ and $\widetilde{A}_{s,s'} = -|A'_{s,s'}|$ for $s \neq s'$. Let $\vec{w}$ be the minimal (unit) eigenvector of $\widetilde{A}$, which by Lemma 4.2.7 has all positive entries. Then

$$\boldsymbol{\lambda}_{\min}(\widetilde{A}) = \langle \widetilde{A}\vec{w}, \vec{w}\rangle > \langle A\vec{w}, \vec{w}\rangle \ge \boldsymbol{\lambda}_{\min}(A) \ge 0,$$

so $\widetilde{A} \succ 0$. Thus $\widetilde{A}$ is invertible, and by part (b) (with $\widetilde{A}$ for A) so is A'. This proves part (c). □

The following part of the proof of Theorem 4.A.5 will be used repeatedly, so we prove it first. It is related to [AEK17a, Lemma 4.3] (namely the operator F appearing there is similar to $\overline{M}$).

Lemma 4.A.11. *For any $\vec{v} \in \mathbb{H}^r$, with $\vec{u} = \vec{u}(\vec{v})$ we have $\overline{M}(\vec{u}) \succ 0$.*

Proof. Taking imaginary parts of (4.119) yields

$$\frac{\lambda_s}{|u_s|^2}\Im(u_s) - \sum_{s'\in\mathscr{S}} \xi''_{s,s'}\Im(u_{s'}) = \Im(v_s). \tag{4.125}$$

Since $\Im(v_s) > 0$, we have $\overline{M}(\vec{u})\Im(\vec{u}) \succ 0$. This implies $\overline{M}(\vec{u}) \succ 0$ by Lemma 4.2.7. □

Lemma 4.A.12. *If $\vec{u} \in \mathbb{H}^r$ and $\overline{M}(\vec{u}) \succeq 0$, then $M(\vec{u})$ is invertible.*

Proof. Since $\overline{M}(\vec{u}) \succeq 0$, $\lambda_s/|u_s|^2 > \xi''_{s,s}$. So, for any $s \in \mathscr{S}$,

$$\left|\frac{\lambda_s}{u_s^2} - \xi''_{s,s}\right| > \frac{\lambda_s}{|u_s|^2} - \xi''_{s,s},$$

where the inequality is strict because $\vec{u} \in \mathbb{H}^r$. Taking $(A, A') = (\overline{M}(\vec{u}), M(\vec{u}))$ in Lemma 4.A.10(c) yields the claim. □

Corollary 4.A.13. *For any $\vec{v} \in \mathbb{H}^r$, with $\vec{u} = \vec{u}(\vec{v})$, $M(\vec{u})$ is invertible.*

Proof. Follows from Lemmas 4.A.11 and 4.A.12. □

Lemma 4.A.14. *There exists $C_0 > 0$ depending on $(\vec{\lambda}, \xi'')$ such that for all $\vec{v} \in \mathbb{H}^r$, $\|\vec{u}(\vec{v})\|_\infty \le C_0$.*

Proof. Let $\vec{u} = \vec{u}(\vec{v})$. By Lemma 4.A.11, $\overline{M}(\vec{u}) \succ 0$, so $\overline{M}(\vec{u})_{s,s} > 0$. Thus $|u_s| \le \sqrt{\lambda_s/\xi''_{s,s}}$. □

4.A.2 Joint continuity of the vector Dyson equation

In this subsection, we let C_0 be as in Lemma 4.A.14, let C_1 be a sufficiently large constant depending on $(\vec{\lambda}, \xi'', C_0)$, and similarly take large C_2 depending on $(\vec{\lambda}, \xi'', C_0, C_1)$ and C_3 large depending on $(\vec{\lambda}, \xi'', C_0, C_1, C_2)$. Given any $S \subseteq \mathscr{S}$, define

$$V_S = \{\vec{v} \in \mathbb{H}^r \ : \ |v_s| \le C_1 \ \forall s \in S \text{ and } |v_s| \ge C_1 \ \forall s \notin S\} \subseteq \mathbb{H}^r. \tag{4.126}$$

We will show Hölder continuity of the restriction of $\Im(\vec{u})$ to each set V_S.

Lemma 4.A.15. *For each $S \subseteq \mathscr{S}$, the restriction of $\vec{u} : \mathbb{H}^r \to \mathbb{H}^r$ to V_S satisfies $\|\Im(\vec{u})|_{V_S}\|_{C^{1/3}} < \infty$.*

Lemma 4.A.15 readily implies that $\|\Im(\vec{u})\|_{C^{1/3}} < \infty$ holds on all of $\mathbb{H}^r$, thus establishing "half of" Theorem 4.A.2. Namely given $\vec{v}, \vec{v}' \in \mathbb{H}^r$, along the path $(\vec{v} + t(\vec{v}' - \vec{v}))_{t\in[0,1]}$ the s-th coordinate's norm switches between $[0, C_1]$ and $[C_1, \infty)$ at most twice. Hence $\Im(\vec{u}(\vec{v})) - \Im(\vec{u}(\vec{v}'))$ can be bounded by applying Lemma 4.A.15 at most $2r + 1$ times along this path.

We prove Lemma 4.A.15 after establishing some helpful intermediate results.

Lemma 4.A.16. *For C_1 as described above (i.e. sufficient large depending on $(\vec{\lambda}, \xi'', C_0)$), the following holds.*

(a) For all $s \in S$, $|u_s| \geq \lambda_s/2C_1$.

(b) For all $s \notin S$, $|u_s| \leq 2\lambda_s/C_1$.

Proof. Equation (4.119) implies

$$|v_s| - \sum_{s' \in \mathscr{S}} \xi''_{s,s'}|u_{s'}| \leq \frac{\lambda_s}{|u_s|} \leq |v_s| + \sum_{s' \in \mathscr{S}} \xi''_{s,s'}|u_{s'}|. \tag{4.127}$$

In light of Lemma 4.A.14, we have

$$\sum_{s' \in \mathscr{S}} \xi''_{s,s'}|u_{s'}| \leq C_1/2$$

for suitably large C_1 depending only on $(\vec{\lambda}, \xi'')$. For $s \in \mathscr{S}$, the right inequality of (4.127) implies $\lambda_s/|u_s| \leq 2C_1$, which implies part (a). For $s \notin \mathscr{S}$, the left inequality implies $\lambda_s/|u_s| \geq C_1/2$, which implies part (b). □

Lemma 4.A.17. *For $s, s' \in S$,*

$$C_2^{-1}\Im(u_s) \leq \Im(u_{s'}) \leq C_2\Im(u_s).$$

Proof. Taking imaginary parts of (4.119) yields (4.125). In light of Lemma 4.A.16(a), this implies

$$\frac{4C_1^2}{\lambda_s}\Im(u_s) \geq \frac{\lambda_s}{|u_s|^2}\Im(u_s) \geq \xi''_{s,s'}\Im(u_{s'}).$$

Since such an inequality holds for all $s, s' \in S$ the conclusion follows. □

Lemma 4.A.18. *The function $\vec{u}$ is differentiable on $\mathbb{H}^r$ with Jacobian $\nabla\vec{u}(\vec{v}) = M(\vec{u}(\vec{v}))^{-1}$ (which is invertible by Corollary 4.A.13).*

Proof. Let $\vec{v} \in \mathbb{H}^r$ and $\vec{u} = \vec{u}(\vec{v}) \in \mathbb{H}^r$. Then $\vec{v}(\vec{u}) = \vec{v}$. The function $\vec{v}(\cdot)$ is clearly continuous, so it maps an open neighborhood $\mathcal{N} \subset \mathbb{H}^r$ of $\vec{u}$ into $\vec{v}(\mathcal{N}) \subset \mathbb{H}^r$. By Lemma 4.A.1, this is a bijective map with inverse $\vec{u}(\cdot)$. Moreover $\vec{v}(\cdot)$ is differentiable, with Jacobian

$$\nabla\vec{v}(\vec{u}) = M(\vec{u}), \tag{4.128}$$

and this is invertible by Corollary 4.A.13. The result follows by the inverse function theorem. □

Lemma 4.A.19. *For each $s_* \in S$, and distinct $\vec{v}, \widetilde{\vec{v}} \in V_S$,*

$$\frac{\|\Im(u_{s_*}(\vec{v})) - \Im(u_{s_*}(\widetilde{\vec{v}}))\|_2}{\|\vec{v} - \widetilde{\vec{v}}\|_2^{1/3}} \leq C_3. \tag{4.129}$$

Proof. Write $\vec{u} = \vec{u}(\vec{v})$. By Lemma 4.A.18, $\nabla\vec{u}(\vec{v}) = M^{-1}(\vec{u})$. We show that for $\vec{v} \in V_S$,

$$\|M^{-1}(\vec{u})\|_{\mathsf{op}} \leq C_3\Im(u_{s_*})^{-2}. \tag{4.130}$$

To deduce (4.129) from this, first note that for any smooth path $\gamma : [0,1] \to V_S$, (4.130) implies

$$\left|\frac{\mathsf{d}}{\mathsf{d}t}\left[\Im\big(u_{s_*}(\gamma(t))\big)^3\right]\right| \leq C_3\gamma'(t).$$

This implies $\Im(u_{s_*})^3$ is Lipschitz on V_S because for any $\vec{v}, \widetilde{\vec{v}} \in V_S$ there exists γ as above with $\big(\gamma(0), \gamma(1)\big) = (\vec{v}, \widetilde{\vec{v}})$ and $\int_0^1 |\gamma'(t)|\mathsf{d}t \leq 10r\|\vec{v} - \widetilde{\vec{v}}\|_2$. Since $\Im(u_{s_*})$ is uniformly bounded by Lemma 4.A.14, the fact that $\Im(u_{s_*})^3$ is Lipschitz immediately yields (4.129).

To show (4.130), with $\varepsilon = C_2 \Im(u_{s_*}) > 0$, we have $\Im(u_s) \geq \varepsilon$ for all $s \in S$ by Lemma 4.A.17. Define the matrix

$$M^\dagger_{s,s'} = \begin{cases} |M_{s,s}| = |\lambda_s/u_s^2 - \xi''_{s,s}|, & s = s' \in \mathscr{S}, \\ -\xi''_{s,s'}, & s \neq s' \in \mathscr{S}. \end{cases}.$$

Thus $M^\dagger$ agrees with M off of the diagonal. On the diagonal, we claim that

$$M^\dagger_{s,s} \geq \overline{M}_{s,s} + \Omega(\varepsilon^2) \cdot \mathbf{1}_{s \in S}.$$

This is easy to see geometrically: given $|u_s|$, the entry $M^\dagger_{s,s}$ varies on a circle, and its radius is $\sqrt{\lambda_s}/|u_s| \asymp 1$ since $s \in S$, and its distance from the center is also $\xi''_{s,s} \asymp 1$. By Lemma 4.A.10, it follows that $M^\dagger$ is strictly positive definite since $\overline{M} \succeq 0$.

We claim that in fact

$$M^\dagger \succeq \Omega(\varepsilon^2) I_r. \tag{4.131}$$

Indeed let $\vec{y} \in \mathbb{R}^r$ be a unit vector and let $\vec{y}_S \in \mathbb{R}^r$ agree with $\vec{y}$ on coordinates in S and have zero coordinates otherwise. If $\|\vec{y}_S\|_2^2 \geq 1/2$, then

$$\langle \vec{y}, M^\dagger \vec{y} \rangle \geq \langle \vec{y}, \overline{M} \vec{y} \rangle + \Omega(\varepsilon^2 \|\vec{y}_S\|_2^2) \geq \Omega(\varepsilon^2).$$

Otherwise, suppose $\|\vec{y}_{S^c}\|_2^2 \geq 1/2$. Define

$$C'_0 = \max_{s,s' \in \mathscr{S}} \xi''_{s,s'}, \qquad C'_1 = \min_{s \in \mathscr{S}} \left\{ \frac{C_1^2}{4\lambda_s} - \xi''_{s,s} \right\}.$$

Lemma 4.A.16(b) implies $M^\dagger_{s,s} \geq C'_1$ for all $s \notin S$ while $M^\dagger_{s,s'} \geq -C'_0$ for all $s, s' \in \mathscr{S}$. So

$$\langle \vec{y}, M^\dagger \vec{y} \rangle \geq C'_1 \|\vec{y}_{S^c}\|_2^2 - r C'_0 \geq 1$$

if C_1 is suitably large. Combining cases proves (4.131) since $\vec{y}$ was an arbitrary unit vector.

Thus $\|(M^\dagger)^{-1}\|_{\mathsf{op}} \leq O(\varepsilon^{-2})$. Applying Lemma 4.A.10(b) with $(A, A') = (M^\dagger, M)$ shows that $\|M^{-1}\|_{\mathsf{op}} \leq \|(M^\dagger)^{-1}\|_{\mathsf{op}}$, establishing (4.130) as desired. □

Proof of Lemma 4.A.15. Let $B(R) \subseteq \mathbb{C}$ denote the radius R ball. Note that given *any* $(\vec{v}, \vec{u}_S) \in \mathbb{H}^r \times B(C_0)^{|S|}$, the complementary vector $\vec{u}_{S^c} = \vec{u} - \vec{u}_S$ may be defined by the equations in (4.119) for $s \in S^c$. Restricting the domain slightly to $\mathcal{D}_S = (\mathbb{C} \backslash B(C_1))^r \times B(C_0)^{|S|}$, this defines a map

$$\varphi_S : \mathcal{D}_S \to B(C_0)^{|S^c|}.$$

Note that the restriction M_{S^c} of M to coordinates $S^c \times S^c$ satisfies $\|M_{S^c}(\vec{u})^{-1}\|_{\mathsf{op}} \geq 1$ on the domain of φ_S since C_1 is large compared to C_0. It follows that $\|\nabla \varphi_S\| \leq O(1)$ holds everywhere on $\mathcal{D}_S$.

Finally just as in the proof of Lemma 4.A.19, for any pair of points in $\mathcal{D}_S$, there is a smooth path $\gamma : [0,1] \to \mathcal{D}_S$ connecting them with total length at most the Euclidean distance between them. Therefore φ_S is $O(1)$-Lipschitz on $\mathcal{D}_S$, and so using Lemma 4.A.19, for any $\vec{v}, \widetilde{\vec{v}} \in V_S$

$$\begin{aligned} \|\vec{u}(\vec{v}) - \vec{u}(\widetilde{\vec{v}})\|_2 &\leq \|\vec{u}_S(\vec{v}) - \vec{u}_S(\widetilde{\vec{v}})\|_2 + \|\vec{u}_{S^c}(\vec{v}) - \vec{u}_{S^c}(\widetilde{\vec{v}})\|_2 \\ &\lesssim \|\vec{u}_S(\vec{v}) - \vec{u}_S(\widetilde{\vec{v}})\|_2 + \big(\|\vec{u}_S(\vec{v}) - \vec{u}_S(\widetilde{\vec{v}})\|_2 + \|\vec{v} - \widetilde{\vec{v}}\|_2\big) \\ &\lesssim \|\vec{v} - \widetilde{\vec{v}}\|_2^{1/3} \cdot (1 + \|\vec{v} - \widetilde{\vec{v}}\|_2). \end{aligned}$$

Recalling that $\vec{u}$ is uniformly bounded now completes the proof. □

It follows immediately from the preceding result that $\Im\vec{u}$ extends to a $C^{1/3}$ function on $\overline{\mathbb{H}}^r$. It remains to show the same for $\Re\vec{u}$. Similarly to [AEK17a, Proposition 5.1], along any given 1-dimensional subspace of $\mathbb{R}^r$, $\Re\vec{u}$ can be obtained via the Stieltjes transform of the continuous boundary extension of $\Im\vec{u}$, which automatically inherits the 1/3-Hölder continuity of $\Im\vec{u}$. Since we aim to show continuity in $\vec{v} \in \overline{\mathbb{H}}^r$, these 1-dimensional Stieltjes transforms must be patched together. Consistency of the extensions to intersecting lines in $\mathbb{R}^r$ will follow from the existence of non-tangential limits as recalled below.

Definition 4.A.20. Given $v \in \mathbb{R}$ and $\theta \in (0, \pi/2)$, define the cone

$$\mathsf{Cone}_\theta(v) = \big\{y \in \mathbb{H} \ : \ \arg(y - v) \in [\pi/2 - \theta, \pi/2 + \theta]\big\} \subseteq \mathbb{H}.$$

Given $\vec{v} \in \mathbb{R}^r$ and $\theta_1, \dots, \theta_r \in (0, \pi/2)$, define the product cone

$$\mathsf{Cone}_{\vec{\theta}}(\vec{v}) = \prod_{s=1}^{r} \mathsf{Cone}_{\theta_s}(v_s) \subseteq \mathbb{H}^r.$$

Proposition 4.A.21 ([SW71, Special Case of Theorem 3.24 in Chapter 2]). *Let $\vec{u} : \mathbb{H}^r \to \mathbb{H}^r$ be a bounded harmonic function. Then for almost every $\vec{v} \in \mathbb{R}^r$, the following **non-tangential limit** exists and is uniformly bounded for all $\vec{\theta} \in (0, \pi/2)^r$:*

$$\vec{u}^{(\mathbb{R})}(\vec{v}) \equiv \lim_{\substack{\vec{y} \to \vec{v} \\ \vec{y} \in \mathsf{Cone}_{\vec{\theta}}(\vec{v})}} \vec{u}(\vec{y}). \tag{4.132}$$

*We call those $\vec{v} \in \mathbb{R}^r$ with this property **regular points** for $\vec{u}$.*

It is well-known that a bounded harmonic function on $\overline{\mathbb{H}}$ can be recovered as the Poisson integral of its boundary values, see e.g. [Rud87, Theorem 11.30(b)]. Below we give an extension to higher dimensions which suffices for our purposes. Define for $y \in \mathbb{H}, \vec{y} \in \mathbb{H}^r$ the univariate and multivariate Poisson kernels:

$$K(y) = \frac{1}{\pi \|y\|^2}, \qquad K(\vec{y}) = \prod_{s=1}^{r} \vec{K}(y_s). \tag{4.133}$$

Note that $K(y)$ is a probability density on each shift $\mathbb{R} + it$ for $t > 0$. We view $K(y)$ as a point mass at y if $y \in \mathbb{R}$, and $\vec{K}$ as the corresponding product measure for $\vec{y} \in \overline{\mathbb{H}}^r$.

Proposition 4.A.22. *Suppose $\vec{u} : \mathbb{R}^r \to \mathbb{R}$ as defined in Proposition 4.A.21 agrees with a continuous bounded function $\vec{u}^{(\mathbb{R})} : \mathbb{R}^r \to \mathbb{R}$ almost everywhere. Then $\vec{u}$ extends to a bounded continuous function on $\overline{\mathbb{H}}^r$ agreeing with $\vec{u}^{(\mathbb{R})}$ on $\mathbb{R}^r$ and admitting the Poisson integral representation*

$$\vec{u}(\vec{y}) = \int_{\mathbb{R}^r} \vec{K}(\vec{y} - \vec{v}) \vec{u}^{(\mathbb{R})}(\vec{v}) \ \mathsf{d}\vec{v}, \quad \vec{y} \in \overline{\mathbb{H}}^r. \tag{4.134}$$

Proof. First, the above definition of the Poisson integral agrees (in the case that $\vec{u}^{(\mathbb{R})}$ is uniformly bounded) with that of [SW71, Chapter 2 page 67] as an iterated application of univariate Poisson integrals. By r-fold application of [SW71, Chapter 2 Theorem 2.1(b)], it follows that the right-hand side of (4.134) is continuous and bounded on $\overline{\mathbb{H}}^r$. Call this right-hand side $\widetilde{\vec{u}}$. Since each probability measure $K(y_s - v_s)\mathsf{d}v_s$ converges weakly to a point mass at $\Re y_s$ as $\Im y_s \downarrow 0$, it follows that $\widetilde{\vec{u}}$ is continuous on $\overline{\mathbb{H}}^r$.

It remains to show that $\vec{u}$ and $\widetilde{\vec{u}}$ agree on $\mathbb{H}^r$. Hence fix $\vec{y} \in \mathbb{H}^r$. Since both functions are harmonic and bounded on $\mathbb{H}^r$, we have the upward-shifted Poisson integral representations for $\varepsilon \in (0, \min_s \Im y_s)$:

$$\begin{aligned} \vec{u}(\vec{y}) &= \int_{\mathbb{R}^r} \vec{K}\big(\vec{y} - \vec{v} - \varepsilon i \vec{1}\big) \vec{u}(\vec{v} + \varepsilon i \vec{1}) \ \mathsf{d}\vec{v}, \\ \widetilde{\vec{u}}(\vec{y}) &= \int_{\mathbb{R}^r} \vec{K}\big(\vec{y} - \vec{v} - \varepsilon i \vec{1}\big) \widetilde{\vec{u}}(\vec{v} + \varepsilon i \vec{1}) \ \mathsf{d}\vec{v}. \end{aligned} \tag{4.135}$$

The functions $\vec{u}(\vec{v} + \varepsilon i \vec{1})$ and $\widetilde{\vec{u}}(\vec{v} + \varepsilon i \vec{1})$ on $\mathbb{R}^r$ are uniformly bounded and converge almost everywhere to the same limit $\vec{u}^{(\mathbb{R})}$ as $\varepsilon \downarrow 0$. Moreover for each fixed $\vec{y} \in \mathbb{H}^r$, the kernel densities $\vec{K}(\vec{y} - \vec{v} - \varepsilon i \vec{1})$ are all probability measures, and they converge in total variation to $\vec{K}(\vec{y} - \vec{v})$ as $\varepsilon \downarrow 0$. It follows that the $\varepsilon \downarrow 0$ limits of the right-hand sides in (4.135) agree. Hence the left-hand sides also agree as desired. □

Proof of Theorem 4.A.2. It follows by Lemma 4.A.15 and the following discussion that $\vec{v} \mapsto \Im(\vec{u}(\vec{v}))$ is a uniformly 1/3-Hölder continuous function on $\mathbb{H}^r$. We use [Gar07, Theorem 3.5] which states that if a bounded holomorphic function $\varphi : \mathbb{H} \to \mathbb{H}$ satisfies

$$\lim_{A \to \infty} -iA\varphi(iA) = W > 0, \tag{4.136}$$

then φ is the Stieltjes transform of a positive measurable density on $\mathbb{R}$ with total integral W, which is given as as almost-everywhere limit of functions $\Im(\varphi(x+i\eta))$ as $\eta \downarrow 0$. We consider for each $\vec{y} \in [1/2,2]^r$ and $\vec{z}_* \in \mathbb{R}^r$ and $s \in \mathscr{S}$ the function

$$\varphi_{\vec{y},s}(z) = \vec{u}_s(\vec{z}_* + z\vec{y}), \quad z \in \mathbb{H}.$$

Then it is easy to see that the condition (4.136) holds with $W = \lambda_s/y_s$.

Consider now the lines $\ell(\vec{z}_*, \vec{y}) = \{\vec{z}_* + z\vec{y}\}_{z\in\mathbb{R}}$ for $\vec{y} \in [1/2,2]^r$ and $\vec{z}_* \in \mathbb{H}^r$. For each $\ell(\vec{z}_*, \vec{y})$, taking the Stieltjes transform of $\varphi_{\vec{y},s}$ gives a function $u_s(\cdot\,; \vec{z}_*, \vec{y}) : \ell(\vec{z}_*, \vec{y}) \to \overline{\mathbb{H}}$. Recall that Stieltjes transforms increase $C^{1/3}$ norms by at most a constant factor (see e.g. [MR08, Section 22]). Since $\Im\vec{u}$ is uniformly $1/3$-Hölder, it follows that each $\vec{u}(\cdot\,; \vec{z}_*, \vec{y})$ and in particular its real part has uniformly bounded $C^{1/3}$ norm on its corresponding domain $\ell(\vec{z}_*, \vec{y})$.

Next whenever $\vec{z}_* + z\vec{y} \in \mathbb{R}^r$ is regular, we have

$$\vec{u}(z; \vec{z}_*, \vec{y}) = \lim_{\varepsilon\downarrow 0} \varphi_{\vec{y},s}(z+i\varepsilon) = \vec{u}^{(\mathbb{R})}(\vec{z}_* + z\vec{y}).. \tag{4.137}$$

(The first equality holds by continuity properties of ordinary Stieltjes transforms, and the second by definition of a regular point.) Thus let $\vec{v}, \vec{v}' \in \mathbb{R}^r$ be regular points for $\vec{u}$, and let $\widetilde{\vec{v}}$ be another regular point such that with $\|\vec{v} - \vec{v}'\|_\infty = M$, we have $\widetilde{v}_s - \vec{v}_s \in [3M, 4M]$ for each s. (Such $\widetilde{\vec{v}}$ exists by Proposition 4.A.21.) Then $\frac{v_s - \widetilde{v}_s}{v_{s'} - \widetilde{v}_{s'}} \in [1/2, 2]$ for each $s, s' \in \mathscr{S}$, which means there is some $\ell(\vec{z}_*, \vec{y})$ passing through $(\vec{v}, \widetilde{\vec{v}})$, and similarly a $\ell(\vec{z}'_*, \vec{y}')$ passing through $(\vec{v}', \widetilde{\vec{v}})$. Using $1/3$-Hölder continuity on these lines together with (4.137) in the latter step, we thus obtain:

$$\begin{aligned}\|\vec{u}^{(\mathbb{R})}(\vec{v}) - \vec{u}^{(\mathbb{R})}(\vec{v}')\| &\le \|\vec{u}^{(\mathbb{R})}(\vec{v}) - \vec{u}^{(\mathbb{R})}(\widetilde{\vec{v}})\| + \|\vec{u}^{(\mathbb{R})}(\vec{v}') - \vec{u}^{(\mathbb{R})}(\widetilde{\vec{v}})\| \\ &\le O(M^{1/3}) = O(\|\vec{v} - \vec{v}'\|_\infty).\end{aligned}$$

Hence the restriction of $\vec{u}^{(\mathbb{R})}$ to regular $\vec{v} \in \mathbb{R}^r$ is uniformly $C^{1/3}$. By Proposition 4.A.21, it follows that $\vec{u}^{(\mathbb{R})}$ admits a bounded continuous extension to all of $\mathbb{R}^r$. By Proposition 4.A.22, $\vec{u}$ extends to a bounded continuous function on $\overline{\mathbb{H}}^r$.

Finally we show $\vec{u} : \overline{\mathbb{H}}^r \to \overline{\mathbb{H}}^r$ as just defined is uniformly $C^{1/3}$ on its full domain. Fix $\vec{v}, \vec{v}' \in \overline{\mathbb{H}}^r$. For each $s \in \mathscr{S}$, let $\boldsymbol{B}_s(t)$ be a standard complex Brownian motion. Define the processes

$$\boldsymbol{V}_s(t) = v_s + \Im(v_s)\boldsymbol{B}_s(t \wedge \tau_s) \text{ and } \boldsymbol{V}'_s(t) = v'_s + \Im(v'_s)\boldsymbol{B}_s(t \wedge \tau_s)$$

where τ_s denotes the first time that $\Im\boldsymbol{B}_s(t) = -1$; thus $\Im(\boldsymbol{V}_s(\tau_s)) = \Im(\boldsymbol{V}'_s(\tau_s)) = 0$. Note $\tau \equiv \max_s \tau_s < \infty$ almost surely.

Since $\vec{u}$ is bounded and holomorphic, it follows that $\vec{u}(\vec{\boldsymbol{V}}(t))$ and $\vec{u}(\vec{\boldsymbol{V}}'(t))$ are both $\mathbb{C}$-valued martingales. (This also follows directly from (4.133).) Using the triangle inequality followed by $C^{1/3}$-boundedness of $\vec{u}$ on $\mathbb{R}^r$ gives

$$\begin{aligned}|\vec{u}(\vec{v}) - \vec{u}(\vec{v}')| &\le \mathbb{E}|\vec{u}(\vec{\boldsymbol{V}}(\tau))) - \vec{u}(\vec{\boldsymbol{V}}'(\tau))| \\ &\lesssim \mathbb{E}\sum_{s=1}^r |\Re(\boldsymbol{V}_s(\tau)) - \Re(\boldsymbol{V}'_s(\tau))|^{1/3} \\ &\lesssim \mathbb{E}\Big[\sum_{s=1}^r |\Re(v_s) - \Re(v'_s)|^{1/3} + \sum_{s=1}^r |\Im(v_s) - \Im(v'_s)|^{1/3} |\Re\boldsymbol{B}_s(\tau_s)|^{1/3}\Big].\end{aligned}$$

The law of $\Re\boldsymbol{B}_s(\tau_s)$ is well known to be a standard symmetric Cauchy random variable (and does not depend on $\vec{v}$ or $\vec{v}'$). In particular it has finite $1/3$ moment. Hence we find that $\|\vec{u}(\vec{v}) - \vec{u}(\vec{v}')\| \le O\big(\|\vec{v} - \vec{v}'\|_\infty^{1/3}\big)$, completing the proof. ☐

Proof of Lemma 4.A.4. Consider a sequence of functions $\vec{u}^\varepsilon(\vec{v}) = \vec{u}(\vec{v} + \varepsilon i\vec{1})$. By Lemma 4.A.18, $\nabla\vec{u}^\varepsilon(\vec{v}) = M(\vec{u}^\varepsilon(\vec{v}))^{-1}$. By Theorem 4.A.2 and invertibility of $M(\vec{u}(\vec{v}))$, as $\varepsilon \downarrow 0$ both $\vec{u}^\varepsilon(\cdot)$ and $M(\vec{u}^\varepsilon(\cdot))^{-1}$ converge locally uniformly to $\vec{u}(\cdot)$ and $M(\vec{u}(\cdot))^{-1}$. The result now follows by e.g. [Rud76, Theorem 7.17], which states that if a family of functions and their derivatives each converge uniformly, then the derivative of the limiting function is the limit of the derivatives. ☐

4.A.3 Solution space of the vector Dyson equation

In this subsection we prove Theorem 4.A.5 and Proposition 4.A.7. We first establish Theorem 4.A.5 in the setting $\vec{u}^*, \vec{v} \in \mathbb{H}^r$.

Lemma 4.A.23. *Let $\vec{u}^* \in \mathbb{H}^r$. There exists $\vec{v} \in \mathbb{H}^r$ such that $\vec{u}^* = \vec{u}(\vec{v})$ if and only if $\overline{M}(\vec{u}^*) \succ 0$ and $\overline{M}(\vec{u}^*)\Im(\vec{u}^*) \succ \vec{0}$.*

Remark 4.A.24. In the setting of this lemma $\Im(\vec{u}^*) \succ \vec{0}$, so $\overline{M}(\vec{u}^*)\Im(\vec{u}^*) \succ \vec{0}$ implies $\overline{M}(\vec{u}^*) \succ 0$ by Lemma 4.2.7. We have written the lemma in this form for consistency with Theorem 4.A.5.

Proof. If $\vec{u}^* = \vec{u}(\vec{v})$ for some $\vec{v} \in \mathbb{H}^r$, Lemma 4.A.11 shows $\overline{M}(\vec{u}^*) \succ 0$, and the proof of that lemma shows $\overline{M}(\vec{u}^*)\Im(\vec{u}^*) \succ \vec{0}$. Conversely, suppose $\overline{M}(\vec{u}^*)\Im(\vec{u}^*) \succ \vec{0}$. Define $\vec{v} = \vec{v}(\vec{u}^*)$. Then

$$v_s = -\frac{\lambda_s}{|u_s^*|^2}\bar{u}_s^* - \sum_{s' \in \mathscr{S}} \xi''_{s,s'} u^*_{s'},$$

so $\Im(\vec{v}) = \overline{M}(\vec{u}^*)\Im(\vec{u}^*) \succ \vec{0}$. Thus $\vec{v} \in \mathbb{H}^r$. By Lemma 4.A.1, $\vec{u}^* = \vec{u}(\vec{v})$. $\square$

Proof of Theorem 4.A.5. We first prove the forward directions of both parts. For part (a), suppose $\vec{u}^* = \vec{u}(\vec{v})$ for some $\vec{v} \in \overline{\mathbb{H}}^r$. By Lemma 4.A.23 and continuity of $\vec{u}$, $\vec{u}^*$ lies in the closure of the set defined by $\overline{M}(\vec{u}^*) \succ 0$ and $\overline{M}(\vec{u}^*)\Im(\vec{u}^*) \succ \vec{0}$, which implies the conclusion. For part (b), suppose $\vec{u}^* = \vec{u}(\vec{v})$ for some $\vec{v} \in \mathbb{R}^r$. Taking imaginary parts of (4.119) yields (4.125), which implies $\overline{M}(\vec{u}^*)\Im(\vec{u}^*) = \vec{0}$. Since $\overline{M}(\vec{u}^*)$ is diagonally signed this implies $\Im(\vec{u}^*) = \vec{0}$ or $\Im(\vec{u}^*) \succ \vec{0}$, i.e. $\vec{u}^* \in \mathbb{R}^r$ or $\vec{u}^* \in \mathbb{H}^r$. By part (a) we also have $\overline{M}(\vec{u}^*) \succeq 0$. If $\vec{u}^* \in \mathbb{R}^r$, then $M(\vec{u}^*) = \overline{M}(\vec{u}^*) \succeq 0$, so conclusion (i) holds. If $\vec{u}^* \in \mathbb{H}^r$, conclusion (ii) holds.

We turn to the converses, beginning with part (a). Suppose $\overline{M}(\vec{u}^*) \succeq 0$ and $\overline{M}(\vec{u}^*)\Im(\vec{u}^*) \succeq \vec{0}$. Let $\vec{v}^* = \vec{v}(\vec{u}^*)$; we will show that $\vec{u}^* = \vec{u}(\vec{v}^*)$.

Similarly to above, $\overline{M}(\vec{u}^*)\Im(\vec{u}^*) \succeq \vec{0}$ implies $\vec{u}^* \in \mathbb{R}^r$ or $\vec{u}^* \in \mathbb{H}^r$. Suppose first that $\vec{u}^* \in \mathbb{R}^r$, and further assume $\overline{M}(\vec{u}^*) \succ 0$. Recall from (4.128) that $\vec{v}$ has Jacobian $\nabla\vec{v}(\vec{u}) = M(\vec{u})$. Because $\vec{u}^* \in \mathbb{R}^r$, we have $M(\vec{u}^*) = \overline{M}(\vec{u}^*) \succ 0$. So, $\nabla\vec{v}$ is invertible in a neighborhood of $\vec{u}^*$. By the inverse function theorem, there is a local inverse $\vec{v}^{-1}$ of $\vec{v}$ satisfying

$$\nabla\vec{v}^{-1}(\vec{v}^*) = \overline{M}(\vec{u}^*)^{-1}.$$

By Lemma 4.A.10(a), $\overline{M}(\vec{u}^*)^{-1}$ has all positive entries. For small $\varepsilon > 0$ define $\vec{v}^\varepsilon = \vec{v}^* + i\varepsilon\vec{1}$ and note that

$$\frac{\mathsf{d}}{\mathsf{d}\varepsilon}\vec{v}^{-1}(\vec{v}^\varepsilon)\big|_{\varepsilon=0} = i\overline{M}(\vec{u}^*)^{-1}\vec{1} \in \mathbb{H}^r.$$

Define $\vec{u}^\varepsilon = \vec{v}^{-1}(\vec{v}^\varepsilon)$. Then, for small $\varepsilon > 0$ we have $\vec{u}^\varepsilon, \vec{v}^\varepsilon \in \mathbb{H}^r$ and $\vec{v}^\varepsilon = \vec{v}(\vec{u}^\varepsilon)$. By Lemma 4.A.1, $\vec{u}^\varepsilon = \vec{u}(\vec{v}^\varepsilon)$. Taking $\varepsilon \to 0$, continuity of $\vec{u}$ implies $\vec{u}^* = \vec{u}(\vec{v}^*)$.

Next suppose $\vec{u}^* \in \mathbb{R}^r$ and $\overline{M}(\vec{u}^*) \succeq 0$ is singular. For small $\varepsilon > 0$ define $\vec{u}^{(\varepsilon)} = (1-\varepsilon)\vec{u}^*$ and $\vec{v}^{(\varepsilon)} = \vec{v}(\vec{u}^{(\varepsilon)})$. Since $\overline{M}(\vec{u}^{(\varepsilon)}) \succ \overline{M}(\vec{u}^*) \succeq 0$, we have just shown $\vec{u}^{(\varepsilon)} = \vec{u}(\vec{v}^{(\varepsilon)})$. Continuity of $\vec{u}$ implies $\vec{u}^* = \vec{u}(\vec{v}^*)$.

Finally, suppose $\vec{u}^* \in \mathbb{H}^r$. As above, for small $\varepsilon > 0$, $\overline{M}(\vec{u}^{(\varepsilon)}) \succ 0$. Moreover, for any $s \in \mathscr{S}$,

$$\frac{\lambda_s}{|u_s^{(\varepsilon)}|^2}\Im(u_s^{(\varepsilon)}) - \sum_{s' \in \mathscr{S}} \xi''_{s,s'}\Im(u_{s'}^{(\varepsilon)}) > \frac{\lambda_s}{|u_s^*|^2}\Im(u^*) - \sum_{s' \in \mathscr{S}} \xi''_{s,s'}\Im(u^*_{s'}) \geq 0,$$

so $\overline{M}(\vec{u}^{(\varepsilon)})\Im(\vec{u}^{(\varepsilon)}) \succ \vec{0}$. Lemma 4.A.23 implies $\vec{u}^{(\varepsilon)} = \vec{u}(\vec{v}^{(\varepsilon)})$. Continuity of $\vec{u}$ implies $\vec{u}^* = \vec{u}(\vec{v}^*)$. This proves the converse to part (a).

Finally, we turn to the converse to part (b). If either of (i), (ii) holds, then $\overline{M}(\vec{u}^*) \succeq 0$ and $\overline{M}(\vec{u}^*)\Im(\vec{u}^*) \succeq \vec{0}$. We have just shown that $\vec{u}^* = \vec{u}(\vec{v}^*)$, where $\vec{v}^* = \vec{v}(\vec{u}^*)$. We easily verify that under (i) or (ii), $\vec{v}^* \in \mathbb{R}^r$, completing the proof. $\square$

Proof of Corollary 4.A.6. Theorem 4.A.5 implies $\overline{M}(\vec{u}) \succeq 0$, so the result follows from Lemma 4.A.12. $\square$

Proof of Proposition 4.A.7

Recall the notation (4.123), which will be frequently used below. We begin with the first assertion of Proposition 4.A.7, namely that singularity of M always corresponds to either an edge or cusp.

Lemma 4.A.25. *Fix $\vec{\chi} \in \mathbb{R}^r_{>0}$ with $\|\vec{\chi}\|_1 = 1$. Then $M(\vec{u}(\vec{v}))$ is singular if and only if 0 is either an edge or cusp for $\widetilde{\mu}_{\vec{\chi}}(\vec{v})$.*

Proof. First if 0 is an edge or cusp, then [AEK17a, Theorem 2.6] makes it clear that $\vec{u}$ is not locally Lipschitz in $\vec{v}$, hence the contrapositive of Lemma 4.A.4 shows $M(\vec{u})$ is singular.

In the other direction, we have seen that singularity of $M(\vec{u})$, for $\vec{u} = \vec{u}(\vec{v})$, implies $\vec{u} \in \mathbb{R}^r$. Moreover since $M(\vec{u})$ is diagonally signed, its singularity implies by Lemma 4.2.7 that there exists $\vec{w} \succ \vec{0} \in \mathbb{R}^r$ with $M(\vec{u})\vec{w} = 0$. Suppose for sake of contradiction that $0 \notin \mathsf{supp}\, \widetilde{\mu}_{\vec{\chi}}(\vec{v})$. Then by Lemma 4.2.12, and the Stieltjes transform definition of $\vec{u}$, it follows that $\gamma \mapsto \vec{u}^{\gamma}_{\vec{\chi}}(\vec{v})$ is Lipschitz for γ in a neighborhood of 0 (since $\log(x)$ is Lipschitz away from 0). A first order Taylor expansion of (4.119) (similarly to Lemma 4.A.18) then implies

$$\lim_{\gamma \downarrow 0} M(\vec{u})\big(\vec{u}^{\gamma}_{\vec{\chi}} - \vec{u}\big)/\gamma = \lim_{\gamma \downarrow 0} \big(\vec{v} + \gamma\vec{\chi} - \vec{v}\big)/\gamma = \vec{\chi}.$$

However the left-hand side above is orthogonal to $\vec{w}$ for all $\gamma \neq 0$, while $\langle \vec{w}, \vec{\chi} \rangle > 0$ since both have strictly positive entries. This is a contradiction and completes the proof. $\square$

Given absolutely continuous $\mu \in \mathcal{P}(\mathbb{R})$ and $q \in (0,1)$, let

$$\boldsymbol{\lambda}_{(q)}(\mu) = \sup\{\lambda \in \mathbb{R} \ : \ \mu((-\infty, \lambda] \leq q)\}$$

be its q-th quantile.

Lemma 4.A.26. *Suppose* $\mathrm{diag}(\vec{\chi}^{-1})(\vec{v} - \vec{v}') \in [a,b]^r$. *Then for all $q \in (0,1)$, we have*

$$\boldsymbol{\lambda}_{(q)}(\widetilde{\mu}_{\vec{\chi}}(\vec{v}')) - \boldsymbol{\lambda}_{(q)}(\widetilde{\mu}_{\vec{\chi}}(\vec{v})) \in [a,b].$$

Proof. Immediate by the Weyl inequalities applied to the eigenvalues of the $N \times N$ random matrices (4.122) whose spectra tend to $\widetilde{\mu}_{\vec{\chi}}(\vec{v}'), \widetilde{\mu}_{\vec{\chi}}(\vec{v})$. Indeed in this context, the shift from $\vec{v}$ to $\vec{v}'$ is equivalent to adding a diagonal matrix with all entries in $[a,b]$. $\square$

Lemma 4.A.27. *Given any $\vec{u}(\vec{v}) \in \mathbb{R}^r$ with $M(\vec{u})$ singular, the following are equivalent:*

(1) For all $\varepsilon_0 > 0$ sufficiently small and all $\vec{\varepsilon} \in (0, \varepsilon_0]^r$,

$$\vec{u}(\vec{v} - \vec{\varepsilon}) \in \mathbb{R}^r.$$

(2) For all $\varepsilon_0 > 0$ sufficiently small, there exists $\vec{\varepsilon} \in (0, \varepsilon_0]^r$ such that

$$\vec{u}(\vec{v} - \vec{\varepsilon}) \in \mathbb{R}^r.$$

(3) For all $\vec{\chi} \in \mathbb{R}^r_{>0}$, the density $\widetilde{\mu}_{\vec{\chi}}(\vec{v})$ has 0 as a left edge.

(4) There exists $\vec{\chi} \in \mathbb{R}^r_{>0}$ such that the density $\widetilde{\mu}_{\vec{\chi}}(\vec{v})$ has 0 as a left edge.

Proof. We will show that point (4) implies point (1), and that point (2) implies point (3), which suffices. In the first direction, if 0 is a left edge for some $\vec{\chi}$, then Lemma 4.A.26 immediately implies point (1).

In the other direction, suppose point (3) does not hold. Singularity of $M(\vec{u})$ implies via Lemma 4.A.25 that 0 is an edge or cusp of $\widetilde{\mu}_{\vec{\chi}}(\vec{v})$. Hence if 0 is not a left edge for some $\vec{\chi}$, it must be a right edge or a cusp. In either case, Lemma 4.A.26 then implies that point (2) does not hold. This completes the proof. $\square$

Proof of Proposition 4.A.7. Note that parts (1) and (2) of Lemma 4.A.27 are both independent of $\vec{\chi}$. It follows that 0 being a left edge, right edge, or cusp for $\widetilde{\mu}_{\vec{\chi}}(\vec{v})$ are also each independent of $\vec{\chi}$. Moreover the left and analogous right edge characterizations in parts (1), (2) of Lemma 4.A.27 directly correspond to case (i) of Proposition 4.A.7. This correspondence implies the result. $\square$

4.A.4 Exponential growth rate of random determinant

This subsection is devoted to the proof of Theorem 4.A.9. We adopt the same notation as in Subsection 4.4.1, setting $\overline{\Psi}(\vec{v}) = \Psi(\vec{x})$ where $\vec{v} = \Lambda^{1/2}\vec{x} \in \mathbb{R}^r$.

Lemma 4.A.28. $\overline{\Psi} : \mathbb{R}^r \to \mathbb{R}$ *is continuously differentiable, with* $\nabla\overline{\Psi}(\vec{v}) = -\Re(\vec{u}(\vec{v}))$.

The following non-rigorous calculation, which we carefully justify below, yields this formula. Due to the identification (4.120), we may freely switch between the notations $\vec{u}(\,\cdot\,;\,\cdot\,)$ and $\vec{u}(\cdot)$ in what follows.

$$\begin{aligned}
\frac{\mathsf{d}}{\mathsf{d}v_s}\overline{\Psi}(\vec{v}) &= \frac{1}{\pi}\int_{\mathbb{R}} \log|\gamma| \sum_{s'\in\mathscr{S}} \lambda_{s'} \Im\left(\frac{\mathsf{d}u_{s'}(\gamma;\vec{v})}{\mathsf{d}v_s}\right)\,\mathsf{d}\gamma \\
&= \frac{1}{\pi}\int_{\mathbb{R}} \log|\gamma| \sum_{s'\in\mathscr{S}} \lambda_{s'} \Im\left(\frac{\mathsf{d}u_{s'}(\vec{v}+\gamma\vec{\lambda})}{\mathsf{d}v_s}\right)\,\mathsf{d}\gamma \\
&\overset{(*)}{=} \frac{1}{\pi}\int_{\mathbb{R}} \log|\gamma| \sum_{s'\in\mathscr{S}} \lambda_{s'} \Im\left(\frac{\mathsf{d}u_{s}(\vec{v}+\gamma\vec{\lambda})}{\mathsf{d}v_{s'}}\right)\,\mathsf{d}\gamma \\
&= \frac{1}{\pi}\int_{\mathbb{R}} \log|\gamma| \Im\left(\frac{\mathsf{d}u_{s}(\vec{v}+\gamma\vec{\lambda})}{\mathsf{d}\gamma}\right)\,\mathsf{d}\gamma \\
&\overset{(\diamond)}{=} -\frac{1}{\pi}\int_{\mathbb{R}} \frac{1}{\gamma}\Im\left(u_s(\vec{v}+\gamma\vec{\lambda})\right)\,\mathsf{d}\gamma \\
&\overset{(\star)}{=} -\Re\left(u_s(\vec{v})\right)\,.
\end{aligned} \tag{4.138}$$

Step $(*)$ uses that $\nabla u(\vec{v}) = M(\vec{u}(\vec{v}))^{-1}$ (recall Lemma 4.A.4) is a symmetric matrix; step $(\diamond)$ integrates by parts and step $(\star)$ is a contour integral. However this is not a rigorous calculation, primarily because $M(\vec{u}(\vec{v}))^{-1}$ may be singular for $\vec{v} \in \mathbb{R}^r$.

To make this calculation rigorous, we first work on the line $\mathbb{R} + i\eta$ for $\eta > 0$, and then send $\eta \downarrow 0$ (see [BBM24, Proposition 4.9] for a similar computation). Recall from Proposition 4.2.10 that the probability densities μ_s solving the Dyson equation are uniformly compactly supported for $\vec{v} \in \mathbb{R}^r$ with $\|\vec{v}\|_\infty \le R$. It follows that for such $\vec{v}$,

$$\begin{aligned}
\Im\left(u_s(\gamma+i\eta,\vec{v})\right) &\le \frac{C(R,\eta,\xi'',\vec{\lambda})}{1+\gamma^2}, \\
\left|\Im\left(\frac{\mathsf{d}}{\mathsf{d}v_s}u_s(\gamma+i\eta,\vec{v})\right)\right| &\le \frac{C(R,\eta,\xi'',\vec{\lambda})}{1+\gamma^2}.
\end{aligned} \tag{4.139}$$

Let $\gamma \in \mathbb{R}$ and $\eta > 0$. By Lemma 4.A.4 and (4.120), $\nabla\vec{u}(\gamma+i\eta;\vec{v}) = M(\vec{v}+(\gamma+i\eta)\vec{\lambda})^{-1}$ (which is invertible by Corollary 4.A.13). Moreover this matrix is symmetric, so

$$\begin{aligned}
&\frac{1}{\pi}\int_{\mathbb{R}} \log|\gamma+i\eta| \sum_{s'\in\mathscr{S}} \lambda_{s'} \Im\left(\frac{\mathsf{d}u_{s'}(\gamma+i\eta;\vec{v})}{\mathsf{d}v_s}\right)\,\mathsf{d}\gamma \\
&= \frac{1}{\pi}\int_{\mathbb{R}} \log|\gamma+i\eta| \sum_{s'\in\mathscr{S}} \lambda_{s'} \Im\left(\frac{\mathsf{d}u_{s}(\gamma+i\eta;\vec{v})}{\mathsf{d}v_{s'}}\right)\,\mathsf{d}\gamma \\
&= \frac{1}{\pi}\int_{\mathbb{R}} \log|\gamma+i\eta|\ \Im\left(\frac{\mathsf{d}u_{s}(\gamma+i\eta;\vec{v})}{\mathsf{d}\gamma}\right)\,\mathsf{d}\gamma \\
&\overset{(\dagger)}{=} -\frac{1}{\pi}\int_{\mathbb{R}} \Re\left(\frac{1}{\gamma+i\eta}\right)\Im\left(u_s(\gamma+i\eta;\vec{v})\right)\,\mathsf{d}\gamma,
\end{aligned} \tag{4.140}$$

Step $(\dagger)$ is an integration by parts which is valid by (4.139). We use the residue theorem to evaluate the last integral. Note that

$$\int_{\mathbb{R}} \Re\left(\frac{1}{\gamma+i\eta}\right)\Im\left(u_s(\gamma+i\eta;\vec{v})\right)\,\mathsf{d}\gamma = \frac{1}{2}\Im\int_{\mathbb{R}}\left(\frac{1}{\gamma+i\eta}+\frac{1}{\gamma-i\eta}\right)u_s(\gamma+i\eta;\vec{v})\,\mathsf{d}\gamma.$$

The latter integral can be evaluated by completing the contour via a radius R semicircle in $\mathbb{H}$; the contribution of this semicircle decays to 0 with R since the uniformly compact support of μ_s implies that $|u_s(z;\vec{v})| \leq O(1/|z|)$ for large z and fixed $\vec{v}$. Hence applying the residue theorem, we complete the calculation (4.140) and obtain:

$$\begin{aligned}\frac{1}{\pi}\int_{\mathbb{R}} \log|\gamma+i\eta| \sum_{s'\in\mathscr{S}} \lambda_{s'}\Im\left(\frac{\mathsf{d}u_{s'}(\gamma+i\eta;\vec{v})}{\mathsf{d}v_s}\right)\ \mathsf{d}\gamma &= -\Re\left(u_s(2i\eta;\vec{v})\right) \\ &= -\Re\left(u_s(\vec{v}+2i\eta\vec{\lambda})\right).\end{aligned} \tag{4.141}$$

We complete the proof of Lemma 4.A.28 by taking $\eta \downarrow 0$ in (4.141). The right-hand side poses no issue since $\vec{u}$ is $1/3$-Hölder continuous by Theorem 4.A.2. For the left-hand side, we differentiate under the integral sign (which is justified using e.g. compact support of μ_s):

$$\begin{aligned}&\frac{1}{\pi}\int_{\mathbb{R}} \log|\gamma+i\eta| \sum_{s'\in\mathscr{S}} \lambda_{s'}\Im\left(\frac{\mathsf{d}u_{s'}(\gamma+i\eta;\vec{v})}{\mathsf{d}v_s}\right)\ \mathsf{d}\gamma \\ &= \frac{\mathsf{d}}{\mathsf{d}v_s}\frac{1}{\pi}\int_{\mathbb{R}} \log|\gamma+i\eta| \sum_{s'\in\mathscr{S}} \lambda_{s'}\Im\left(u_{s'}(\gamma+i\eta;\vec{v})\right)\ \mathsf{d}\gamma.\end{aligned}$$

We then take the $\eta \downarrow 0$ limit for the latter integrand.

Proposition 4.A.29. *Locally uniformly over $\vec{v} \in \mathbb{R}^r$,*

$$\lim_{\eta\downarrow 0}\int_{\mathbb{R}} \log|\gamma+i\eta| \sum_{s'\in\mathscr{S}} \lambda_{s'}\Im\left(u_{s'}(\gamma+i\eta;\vec{v})\right)\ \mathsf{d}\gamma = \int_{\mathbb{R}} \log|\gamma| \sum_{s'\in\mathscr{S}} \lambda_{s'}\Im\left(u_{s'}(\gamma;\vec{v})\right)\ \mathsf{d}\gamma$$

Proof. This follows directly by dominated convergence. The large γ contributions are controlled by (4.139), while the $\log 0$ singularity is integrable hence causes no issues. □

Proof of Lemma 4.A.28. Define

$$f_s(\eta;\vec{v}) = \int_{\mathbb{R}} \log|\gamma+i\eta| \sum_{s'\in\mathscr{S}} \lambda_{s'}\Im\left(u_{s'}(\gamma+i\eta;\vec{v})\right)\ \mathsf{d}\gamma, \quad \eta \geq 0.$$

We have shown above that:

(1) $\lim_{\eta\downarrow 0} f_s(\eta;\vec{v}) = f_s(0;\vec{v})$ holds locally uniformly in $\vec{v}$.

(2) For $\eta > 0$, we have $\frac{\mathsf{d}}{\mathsf{d}v_s} f_s(\eta;\vec{v}) = -\Re(u_s(\vec{v}+2i\eta\vec{\lambda}))$.

(3) $\Re(u_s(\vec{v}+2i\eta\vec{\lambda}))$ is continuous on $\eta \geq 0$, locally uniformly in $\vec{v}$.

Recall from e.g. [Rud76, Theorem 7.17] that if a family of functions and their derivatives each converge uniformly, then the derivative of the limiting function is the limit of the derivatives. This shows $\frac{\mathsf{d}}{\mathsf{d}v_s}\overline{\Psi}(\vec{v}) = -\Re(u_s(\vec{v}))$. Finally $\vec{u}(\cdot)$ is continuous by Theorem 4.A.2, concluding the proof. □

Proof of Theorem 4.A.9. We claim that

$$G(\vec{v}) = \overline{\Psi}(\vec{v}) - \frac{1}{2}\Re\left(\langle \vec{u}(\vec{v}), \xi''\vec{u}(\vec{v})\rangle\right) + \sum_{s\in\mathscr{S}} \lambda_s \log|u_s(\vec{v})|$$

vanishes identically on $\vec{v} \in \mathbb{R}^r$. We first check that $|G(\vec{v})| \to 0$ as $\min_s |v_s| \to \infty$. In this limit, (4.46) implies that

$$\mathbb{W}_\infty\Big(\overline{\mu}(\vec{v}), \sum_{s\in\mathscr{S}} \lambda_s \delta_{-v_s/\lambda_s}\Big)$$

is bounded independently of $\vec{v}$. Hence $\overline{\Psi}(\vec{v}) - \sum_{s\in\mathscr{S}} \lambda_s \log(|v_s|/\lambda_s)$ tends to 0 as $\min_s v_s \to \infty$. Furthermore $\vec{u}(\vec{v}) = \vec{u}(0;\vec{v}) \to 0$ in this limit, so (4.119) implies that in fact $u_s v_s \to -\lambda_s$. Thus $\overline{\Psi}(\vec{v}) + \sum_{s\in\mathscr{S}} \lambda_s \log|u_s|$ indeed tends to 0 with $\min_s |v_s|$.

Next let $\mathcal{T} \subseteq \mathbb{R}^r$ denote the set of points at which $\det M(\vec{u}(\vec{v})) = 0$. Since $\det M(\vec{u}(\vec{v}))$ is continuous, for any $\vec{v} \notin \mathcal{T}$, we may differentiate $G(\vec{v})$ using Lemma 4.A.4 to obtain $\nabla G(\vec{v}) = \vec{0}$. Indeed, letting $\vec{u} = \vec{u}(\vec{v})$, one directly verifies

$$
\begin{aligned}
&\nabla_{\vec{v}\in\mathbb{R}^r}\left(-\frac{1}{2}\Re\langle\vec{u}(\vec{v}),\xi''\vec{u}(\vec{v})\rangle + \sum_{s\in\mathscr{S}}\lambda_s \log|u_s(\vec{v})|\right)\\
&= \Re\left(\nabla_{\vec{v}\in\overline{\mathbb{H}}^r}\left(-\frac{1}{2}\langle\vec{u}(\vec{v}),\xi''\vec{u}(\vec{v})\rangle + \sum_{s\in\mathscr{S}}\lambda_s \log u_s(\vec{v})\right)\right)\\
&= \Re\left(M(\vec{u})^{-1}\nabla_{\vec{u}\in\overline{\mathbb{H}}^r}\left(-\frac{1}{2}\langle\vec{u},\xi''\vec{u}\rangle + \sum_{s\in\mathscr{S}}\lambda_s \log u_s\right)\right)\\
&= \Re\left(M(\vec{u})^{-1}\left((\lambda_s/u_s)_{s\in\mathscr{S}} - \xi''\vec{u}\right)\right)\\
&= \Re(\vec{u})\\
&= -\nabla_{\vec{v}\in\mathbb{R}^r}\overline{\Psi}(\vec{v}).
\end{aligned}
$$

The gradient subscripts indicate whether we consider the expression as a gradient of a smooth function defined on $\mathbb{R}^r$ or of a holomorphic function on $\overline{\mathbb{H}}^r$.

It follows that $G(\vec{v})$ is locally constant on $\mathbb{R}^r\backslash\mathcal{T}$. Moreover Proposition 4.2.15 implies that G is continuous. Finally, Proposition 4.2.10(b) implies that each line $\vec{v}(t) = \vec{v}(0)+t\vec{\lambda}$ intersects $\mathcal{T}$ at only finitely many points. Combining the above implies $G(\vec{v}) = 0$ for all $\vec{v} \in \mathbb{R}^r$ as desired. $\square$

Part II

Algorithms

Chapter 5

Algorithmic threshold for multi-species spherical spin glasses

Abstract – We study efficient optimization of the Hamiltonians of multi-species spherical spin glasses. Our results characterize the maximum value attained by algorithms that are suitably Lipschitz with respect to the disorder through a variational principle that we study in detail. We rely on the branching overlap gap property introduced in our previous work [HS25] and develop a new method to establish it that does not require the interpolation method. Consequently our results apply even for models with non-convex covariance, where the Parisi formula for the true ground state remains open. As a special case, we obtain the algorithmic threshold for all single-species spherical spin glasses, which was previously known only for even models. We also obtain closed-form formulas for pure models which coincide with the E_∞ value previously determined by the Kac-Rice formula.

5.1 Introduction

This paper studies the efficient optimization of a family of random functions H_N which are high-dimensional and extremely non-convex. The computational complexity of such random optimization problems remains poorly understood in the majority of cases as most impossibility results concern worst-case rather than average-case behavior.

We focus on a general class of such problems: the Hamiltonians of multi-species spherical spin glasses. Mean-field spin glasses have been studied since [SK75] as models for disordered magnetic systems and are also closely linked to random combinatorial optimization problems [KMRT$^+$07, DMS17, Pan18]. Simply put, their Hamiltonians are certain polynomials in many variables with independent centered Gaussian coefficients.

Multi-species spin glasses such as the bipartite SK model [KC75, KS85, FKS87a, FKS87b] open the door to yet richer behavior and as discussed below remain poorly understood from a rigorous viewpoint. Our main result gives, for all multi-species spherical spin glasses, an exact *algorithmic threshold* ALG for the maximum Hamiltonian value obtained by a natural class of *stable* optimization algorithms.

For the more well-known single-species spin glasses, the celebrated Parisi formula [Par79, Tal06b, Tal06a, AC17] gives the limiting maximum value of H_N as a certain variational formula. In previous work [HS25] we obtained the algorithmic threshold for these models restricted to have only even degree interactions, given by the same variational formula over an extended state space. The central idea was to show H_N obeys a branching version of the *overlap gap property* (OGP): the absence of a certain geometric configuration of high-energy inputs [GS17a, Gam21]. The proofs of the Parisi formula [Tal06b, Tal06a, AC17], the branching OGP in [HS25], and other results (e.g. [GT02, BGT10]) require the so-called interpolation method, which is known to fail when the model's covariance is not convex. Due to this limitation of the interpolation method, the proof of our previous result does not generalize to single-species spin glasses with odd interactions, nor to multi-species spin glasses. For the same reason, the Parisi formula for the ground state of a multi-species spin glass is known only in restricted cases [BCMT15, Pan15, BL20, Sub23b, BS22b].

We develop a new method to establish the branching OGP which does not use the interpolation method. Instead, we recursively apply a uniform concentration idea introduced in [Sub24]. Consequently we are able to determine ALG for all multi-species spherical spin glasses, including those whose ground state energy is not known. As a special case, this removes the even interactions condition from [HS25] for spherical models and is the first OGP that applies to mean-field spin glasses with odd interactions.

Our results strengthen a geometric picture put forth in [HS25, Section 1.4] that in mean-field random optimization problems, the tractability of optimization to value E coincides with the presence of densely branching ultrametric trees within the super-level set at value E. On the hardness side, such trees are precisely what the branching OGP forbids. On the algorithmic side, it will be clear from our methods (see the end of Subsection 5.1.5) that efficient algorithms can be designed to descend such trees whenever they exist, thereby achieving value E.

Our algorithmic threshold for multi-species models is expressed as the maximum of a somewhat different variational principle. We analyze our algorithmic variational principle in detail, showing that maximizers are formed by joining the solutions to a pair of differential equations, and are explicit and unique for single-species and pure models. To our surprise the maximizers are not unique in general, a behavior we term *algorithmic symmetry breaking*.

5.1.1 Problem description and the value of ALG

Fix a finite set $\mathscr{S} = \{1, \dots, r\}$. For each positive integer N, fix a deterministic partition $\{1, \dots, N\} = \sqcup_{s\in\mathscr{S}} \mathcal{I}_s$ with $\lim_{N\to\infty} |\mathcal{I}_s|/N = \lambda_s$ where $\vec{\lambda} = (\lambda_1, \dots, \lambda_r) \in \mathbb{R}^{\mathscr{S}}_{>0}$ sum to 1. For $s \in \mathscr{S}$ and $\boldsymbol{x} \in \mathbb{R}^N$, let $\boldsymbol{x}_s \in \mathbb{R}^{\mathcal{I}_s}$ denote the restriction of $\boldsymbol{x}$ to coordinates $\mathcal{I}_s$. We consider the state space

$$\mathcal{B}_N = \left\{\boldsymbol{x} \in \mathbb{R}^N : \|\boldsymbol{x}_s\|_2^2 \le \lambda_s N \ \forall s \in \mathscr{S}\right\}.$$

Fix $\check{h} = (h_1, \dots, h_r) \in \mathbb{R}^{\mathscr{S}}_{\ge 0}$ and let $\mathbf{1} = (1, \dots, 1) \in \mathbb{R}^N$. For each $k \ge 2$ fix a symmetric tensor $\Gamma^{(k)} = (\gamma_{s_1,\dots,s_k})_{s_1,\dots,s_k\in\mathscr{S}} \in (\mathbb{R}^{\mathscr{S}}_{\ge 0})^{\otimes k}$ with $\sum_{k\ge 2} 2^k \big\|\Gamma^{(k)}\big\|_\infty < \infty$, and let $\boldsymbol{G}^{(k)} \in (\mathbb{R}^N)^{\otimes k}$ be a tensor with i.i.d. standard Gaussian entries. For $A \in (\mathbb{R}^{\mathscr{S}})^{\otimes k}$, $B \in (\mathbb{R}^N)^{\otimes k}$, define $A \diamond B \in (\mathbb{R}^N)^{\otimes k}$ to be the tensor with entries

$$(A \diamond B)_{i_1,\dots,i_k} = A_{s(i_1),\dots,s(i_k)} B_{i_1,\dots,i_k}, \tag{5.1}$$

where $s(i)$ denotes the $s \in \mathscr{S}$ such that $i \in \mathcal{I}_s$. Let $\boldsymbol{h} = \check{h} \diamond \mathbf{1}$. We consider the mean-field multi-species spin glass Hamiltonian

$$H_N(\boldsymbol{\sigma}) = \langle \boldsymbol{h}, \boldsymbol{\sigma}\rangle + \widetilde{H}_N(\boldsymbol{\sigma}), \quad \text{where} \tag{5.2}$$

$$\begin{aligned}\widetilde{H}_N(\boldsymbol{\sigma}) &= \sum_{k\ge 2} \frac{1}{N^{(k-1)/2}} \langle \Gamma^{(k)} \diamond \boldsymbol{G}^{(k)}, \boldsymbol{\sigma}^{\otimes k}\rangle \qquad (5.3)\\ &= \sum_{k\ge 2} \frac{1}{N^{(k-1)/2}} \sum_{i_1,\dots,i_k=1}^{N} \gamma_{s(i_1),\dots,s(i_k)} \boldsymbol{G}^{(k)}_{i_1,\dots,i_k} \sigma_{i_1}\cdots\sigma_{i_k}\end{aligned}$$

with inputs $\boldsymbol{\sigma} = (\sigma_1, \dots, \sigma_N) \in \mathcal{B}_N$. For $\boldsymbol{\sigma}, \boldsymbol{\rho} \in \mathcal{B}_N$, define the species s overlap and overlap vector

$$R_s(\boldsymbol{\sigma}, \boldsymbol{\rho}) = \frac{\langle \boldsymbol{\sigma}_s, \boldsymbol{\rho}_s\rangle}{\lambda_s N}, \qquad \vec{R}(\boldsymbol{\sigma}, \boldsymbol{\rho}) = (R_1(\boldsymbol{\sigma}, \boldsymbol{\rho}), \dots, R_r(\boldsymbol{\sigma}, \boldsymbol{\rho})). \tag{5.4}$$

Let $\odot$ denote coordinate-wise product. For $\vec{x} = (x_1, \dots, x_r) \in \mathbb{R}^{\mathscr{S}}$, let

$$\begin{aligned}\xi(\vec{x}) &= \sum_{k\ge 2} \langle \Gamma^{(k)} \odot \Gamma^{(k)}, (\vec{\lambda} \odot \vec{x})^{\otimes k}\rangle\\ &= \sum_{k\ge 2} \sum_{s_1\dots,s_k\in\mathscr{S}} \gamma^2_{s_1,\dots,s_k} (\lambda_{s_1} x_{s_1}) \cdots (\lambda_{s_k} x_{s_k}).\end{aligned}$$

The random function $\widetilde{H}_N$ can also be described as the Gaussian process on $\mathcal{B}_N$ with covariance

$$\mathbb{E}\widetilde{H}_N(\boldsymbol{\sigma})\widetilde{H}_N(\boldsymbol{\rho}) = N\xi(\vec{R}(\boldsymbol{\sigma}, \boldsymbol{\rho})).$$

It will be useful to define, for $s \in \mathscr{S}$,

$$\xi^s(\vec{x}) = \lambda_s^{-1} \partial_{x_s} \xi(\vec{x}).$$

Our main result is a characterization of the largest energy attainable by algorithms with $O(1)$-Lipschitz dependence on the disorder coefficients. To define this class of algorithms, we consider the following distance on the space $\mathscr{H}_N$ of Hamiltonians H_N. We identify H_N with its disorder coefficients $(\boldsymbol{G}^{(k)})_{k\geq 2}$, which we concatenate in an arbitrary but fixed order into an infinite vector $\boldsymbol{g}(H_N)$. We equip $\mathscr{H}_N$ with the (possibly infinite) distance

$$\|H_N - H_N'\|_2 = \|\boldsymbol{g}(H_N) - \boldsymbol{g}(H_N')\|_2.$$

In other words we identify H with functions of the form (5.2) and measure distance using the Euclidean norm on the infinite sequence of coefficients, treating $\boldsymbol{h}$ as constant.[1]

We equip $\mathcal{B}_N$ with the ℓ_2 distance. For each $\tau > 0$, these distances define a class of τ-Lipschitz functions $\mathcal{A}_N : \mathscr{H}_N \to \mathcal{B}_N$, satisfying

$$\|\mathcal{A}_N(H_N) - \mathcal{A}_N(H_N')\|_2 \leq \tau \|H_N - H_N'\|_2, \quad \forall\ H_N, H_N' \in \mathscr{H}_N.$$

Note that this inequality holds vacuously for pairs (H_N, H_N') where the latter distance is infinite. As explained in [HS25, Section 8], the class of $O(1)$-Lipschitz algorithms includes gradient descent and Langevin dynamics for the Gibbs measure $e^{\beta H_N(\boldsymbol{\sigma})}\mathsf{d}\boldsymbol{\sigma}$ (with suitable reflecting boundary conditions) run on constant time scales.[2] The behavior of such dynamics has been a major focus of study in its own right, see e.g. [SZ81, CK94, BG95, BG97b, BDG01, BDG06, BGJ20, DS20, DG21, DLZ21, CCM21, Sel24b].

We will characterize the largest energy attainable by a τ-Lipschitz algorithm, where τ is an arbitrarily large constant independent of N, in terms of the following variational principle. For $0 \leq q_0 \leq q_1 \leq 1$, let $\mathbb{I}(q_0, q_1)$ be the set of non-decreasing, continuously differentiable functions $f : [q_0, q_1] \to [0, 1]$. Let $\mathsf{Adm}(q_0, q_1) \subset \mathbb{I}(q_0, q_1)^{\mathscr{S}}$ be the set of coordinate-wise non-decreasing, continuously differentiable functions $\Phi : [q_0, q_1] \to [0, 1]^{\mathscr{S}}$ which satisfy, for all $q \in [q_0, q_1]$,

$$\langle \vec{\lambda}, \Phi(q) \rangle = q. \tag{5.5}$$

We say Φ is *admissible* if it satisfies (5.5). For $p \in \mathbb{I}(q_0, 1)$, $\Phi \in \mathsf{Adm}(q_0, 1)$, define the algorithmic functional

$$\mathbb{A}(p, \Phi; q_0) \equiv \sum_{s \in \mathscr{S}} \lambda_s \left[h_s \sqrt{\Phi_s(q_0)} + \int_{q_0}^1 \sqrt{\Phi_s'(q)(p \times \xi^s \circ \Phi)'(q)}\ \mathsf{d}q \right] \tag{5.6}$$

where $(p \times \xi^s \circ \Phi)(q) = p(q)\xi^s(\Phi(q))$. (See the end of this subsection for an interpretation of this formula.) We can now state the algorithmic threshold for multi-species spherical spin glasses:

$$\mathsf{ALG} \equiv \sup_{q_0 \in [0,1]} \sup_{\substack{p \in \mathbb{I}(q_0,1) \\ \Phi \in \mathsf{Adm}(q_0,1)}} \mathbb{A}(p, \Phi; q_0). \tag{5.7}$$

The following theorem is our main result. Together with Theorem 5.1.2 in our companion work [HS24], we find that ALG is the largest energy attained by an $O(1)$-Lipschitz algorithm. Here and throughout, all implicit constants may depend also on $(\xi, \check{h}, \vec{\lambda})$.

Theorem 5.1.1. *Let $\tau, \varepsilon > 0$ be constants. For $N \geq N_0$ sufficiently large, any τ-Lipschitz $\mathcal{A}_N : \mathscr{H}_N \to \mathcal{B}_N$ satisfies*

$$\mathbb{P}[H_N(\mathcal{A}_N(H_N))/N \geq \mathsf{ALG} + \varepsilon] \leq \exp(-cN), \qquad c = c(\varepsilon, \tau) > 0.$$

Theorem 5.1.2 ([HS24, Theorem 1])**.** *For any $\varepsilon > 0$, there exists an efficient and $O_\varepsilon(1)$-Lipschitz algorithm $\mathcal{A}_N : \mathscr{H}_N \to \mathcal{B}_N$ such that*

$$\mathbb{P}[H_N(\mathcal{A}_N(H_N))/N \geq \mathsf{ALG} - \varepsilon] \geq 1 - \exp(-cN), \quad c = c(\varepsilon) > 0.$$

[1] Technically, under this definition multiple "Hamiltonians" may correspond to the same function $\mathcal{B}_N \to \mathbb{R}$ due to redundancy between coefficients $G^{(k)}_{i_1,\ldots,i_k}$ and $G^{(k)}_{i_{\pi(1)},\ldots,i_{\pi(k)}}$ for permutations π. However this does not cause any issues for us. If one prefers that $\mathcal{A}(\cdot)$ depend only on the underlying function, it suffices to average the coefficients over permutations π before applying $\mathcal{A}$.

[2] Up to modification on a set of probability at most e^{-cN}, which suffices just as well for our purposes.

Our proof of Theorem 5.1.2 in [HS24] uses approximate message passing (AMP), a general family of gradient-based algorithms, following a recent line of work [Sub21a, Mon21, AMS21, AS22, Sel24a].

In fact, in Theorem 5.1.1 we will not require the full Lipschitz assumption on $\mathcal{A}_N$. Theorem 5.1.1 holds for all algorithms satisfying an *overlap concentration* property (see Definition 5.2.2, Theorem 5.2.3), that for any fixed correlation $p \in [0,1]$ between the disorder coefficients of H_N^1 and H_N^2, the overlap vector $\vec{R}(\mathcal{A}_N(H_N^1), \mathcal{A}_N(H_N^2))$ concentrates tightly around its mean. This property holds automatically for $O(1)$-Lipschitz $\mathcal{A}_N$ due to Gaussian concentration of measure.

Interpretation of the algorithmic functional $\mathbb{A}$ Suppose first that $\check{h} = \vec{0}$. We will see (Theorem 5.1.12) that ALG is maximized at $q_0 = 0$, $p \equiv 1$, in which case

$$\mathbb{A}(p, \Phi; q_0) = \sum_{s \in \mathscr{S}} \lambda_s \int_0^1 \sqrt{\Phi_s'(q)(\xi^s \circ \Phi)'(q)}\, \mathsf{d}q. \tag{5.8}$$

In a single-species spherical spin glass, we have $\lambda_1 = 1$ and $\Phi(q) = q$, so (5.8) reduces to the formula $\mathsf{ALG} = \int_0^1 \xi''(q)^{1/2}\, \mathsf{d}q$ derived in [HS25]. This energy is attained by the algorithm of Subag [Sub21a], which starts from the origin and explores to the surface of the sphere by small orthogonal steps in the direction of the largest eigenvector of the local tangential Hessian.

In multi-species models, (5.8) is the energy attained by a generalization of Subag's algorithm, which is essentially shown in Proposition 5.3.3. Instead of computing a maximal eigenvector at each step, given the current iterate $\boldsymbol{x}^t$ this algorithm chooses $\boldsymbol{x}^{t+1}$ to maximize $\langle \nabla^2 H_N(\boldsymbol{x}^t), (\boldsymbol{x}^{t+1} - \boldsymbol{x}^t)^{\otimes 2}\rangle$ on a product of r small spheres centered at $\boldsymbol{x}^t$. This algorithm may advance through different species at different speeds by tuning the radii of the spheres at each step, and the function Φ is a "radius schedule" whose image specifies the path of depths $(\|\boldsymbol{x}_s^t\|_2^2/\lambda_s N)_{s \in \mathscr{S}}$ traced by the iterates $\boldsymbol{x}^t$. Thus each $\Phi \in \mathsf{Adm}(0,1)$ corresponds to an algorithm, and Theorem 5.1.1 essentially states that the algorithmic threshold is the energy attained by the multi-species Subag algorithm with the best Φ.

The function p arises from a further generalization of this algorithm, which becomes necessary in the presence of external field $\check{h} \neq \vec{0}$. The idea is to reveal the disorder coefficients of H_N gradually (in the sense of progressively less noisy Gaussian observations, see (5.29)) and in tandem with the iterates $\boldsymbol{x}^t$. Though counterintuitive, this allows the algorithm to take advantage of the gradients of the newly revealed part of H_N at each step. The iterate $\boldsymbol{x}^{t+1}$ is now chosen to maximize the sum

$$\langle \nabla(H_N^{t+1} - H_N^t)(\boldsymbol{x}^t), \boldsymbol{x}^{t+1} - \boldsymbol{x}^t\rangle + \frac{1}{2}\langle \nabla^2 H_N^t(\boldsymbol{x}^t), (\boldsymbol{x}^{t+1} - \boldsymbol{x}^t)^{\otimes 2}\rangle \tag{5.9}$$

of a gradient contribution from the new component and a Hessian contribution from the previously revealed components. The function p is an "information schedule" that determines the rate at which entries of H_N are revealed. Moreover, to take advantage of the external field, the algorithm starts from a point $\boldsymbol{x}^0$ correlated with $\boldsymbol{h}$ whose norm is $q_0\sqrt{N}$; the first term in (5.6) is exactly the value $\langle \boldsymbol{h}, \boldsymbol{x}^0\rangle/N$ (see (5.47)).

Technically it is not obvious whether these generalized Subag algorithms can be directly made suitably Lipschitz. This is one reason we prove Theorem 5.1.2 using AMP in [HS24].

5.1.2 Description of maximizers to the algorithmic variational problem

In this subsection we describe the detailed properties of the maximizers (p, Φ, q_0) of (5.7), culminating in an explicit description in Theorem 5.1.12 as a piecewise combination of solutions to two ordinary differential equations.

For intuition, it may help to recall the famous ansatz that spin glass Gibbs measures are asymptotically ultrametric, corresponding to orthogonally branching trees in $\mathbb{R}^N$ (see e.g. [MV85, Pan13a, Jag17, CS21]). When $\check{h} = \vec{0}$, the associated tree is rooted at the origin; otherwise the root's location is correlated with $\boldsymbol{h}$ but random. Theorem 5.1.12 below shows that maximizers of $\mathbb{A}$ consist of a "root-finding" component and a "tree-descending" component; the corresponding algorithms first locate an analogous root, and then descend an algorithmic analog of a low-temperature ultrametric tree.

This description holds under the following generic assumption.

Assumption 5.1.3. All quadratic and cubic interactions participate in H, i.e. $\Gamma^{(2)}, \Gamma^{(3)} > 0$ coordinate-wise. We will call such models **non-degenerate**.

Note that ALG is continuous in the parameters $\xi, \check{h}$ (for a simple proof, first observe that $\mathbb{A}$ and hence ALG are monotone and subadditive in $(\xi, \check{h})$). Since Assumption 5.1.3 is a dense condition, to determine the value of ALG it suffices to do so under this assumption. Under this assumption, we will describe in detail the maximizing triples $(p, \Phi; q_0)$, which always exist but need not be unique. Non-degeneracy removes extraneous symmetries among the maximizers of $\mathbb{A}$ when e.g. ξ is a sum of polynomials in disjoint sets of variables.

Definition 5.1.4. A symmetric matrix $M \in \mathbb{R}^{\mathscr{S}\times\mathscr{S}}$ is **diagonally signed** if $M_{i,i} \geq 0$ and $M_{i,j} < 0$ for all $i \neq j$.

Definition 5.1.5. A diagonally signed matrix M is **super-solvable** if it is positive semidefinite, and **solvable** if it is furthermore singular; otherwise M is **strictly sub-solvable**. A point $\vec{x} \in (0,1]^{\mathscr{S}}$ is super-solvable, solvable, or strictly sub-solvable if $M^*_{\mathsf{sym}}(\vec{x})$ is, where

$$M^*_{\mathsf{sym}}(\vec{x}) = \operatorname{diag}\left(\left(\frac{\partial_{x_s}\xi(\vec{x}) + \lambda_s h_s^2}{x_s}\right)_{s\in\mathscr{S}}\right) - \left(\partial_{x_s,x_{s'}}\xi(\vec{x})\right)_{s,s'\in\mathscr{S}}. \tag{5.10}$$

We also adopt the convention that $\vec{0}$ is always super-solvable, and solvable if $\check{h} = \vec{0}$.

Solvability plays a central role in our description below of optimal solutions to the variational problem. In brief, p reaches 1 exactly at the point where Φ switches from being sub-solvable to super-solvable. Further, in the sub-solvable region (p, Φ) obeys a first order *root-finding* ODE, while in the super-solvable region $p = 1$ and Φ obeys a second order *tree-descending* ODE. If $\vec{0}$ is solvable then there is no root-finding phase, while if $\vec{0}$ is strictly sub-solvable then there is no tree-descending phase.

Remark 5.1.6. It is possible to extend the notions of (super, strict sub)-solvability to all of $[0,1]^{\mathscr{S}}$ by using the alternative characterization from Corollary 5.4.5. However this will not be necessary, as our results only use these notions for $\vec{x} \in (0,1]^{\mathscr{S}} \cup \{\vec{0}\}$.

Definition 5.1.7. Suppose $\vec{x} \in (0,1]^{\mathscr{S}}$ is super-solvable with $\langle \vec{\lambda}, \vec{x}\rangle = q_1$. A **root-finding trajectory** with endpoint $\vec{x}$ is a pair $(p, \Phi) \in \mathbb{I}(q_0, q_1) \times \mathsf{Adm}(q_0, q_1)$, for some $q_0 \in [0, q_1]$, satisfying $p(q_1) = 1$, $\Phi(q_1) = \vec{x}$, $p(q_0) = 0$, and for all $q \in [q_0, q_1]$:

$$\frac{(p \times \xi^s \circ \Phi)'(q)}{\Phi_s'(q)} = L_s \equiv \frac{\xi^s(\vec{x}) + h_s^2}{x_s}, \quad \forall s \in \mathscr{S}. \tag{5.11}$$

Assuming for now that $p, \Phi_s \in C^1([q_0, 1])$, (5.11) together with admissibility can be written for each $q \in [q_0, q_1]$ as the ordinary differential equation

$$p'(q)\xi^s(\Phi(q)) + p(q)\sum_{s'\in\mathscr{S}} \partial_{x_{s'}}\xi^s(\Phi(q))\,\Phi_{s'}'(q) = L_s\Phi_s'(q), \quad \forall s \in \mathscr{S}; \tag{5.12}$$

$$\sum_{s\in\mathscr{S}} \lambda_s \Phi_s'(q) = 1; \tag{5.13}$$

$$p'(q), \Phi_s'(q) \geq 0. \tag{5.14}$$

Here $\vec{L}$ is treated as fixed, as it is determined by the boundary condition at q_1. We note that the value q itself does not explicitly appear in equation (5.12), except that $\Phi(q)$ determines q via admissibility; thus admissibility functions as a choice of time-parametrization. In fact (5.12) is equivalent to a well-posed ordinary differential equation (away from $\vec{0}$, which it never reaches by Proposition 5.1.9). Moreover as shown in Proposition 5.1.9(a), solving this ODE from a super-solvable initial condition always yields a valid root-finding trajectory (e.g. the resulting p is actually non-decreasing on $[q_0, q_1]$).

Proposition 5.1.8. *For $(p(q), \Phi(q))$ in compact subsets of $[0,1] \times (0,1]^{\mathscr{S}}$ the equation (5.12) has a unique solution $(p'(q), \Phi'(q))$ which is Lipschitz in $(p(q), \Phi(q))$ (where the Lipschitz constant may depend on the compact set).*

Proposition 5.1.9. *$\check{h} \neq \vec{0}$ if and only if there exists a super-solvable $\vec{x} \in (0,1]^{\mathscr{S}}$. If this holds, for each such $\vec{x}$:*

(a) Let $q_1 = \langle \vec{\lambda}, \vec{x} \rangle > 0$. There is a unique root-finding trajectory (p, Φ) with endpoint $\vec{x}$. It is obtained by solving (5.12) backward in time from initial condition $p(q_1) = 1$, $\Phi(q_1) = \vec{x}$ until reaching $p(q_0) = 0$. Moreover the resulting p is non-decreasing and concave on $[q_0, q_1]$.

(b) $q_0 > 0$, and in fact $\Phi_s(q_0) > 0$ if and only if $h_s > 0$.

Definition 5.1.10. Suppose $\vec{x} \in (0,1]^{\mathscr{S}} \cup \{\vec{0}\}$ is solvable with $\langle \vec{\lambda}, \vec{x} \rangle = q_1$. A **tree-descending trajectory** with endpoint $\vec{x}$ is a pair $(p, \Phi) \in \mathbb{I}(q_1, q_2) \times \mathsf{Adm}(q_1, q_2)$ satisfying $p \equiv 1$, $\Phi(q_1) = \vec{x}$, $M^*_{\mathsf{sym}}(\vec{x})\Phi'(q_1) = \vec{0}$, $\|\Phi_s(q_2)\|_\infty = 1$ and

$$\frac{1}{\Phi'_s(q)} \frac{\mathrm{d}}{\mathrm{d}q} \sqrt{\frac{\Phi'_s(q)}{(\xi^s \circ \Phi)'(q)}} = \frac{1}{\Phi'_{s'}(q)} \frac{\mathrm{d}}{\mathrm{d}q} \sqrt{\frac{\Phi'_{s'}(q)}{(\xi^{s'} \circ \Phi)'(q)}} \tag{5.15}$$

for all $s, s' \in \mathscr{S}$ and $q \in [q_1, q_2]$. Moreover, (p, Φ) is *targeted* if $\Phi(1) = \vec{1}$ (i.e. $q_2 = 1$).

Similarly to (5.12), assuming Φ'' is defined, (5.15) together with the admissibility constraint

$$\sum_{s \in \mathscr{S}} \lambda_s \Phi''_s(q) = 0 \tag{5.16}$$

is equivalent to a second order differential equation. We show in Subsection 5.4.6 and Appendix 5.C.3 that this equation is suitably well-posed and obtain the following results.

Proposition 5.1.11. *Suppose Assumption 5.1.3 holds and $\vec{x} \in (0,1]^{\mathscr{S}} \cup \{\vec{0}\}$ is solvable with $\langle \vec{\lambda}, \vec{x} \rangle = q_1$.*

(a) If $\check{h} \neq \vec{0}$, then $\vec{x} \in (0,1]^{\mathscr{S}}$ and $q_1 > 0$. There is a unique $\vec{v} \in \mathbb{R}^{\mathscr{S}}_{\geq 0}$ satisfying

$$M^*_{\mathsf{sym}}(\vec{x})\vec{v} = \vec{0}, \tag{5.17}$$

$$\langle \vec{\lambda}, \vec{v} \rangle = 1. \tag{5.18}$$

There is a unique tree-descending trajectory with endpoint $\vec{x}$, which is obtained by solving (5.15) forward in time from $\Phi(q_1) = \vec{x}$, $\Phi'(q_1) = \vec{v}$ until reaching $\|\Phi_s(q_2)\|_\infty = 1$.

(b) If $\check{h} = \vec{0}$, then $\vec{x} = \vec{0}$ and $q_1 = 0$. For any $\vec{v} \in \mathbb{R}^{\mathscr{S}}_{\geq 0}$ satisfying (5.18), there is a unique tree-descending trajectory with $\Phi(0) = \vec{0}$ and $\Phi'(0) = \vec{v}$, which is obtained by solving (5.15) forward in time from these conditions until reaching $\|\Phi_s(q_2)\|_\infty = 1$.

The following theorem is our main result describing maximizers of (5.7).

Theorem 5.1.12. *Suppose Assumption 5.1.3 holds. Then a maximizer (p, Φ, q_0) of (5.7) exists, and all maximizers are continuously differentiable on $[q_0, 1]$. There exists $q_1 \in [q_0, 1]$ such that $\Phi(q_1) \in (0,1]^{\mathscr{S}} \cup \{\vec{0}\}$ and furthermore (p, Φ) is the root-finding trajectory with endpoint $\Phi(q_1)$ on $[q_0, q_1]$ and a (targeted) tree-descending trajectory with endpoint $\Phi(q_1)$ on $[q_1, 1]$. ALG is given by*

$$\mathsf{ALG} = \mathbb{A}(p, \Phi; q_0) = \sum_{s \in \mathscr{S}} \lambda_s \left[\sqrt{\Phi_s(q_1)(\xi^s(\Phi(q_1)) + h_s^2)} + \int_{q_1}^1 \sqrt{\Phi'_s(q)(\xi^s \circ \Phi)'(q)}\, \mathsf{d}q \right]. \tag{5.19}$$

Finally the value of q_1 is described as follows:

(a) If $\vec{1}$ is super-solvable then $q_1 = 1$, i.e. (p, Φ) is the root-finding trajectory with endpoint $\vec{1}$.

(b) If $\vec{1}$ is sub-solvable and $\check{h} \neq \vec{0}$, then $q_1 \in (q_0, 1)$, i.e. (p, Φ) contains both root-finding and tree-descending trajectories.

(c) If $\check{h} = \vec{0}$, then $\vec{1}$ is sub-solvable and $q_1 = 0$, i.e. (p, Φ) is a (targeted) tree-descending trajectory with endpoint $\vec{0}$.

Note that in case (b), $\Phi(q_1) \in (0,1]^{\mathscr{S}}$ if $h_s > 0$ for **any** s. Examples of each of these cases are given in Figure 5.1.1.

Remark 5.1.13. The choice of state space $\mathcal{B}_N$ is a natural though arbitrary normalization. For any $\vec{a} \in \mathbb{R}_{>0}^{\mathscr{S}}$, we could just as well consider the state space

$$\mathcal{B}_N(\vec{a}) = \left\{ \boldsymbol{x} \in \mathbb{R}^N : \|\boldsymbol{x}_s\|_2^2 \le a_s \lambda_s N \ \forall s \in \mathscr{S} \right\}. \tag{5.20}$$

Clearly optimizing the model described by $\xi, \check{h}$ over this space is equivalent to optimizing the model described by[3]

$$\tilde{\xi}(\vec{x}) = \xi(\vec{x} \odot \sqrt{\vec{a}}), \qquad \tilde{\check{h}} = \check{h} \odot \sqrt{\vec{a}} \tag{5.21}$$

over $\mathcal{B}_N$, so changing the problem in this way does not add any complexity. However, from this point of view we can see that the requirement $\Phi(1) = \vec{1}$ in Equation (5.7) and Theorem 5.1.12 is merely a consequence of the normalization. If we wished to optimize over $\mathcal{B}_N(\vec{a})$, equation (5.7) and Theorem 5.1.12 still hold with the right endpoint of Φ changed to $\vec{a}$, which is easily proved by the transformation (5.21). Thus the non-targeted trajectories in Figure 5.1.1 describe optimal algorithms for other state spaces $\mathcal{B}_N(\vec{a})$.

Remark 5.1.14. Because the root-finding and tree-descending ODEs are well-posed, the results above give a natural approach to solve the N-independent problem of approximately maximizing $\mathbb{A}$ to ε error. If $\vec{1}$ is super-solvable then ALG is given directly by (5.19). If $\check{h} \neq \vec{0}$, then it suffices to brute-force search for the value $\Phi(q_1)$ over a δ-net of solvable $\vec{x} \in [0,1]^{\mathscr{S}}$ and solve each of the two ODEs above; note that the vector $\Phi'(q_1)$ is determined by (5.12). Finally if $\check{h} = \vec{0}$, since $q_1 = 0$ it suffices to brute-force search over all $\Phi'(0)$ satisfying (5.13).

Remark 5.1.15. In models where $\vec{1}$ is super-solvable, the formula (5.19) simplifies to

$$\mathsf{ALG} = \sum_{s \in \mathscr{S}} \lambda_s \sqrt{\Phi_s(1)(\xi^s(\Phi(1)) + h_s^2)}. \tag{5.22}$$

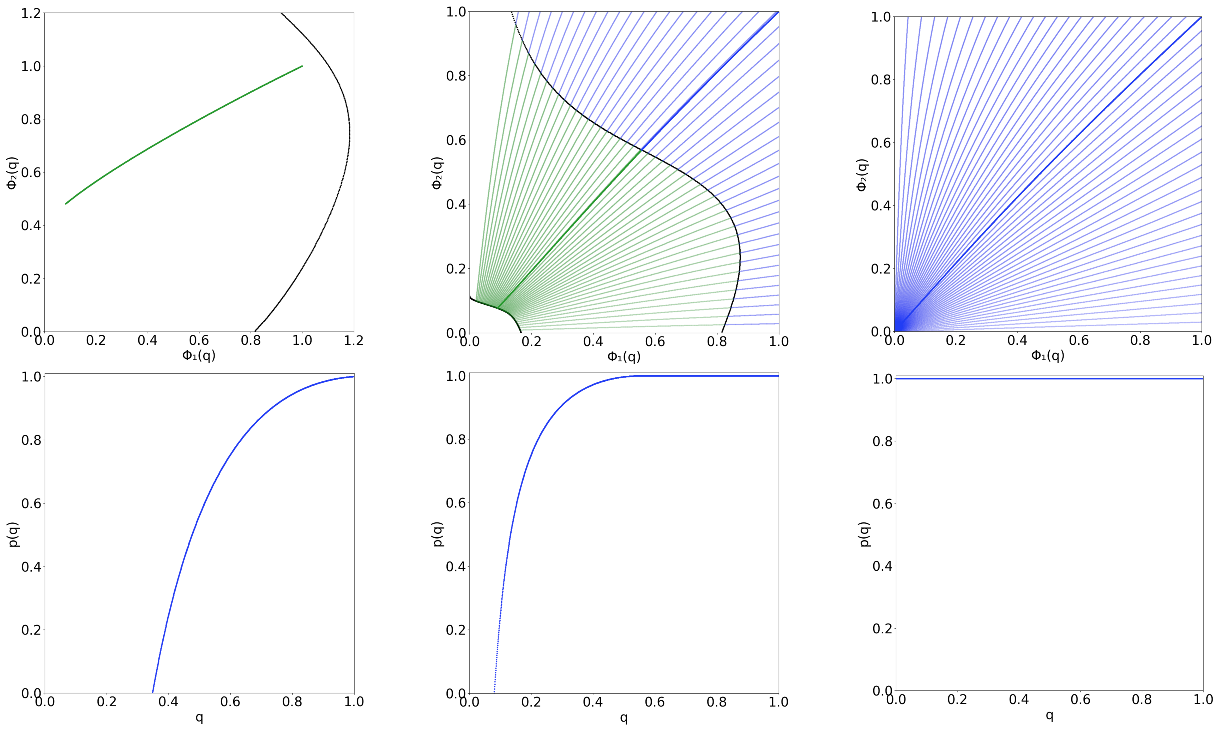
As shown in our companion paper [HS23c, Theorem 1.6], this coincides with the true maximum value OPT. Moreover the models where $\vec{1}$ is strictly super-solvable are precisely the *topologically trivial* ones, where with high probability the number of critical points is exactly 2^r, the minimum number possible for a Morse function on a product of r spheres. This generalizes an observation from [HS25] that in an analogous regime of single-species models, $\mathsf{ALG} = \mathsf{OPT}$ and, as shown in [Fyo15, BČNS22], the model is topologically trivial.

Remark 5.1.16. Recall the algorithmic interpretation of (p, Φ, q_0) discussed around (5.9). For any $q \in [q_0, q_1]$, the iterate of this algorithm at radii $\Phi(q)$ is an approximate maximizer of the Hamiltonian revealed up to that point (whose disorder coefficients have variance $p(q)$) on the product of spheres $\mathcal{B}_N(\Phi(q))$. Indeed the energy attained by these iterates is calculated in Corollary 5.4.29 and coincides with (5.22) with $\Phi(q), p(q)\xi$ in place of $\Phi(1), \xi$.

5.1.3 Explicit solutions in special cases

While the formulas (5.7), (5.19) for ALG involve the solution to a variational problem, ALG can be written explicitly in the important special cases of single-species models where $r = 1$ and $\vec{\lambda} = (1)$, and pure models where ξ is a monomial.

[3] Here and throughout this paper, powers of vectors such as $\sqrt{\vec{a}}$ are taken coordinate-wise.



(a) $\check{h} = (0.4, 1.4)$, $\vec{1}$ super-solvable. (b) $\check{h} = (0.4, 0.4)$, $\vec{1}$ sub-solvable. (c) $\check{h} = (0, 0)$.

Figure 5.1.1: Examples of Theorem 5.1.12. Consider the model $\vec{\lambda} = (\frac{1}{3}, \frac{2}{3})$, $\xi(x_1, x_2) = \nu(\lambda_1 x_1, \lambda_2 x_2)$, and various $\check{h}$ specified in the captions above, where $\nu(x_1, x_2) = x_1^2 + x_1x_2 + x_2^2 + x_1^4 + x_1x_2^3$. These are described by parts (a), (b), and (c) of Theorem 5.1.12, respectively. The top diagrams plot $\Phi(q)$ with root-finding components green and tree-descending components blue. The optimal Φ, which passes through $(1,1)$, is bold. Figures 5.1.1b and 5.1.1c show non-targeted trajectories otherwise described by Theorem 5.1.12. In Figure 5.1.1a the black curve comprises the solvable points and $(1,1)$ is inside of this curve. In Figure 5.1.1b the outer black curve comprises the solvable points, which are the possible values of $\Phi(q_1)$, and $(1,1)$ is outside of this curve. The inner black curve of Figure 5.1.1b comprises the corresponding values of $\Phi(q_0)$. The bottom diagrams plot $(q, p(q))$ for the optimal p.

Single-species models

In single-species models, $\xi(q)$ is a univariate function and (5.5) implies $\Phi(q) = q$. Let $\check{h} = (h)$.

Corollary 5.1.17 (Algorithmic threshold of single-species models)**.** *If $\xi'(1) + h^2 \geq \xi''(1)$, then*

$$\mathsf{ALG} = (\xi'(1) + h^2)^{1/2}.$$

The variational formula (5.7) is maximized at $q_0 = \frac{h^2}{\xi'(1)+h^2}$, $p(q) = \frac{q(\xi'(1)+h^2)-h^2}{\xi'(q)}$ for $q \in [q_0, 1]$. Otherwise there is a unique $q_1 \in [0, 1)$ satisfying $\xi'(q_1) + h^2 = q_1\xi''(q_1)$, and

$$\mathsf{ALG} = q_1\xi''(q_1)^{1/2} + \int_{q_1}^{1} \xi''(q)^{1/2}\, \mathsf{d}q.$$

The variational formula (5.7) is maximized at

$$q_0 = \frac{h^2}{\xi''(q_1)}, \qquad p(q) = \begin{cases} \frac{q\xi''(q_1)-h^2}{\xi'(q)} & q \in [q_0, q_1], \\ 1 & q \in [q_1, 1]. \end{cases}$$

Except for the formulas for q_0 and $p(q)$, this corollary follows readily from Theorem 5.1.12; note that super-solvability of $\vec{1}$ generalizes the inequality $\xi'(1) + h^2 \geq \xi''(1)$ and solvability of $\Phi(q_1)$ generalizes $\xi'(q_1) + h^2 = q_1\xi''(q_1)$. The formulas for q_0 and $p(q)$ follow from (5.83) and (5.81).

The formula for ALG in Corollary 5.1.17 matches [HS25, Proposition 2.2]. Whereas [HS25] proves this formula for even ξ, we obtain it in full generality. This formula also matches the ground state energy in full replica symmetric breaking models as obtained in [CS17, Proposition 2].

Direct proof for single-species models without external field

In the case $h = 0$, the formula for ALG can be directly recovered from the variational formula (5.7). First, we should clearly take $q_0 = 0$, so $\mathsf{ALG} = \sup_{p\in\mathbb{I}(0,1)} \int_0^1 (p\xi')'(q)^{1/2}\ \mathsf{d}q$. Then, because

$$\int_t^1 (p\xi')'(q)\ \mathsf{d}q = \xi'(1) - p(t)\xi'(t) \geq \xi'(1) - \xi'(t) = \int_t^1 \xi''(q)\ \mathsf{d}q$$

for all $t \in [0,1]$ with equality at $t = 0$, the function $(p\xi')'$ majorizes ξ'' (see e.g. [Joe92] for precise definitions of majorization in non-discrete settings). Here we use that ξ'' is increasing, but do not assume that $(p\xi')'$ is. By Karamata's inequality,

$$\int_0^1 (p\xi')'(q)^{1/2}\ \mathsf{d}q \leq \int_0^1 \xi''(q)^{1/2}\ \mathsf{d}q$$

with equality at $p \equiv 1$.

Pure models

Finally we give in Theorem 5.1.18 below an explicit formula for ALG for *pure* ξ consisting of a single monomial, and moreover identify the unique maximizer to $\mathbb{A}$. Our proof in Subsection 5.4.7 takes advantage of scale invariance to relate values of ALG at different radii (see Remark 5.1.13). Recently [Sub23b] used a similar scale invariance (and other ideas) to compute the free energy in such models under the mild assumption of convergence as $N \to \infty$. Intriguingly for all pure models, the value ALG agrees with the threshold E_∞ arising from critical point asymptotics in [ABČ13] and determined in the multi-species setting by [McK24].

It should be noted that Assumption 5.1.3 on non-degeneracy is false for pure models, so we cannot rely on the structural results of Theorem 5.1.12. Additionally, note that although the optimal trajectories Φ stated in Theorem 5.1.18 are not admissible, this does not present a problem; Lemma 5.4.9 shows that admissibility is just a convenient choice of time parametrization and deviating from it does not affect the value of $\mathbb{A}$.

Theorem 5.1.18. *Suppose $\check{h} = \vec{0}$ and*

$$\xi(x_1, \dots, x_r) = \prod_{s\in\mathscr{S}} x_s^{a_s}$$

for positive integers $a_1, \dots, a_r$ with $r \geq 2$ and $\sum_{s\in\mathscr{S}} a_s \geq 3$. Define the exponents b_s by

$$b_s = \frac{1 - \sqrt{\frac{a_s}{a_s + L\lambda_s}}}{2}, \quad s \in \mathscr{S} \tag{5.23}$$

where $L = L(\vec{a}) > 0$ is the unique value such that $\sum_{s\in\mathscr{S}} a_s b_s = 1$. Then ALG and the (p, Φ, q_0) maximizing $\mathbb{A}$ are

$$\mathsf{ALG} = \sum_{s\in\mathscr{S}} \frac{\lambda_s\sqrt{La_s}}{\sqrt{a_s + L\lambda_s}},$$
$$\big(p(q)\,, \Phi(q),\, q_0\big) = \big(1,\, (q^{b_1}, \dots, q^{b_r}),\, 0\big).$$

In the case $\xi(x_1, x_2) = x_1x_2$ we have

$$\mathsf{ALG} = \sqrt{\lambda_1} + \sqrt{\lambda_2},$$
$$\big(p(q)\,, \Phi(q),\, q_0\big) = \big(1,\, (q, q),\, 0\big).$$

Moreover the optimal (p, Φ, q_0) is always unique up to reparametrization.

Theorem 5.1.18 simplifies in the special case that $\frac{a_s}{\lambda_s}$ is independent of s, i.e. $\lambda_s = \frac{a_s}{\sum_{s\in\mathscr{S}} a_s}$. In particular ALG depends only on the total degree $\sum_{s\in\mathscr{S}} a_s$. Note that the formula (5.23) gives $b_s = \frac{1}{\sum_{s\in\mathscr{S}} a_s}$, which is equivalent by reparametrization to $b_s = 1$ as stated below.

Corollary 5.1.19. *For pure models with* $\lambda_s = \frac{a_s}{\sum_{s'\in\mathscr{S}} a_{s'}}$, $\Phi(q) = (q,\dots,q)$ *uniquely maximizes* $\mathbb{A}$ *and*

$$\mathsf{ALG} = 2\sqrt{\frac{\left(\sum_{s\in\mathscr{S}} a_s\right) - 1}{\sum_{s\in\mathscr{S}} a_s}}.$$

For all pure models, the value ALG in Theorem 5.1.18 agrees with the threshold E_∞ defined as follows. We denote by ∇_{sp} the gradient on the product of spheres $\mathcal{S}_N \equiv \{\boldsymbol{x} \in \mathcal{B}_N \ : \ \vec{R}(\boldsymbol{x},\boldsymbol{x}) = \vec{1}\}$, and ∇^2_{sp} the Riemannian Hessian. Below the *index* of a square matrix denotes the number of non-negative eigenvalues.

Definition 5.1.20. For $\check{h} = \vec{0}$ and any ξ, the value E_∞ is given by $E_\infty = \lim_{k\to\infty} E_k \geq 0$. Here $E_k \geq 0$ is the minimal value such that for any $E > E_k$,

$$\lim_{N\to\infty} \frac{1}{N} \log \mathbb{E}\left[\left|\left\{\boldsymbol{\sigma} \in \mathcal{S}_N \ : \ H_N(\boldsymbol{\sigma}) \geq EN,\ \nabla_{\mathsf{sp}} H_N(\boldsymbol{\sigma}) = 0,\ \mathrm{index}(\nabla^2_{\mathsf{sp}} H_N(\boldsymbol{\sigma})) \geq k\right\}\right|\right] < 0.$$

Informally, E_∞ is the threshold above which critical points of unbounded index cease to exist in an annealed sense. For multi-species spin glasses, E_∞ is given by the somewhat complicated formula [McK24, Equation (2.7)] which involves the solution to a matrix Dyson equation, recalled in Subsection 5.4.7. This generalizes the single-species formulas in [ABČ13, BSZ20]. We note that for pure *single-species models*, [AG20, Theorem 1.4] claims (without a full proof yet) that for any $E < E_\infty$, critical points of bounded index (depending only on E) exist above energy E with high probability.

Corollary 5.1.21. *For all pure* ξ*, we have*

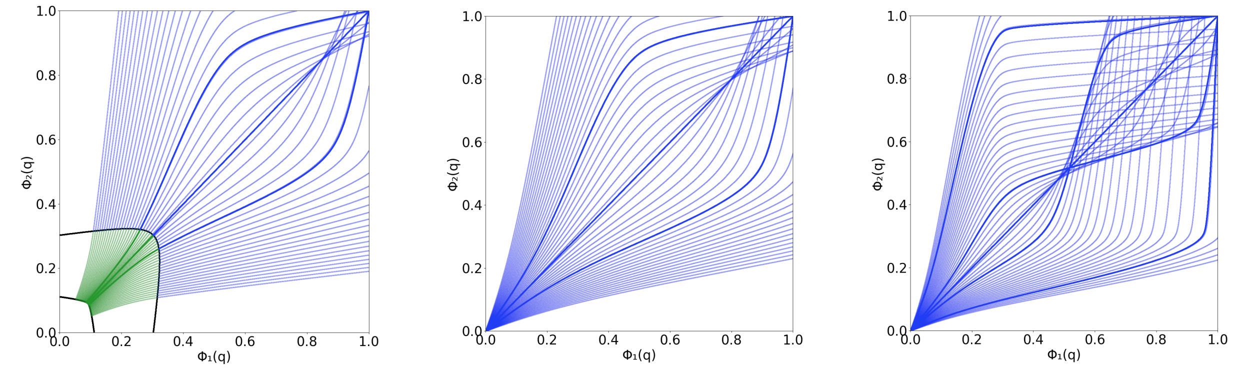
$$\mathsf{ALG} = E_\infty.$$

In the single species case, Corollary 5.1.21 holds for the pure p-spin model $\xi(x) = x^p$ with $\mathsf{ALG} = E_\infty = 2\sqrt{\frac{p-1}{p}}$ identified in [ABČ13], as discussed in [HS25, Section 2.3]. While the single-species formula is simple, Corollary 5.1.21 is much less obvious in general. In our companion works [HS24, HS23c] we give a more general approach to this connection by showing that the top of the bulk spectrum of $\nabla^2_{\mathsf{sp}} H_N(\boldsymbol{\sigma})$ is approximately 0 for $\boldsymbol{\sigma}$ the output of an explicit optimization algorithm attaining value ALG. This statement holds for all ξ and implies that ALG in general lies in an interval denoted $[E_\infty^-, E_\infty^+]$ in [AB13]. Also relatedly, [Sel24b] shows that low-temperature Langevin dynamics (run for large dimension-free time) suffices to attain energy $\mathsf{ALG} = E_\infty$ in pure models. (The result is stated for 1 species but extends with almost no changes to multi-species pure models.) This is not expected to generalize to mixed models as discussed at the end of Subsection 1.1 therein.

5.1.4 Non-uniqueness of maximizers and algorithmic symmetry breaking

In cases (b) and (c) of Theorem 5.1.12, the ODE description of maximizers does *not* uniquely determine (p,Φ). In case (b), each (p,Φ) described by Theorem 5.1.12 is specified by the point $\vec{x} = \Phi(q_1)$, which must be solvable and have the property that the tree-descending trajectory with endpoint $\vec{x}$ (unique by Proposition 5.1.11) is targeted. In case (c), each (p,Φ) is specified by the velocity $\vec{v} = \Phi'(0)$, which must satisfy (5.18) and have the property that the tree-descending trajectory with endpoint $\vec{0}$ and starting velocity $\vec{v}$ (unique by Proposition 5.1.11) is targeted. There may be multiple possible $\vec{x}$ or $\vec{v}$; see Figure 5.1.2 for examples.

In fact, even in *symmetric* two-species models – where $\vec{\lambda} = (\frac{1}{2},\frac{1}{2})$, $\check{h} = (h,h)$, and $\xi(q_1,q_2)$ is symmetric in q_1,q_2 – there may be many (p,Φ) described by Theorem 5.1.12. Moreover, surprisingly, the maximizer of (5.7) need not be symmetric! The only possible symmetric maximizer is $\Phi(q) = (q,q)$, which (for suitable p) satisfies the properties in Theorem 5.1.12. In Figures 5.1.2a and 5.1.2b we give examples of models, corresponding to cases (b) and (c) of Theorem 5.1.12, where a pair of asymmetric Φ numerically outperform



(a) $h = 1.5$, $a = 3$. Here $E_0 \approx 7.1755$, $E_1 \approx 7.1767$.

(b) $h = 0$, $a = 3$. Here $E_0 \approx 6.9230$, $E_1 \approx 6.9254$.

(c) $h = 0$, $a = 5$. Here $E_0 \approx 17.0286$, $E_1 \approx 17.0642$, $E_2 \approx 17.0292$.

Figure 5.1.2: Plots of $\Phi(q)$ with algorithmic symmetry breaking. Consider $\vec{\lambda} = (\frac{1}{2}, \frac{1}{2})$, $\check{h} = (h, h)$, and $\xi(x_1, x_2) = \nu(a\lambda_1 x_1, a\lambda_2 x_2)$ for h, a given in the captions above, where $\nu(x_1, x_2) = x_1^2 + x_1 x_2 + x_2^2 + x_1^4 + x_2^4$. Figure 5.1.2a shows an example with external field (Theorem 5.1.12(b)), Figure 5.1.2b shows an example without external field (Theorem 5.1.12(c)), and Figure 5.1.2c shows an example with several symmetry-breaking trajectories. Targeted trajectories are bold and colors have the same meaning as in Figure 5.1.1. Numerical estimates of the energy $\mathbb{A}(p, \Phi; q_0)$ attained by each bold path are given in the captions: E_0 is the energy of the diagonal trajectory and E_k is the energy of the asymmetric trajectories that intersect the diagonal k times not including $(0, 0)$. In all cases the asymmetric trajectories outperform the symmetric trajectory, and in Figure 5.1.2c the asymmetric trajectories farthest from diagonal perform the best.

the symmetric Φ. We name this phenomenon *algorithmic symmetry breaking*.[4] The presence of algorithmic symmetry breaking implies that there exist symmetric models where the best instantiation of the multi-species Subag algorithm advances through the species asymmetrically. Note that it is impossible for solutions to a first order ODE to cross, but the tree-descending ODE is second order which enables this behavior.

It is also possible to have several trajectories satisfying the ODE description in Theorem 5.1.12 and we expect an unbounded number can coexist, see Figure 5.1.2c. While it is a priori unclear that the extremal trajectories attaining value E_1 (defined in the caption) outperform the diagonal trajectory, there is a simple reason the diagonal-crossing trajectories attaining E_2 cannot be optimal: if these two trajectories were optimal, then joining their above-diagonal parts would yield another global maximizer which is not C^1 and in particular does not satisfy the ODE description of Theorem 5.1.12. (Note also that different trajectories must have different derivatives where they meet, given their description by a second order ODE.) We leave the question of characterizing global maximizers in the presence of algorithmic symmetry breaking for future work.

We emphasize that algorithmic symmetry breaking is not a barrier to efficient/Lipschitz algorithms, as the optimal (p, Φ, q_0) for the variational principle needs to be computed only once. Moreover ξ is convex in the examples shown in Figure 5.1.2, so algorithmic symmetry breaking is not related to the failure of the interpolation method to determine the free energy (obtained for convex ξ in [BS22b]).

However the presence of algorithmic symmetry breaking *does* mean that optimal Lipschitz algorithms cannot respect the symmetry between species. Namely if $|\mathcal{I}_1| = |\mathcal{I}_2|$ (so that $\vec{\lambda} = (1/2, 1/2)$), it is natural to define $T(\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2) = (\boldsymbol{\sigma}_2, \boldsymbol{\sigma}_1)$. And to study Lipschitz $\mathcal{A}$ obeying the additional condition that $\mathcal{A}(\breve{H}_N) = T(\mathcal{A}(H_N))$ where $\breve{H}_N(\boldsymbol{\sigma}) = H_N(T(\boldsymbol{\sigma}))$. However it easily follows from our methods that such algorithms cannot outperform the best value of $\mathbb{A}$ attained by $\Phi(q) = (q, q)$. In particular, such methods are suboptimal when algorithmic symmetry breaking holds.

Assuming non-degeneracy, we show that algorithmic symmetry breaking does not occur sufficiently close to $\vec{0}$. To make this precise, let $\Delta^r = \{\vec{v} \in \mathbb{R}_{\geq 0}^{\mathscr{S}} : \langle \vec{\lambda}, \vec{v} \rangle = 1\}$ denote the simplex of admissible Φ' vectors.

[4] While we don't prove rigorously that these examples exhibit algorithmic symmetry breaking, it can be verified explicitly that for the model $\xi(x, y) = x^4 + y^4 + 24xy$, $\check{h} = (0, 0)$ with endpoint $\vec{a} = (10, 10)$ (cf. Remark 5.1.13), the symmetric path $\Phi(q) = (q, q)$ is not even a local optimum as witnessed by $\Phi^{\varepsilon}(q) = (q + \varepsilon \sin(\pi q/10), q)$.

Then if $\check{h} = \vec{0}$, we define a map $F_t : \Delta^r \to \Delta^r$ given by

$$F_t(\vec{v}) = \Phi(t)/t \tag{5.24}$$

where Φ is the tree-descending trajectory with endpoint $\Phi(0) = \vec{0}$, $\Phi'(q) = \vec{v}$. The next proposition shows that F_t is injective for small t, i.e. algorithmic symmetry breaking is absent sufficiently close to the origin, and is surjective for all t.

Proposition 5.1.22. *Assume ξ is non-degenerate and $\check{h} = \vec{0}$. There exists $\varepsilon > 0$ such that the map F_t defined in (5.24) is injective for $t \in (0, \varepsilon]$. Moreover F_t is surjective for all $t > 0$.*

We expect that similar non-uniqueness is possible for q_0 and p as well as Φ. For example, this likely holds for carefully chosen asymmetric perturbations of those ξ in Figure 5.1.2. However we do not know of specific examples.

5.1.5 Branching overlap gap property as a tight barrier to algorithms

Mean-field spin glasses, including the multi-species models we focus on here, are natural examples of random optimization problems. Other examples are random constraint satisfaction problems such as random (max)-k-SAT and random perceptron models. For any such problem, a basic property to understand is the maximum objective that an efficient algorithm can find.

Since the early 2000s, there has been extensive heuristic work in the physics and computer science communities aiming to understand this question in terms of geometric properties of these problems' solution spaces [KMRT$^+$07, ZK07, AC08]. The first rigorous link from solution geometry to hardness was obtained by Gamarnik and Sudan [GS17a], in the form of the Overlap Gap Property (OGP). An OGP argument shows that the absence of a certain geometric constellation in the super-level set $S_E(H_N) = \{\boldsymbol{\sigma} : H_N(\boldsymbol{\sigma})/N \geq E\}$ implies that suitably stable algorithms cannot find objectives larger than E. The proof is by contradiction, showing that a stable algorithm attaining value E can construct the forbidden constellation.

The value E at which the constellation disappears (and at which hardness is shown) depends on the constellation and does not generally equal the value ALG found by the best efficient algorithm. The first OGP works used as the constellation a pair of solutions with medium overlap [GS17a, GJ21, CGPR19, GJW20]. Subsequent work considered constellations with more points, arranged in a "star" [RV17, GS17b, GK23, GKPX22] or "ladder" [Wei22, BH22] configuration; these constellations vanish at smaller E, thereby showing hardness closer to ALG. In particular, [RV17, Wei22] identify the computational threshold of maximum independent set on $G(N, d/N)$ within a $1 + o_d(1)$ factor, and [BH22] identifies that of random k-SAT within a constant factor clause density. We refer the reader to [HS25, Sections 1.2 and 1.3] for a more detailed discussion and [Gam21] for a survey of OGP.

Our previous work [HS25] introduced the *branching OGP*, where the forbidden constellation is a densely branching ultrametric tree. For mixed even p-spin models, this work showed that this constellation is absent for any $E > \mathsf{ALG}$, and therefore Lipschitz algorithms cannot surpass ALG. It was further shown that for these models, any ultrametric constellation that is not densely branching is *not* forbidden at all $E > \mathsf{ALG}$, and thus the branching OGP is necessary to show hardness at ALG. As discussed previously, the hardness proof of [HS25] uses interpolation to upper bound the maximum energy of the ultrametric constellation, and hence does not apply with odd interactions or more generally in multi-species models.

In Section 5.3, we develop a new method to establish the branching OGP which does not rely on interpolation. Instead we recursively apply a uniform concentration idea of Subag [Sub24] (see Lemma 5.3.2) to show that among all densely branching ultrametric constellations, the highest energy ones can be constructed *greedily.* Roughly speaking, in such constellations the children $\boldsymbol{x}^1, \ldots, \boldsymbol{x}^k$ of a point $\boldsymbol{x}$ lie on a small sphere centered at $\boldsymbol{x}$ such that the increments $\boldsymbol{x}^i - \boldsymbol{x}$ are orthogonal to $\boldsymbol{x}$ and to each other, and approximately maximize H_N on this set. Because the aforementioned generalized Subag algorithm traces a root-to-leaf path of this tree, this method automatically finds a matching algorithm and lower bound (again modulo that the greedy algorithm is not clearly Lipschitz; our AMP algorithm in [HS24] also descends this tree). In other words, the optimal algorithm can be read off from the proof of the lower bound.

We remark that in the branching OGP (and many previous OGPs) one must actually consider a family of correlated Hamiltonians. In the branching OGP the correlation structure of these Hamiltonians is also ultrametric. The function p in (5.6) enters to parametrize the correlation structure of this Hamiltonian family, see Subsection 5.2.2.

Finally, let us point out that the branching OGP is somewhat of a counterpart to the ultrametricity of low-temperature Gibbs measures mentioned previously. One essentially expects that $\mathsf{ALG} = \mathsf{OPT}$ holds whenever the Gibbs measure branches at all depths in a suitable zero-temperature limit, which is a strong form of *full replica symmetry breaking*. However, in general the true Gibbs measures may not exhibit full RSB and may even have finite combinatorial depth, whereas the algorithmic trees we consider must always branch continuously.

5.1.6 Other related work

Following the introduction of mean-field spin glasses in [SK75], a great deal of effort has been devoted to computing their free energy. In [Par79], Parisi conjectured the value of the free energy based on his celebrated ultrametric ansatz. Following progress by [MV85, Rue87, GT02, ASS03], the Parisi formula was confirmed by [Tal06b, Tal06a, Pan13a], and the zero-temperature Parisi formula for the ground state energy by [AC17, CS17]. An understanding of the high temperature regime was obtained earlier in [ALR87, CN95] and through Talagrand's cavity method [Tal10].

Another important line of work is the landscape complexity, i.e. the determination of the exponential growth rate of critical points of H_N at each energy level. Such asymptotics were put forward in [CLR03, CLR05, Par06] followed by much rigorous progress in [ABČ13, AB13, Sub17a, BSZ20, McK24, Kiv23, SZ21]. The dynamical behavior of spin glasses is also of great interest; as previously mentioned, the behavior of e.g. Langevin dynamics has been described on dimension-free time-scales. At high temperature, fast mixing has been recently established in [EKZ22, AJK$^+$22, ABXY24].

The first multi-species spin glass to be introduced was the bipartite Sherrington-Kirkpatrick model in [KC75]. It was later studied further in [KS85, FKS87a, FKS87b]. While the analogous lower bound to the Parisi formula applies in general with a similar proof [Pan15], the upper bound is known only in special cases: models where ξ is convex in the positive orthant [BCMT15, BL20], pure spherical models assuming the $N \to \infty$ limit exists [Sub23b], and spherical models for which $\vec{1}$ is super-solvable [HS23c]. A different free energy upper bound, in the form of an infinite-dimensional Hamilton-Jacobi equation, was recently proved by Mourrat [Mou23].

In the large degree limit, the maxima of random constraint satisfaction problems such as max-k-SAT and MaxCut are described by Ising mean-field models [DMS17, Pan18]. See [AMS23a, JMSS23] for algorithmic analogs.

5.1.7 Notations and preliminaries

Throughout this paper we adopt the following notational conventions. For $\boldsymbol{x} \in \mathbb{R}^N$, $\boldsymbol{x}_s \in \mathbb{R}^{\mathcal{I}_s}$ denotes the restriction of $\boldsymbol{x}$ to the coordinates $\mathcal{I}_s$. The symbol $\odot$ denotes coordinate-wise product, and the symbol $\diamond$ denotes the operation defined in (5.1). The all-0 and all-1 vectors in $\mathbb{R}^{\mathscr{S}}$ are denoted $\vec{0}, \vec{1}$, and those in $\mathbb{R}^N$ are denoted $\mathbf{0}, \mathbf{1}$. For vectors $\vec{x}, \vec{y} \in \mathbb{R}^{\mathscr{S}}$, $\vec{x} \preceq \vec{y}$ denotes the coordinate-wise inequality, and for matrices $\preceq$ denotes the Loewner order. Vector operations such as $\sqrt{\vec{x}}$ are always coordinate-wise.

Let $S_N = \{\boldsymbol{x} \in \mathbb{R}^N : \|\boldsymbol{x}\|_2^2 = N\}$. For any tensor $\boldsymbol{A} \in (\mathbb{R}^N)^{\otimes k}$, we define the operator norm

$$\|\boldsymbol{A}\|_{\mathsf{op}} = \frac{1}{N} \max_{\boldsymbol{\sigma}^1, \dots, \boldsymbol{\sigma}^k \in S_N} \left| \langle \boldsymbol{A}, \boldsymbol{\sigma}^1 \otimes \cdots \otimes \boldsymbol{\sigma}^k \rangle \right|.$$

The following proposition shows that with all but exponentially small probability, the operator norms of all constant-order gradients of H_N are bounded and $O(1)$-Lipschitz.

Proposition 5.1.23. *For any fixed model $(\xi, \check{h})$ there exists a constant $c > 0$, sequence $(K_N)_{N \ge 1}$ of convex sets $K_N \subseteq \mathscr{H}_N$, and sequence of constants $(C_k)_{k \ge 1}$ independent of N, such that the following properties hold.*

(a) $\P[H_N \in K_N] \ge 1 - e^{-cN}$;

(b) *For all $H_N \in K_N$ and $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{B}_N$ and $k \geq 1$:*

$$\left\|\nabla^k H_N(\boldsymbol{x})\right\|_{\mathsf{op}} \leq C_k, \tag{5.25}$$

$$\left\|\nabla^k H_N(\boldsymbol{x}) - \nabla^k H_N(\boldsymbol{y})\right\|_{\mathsf{op}} \leq \frac{C_{k+1}}{\sqrt{N}} \|\boldsymbol{x}-\boldsymbol{y}\|_2. \tag{5.26}$$

Proof. Note that the conditions (5.25) and (5.26) are convex in H_N. Defining K_N to be the set of H_N such that the estimates (5.25), (5.26) hold with suitably large implicit constants, it remains to show point (a). For this, by Slepian's lemma it suffices to consider the case where $\gamma_{s_1,\dots,s_k}$ is replaced by the maximal entry in $\Gamma^{(k)}$. The result then follows by [HS25, Proposition 2.3] since we assumed at the outset that $\sum_{k\geq 2} 2^k \left\|\Gamma^{(k)}\right\|_\infty < \infty$. □

Acknowledgements

We thank Mehtaab Sawhney for pointing us to [BBvH23], and Jean-Christophe Mourrat and Nike Sun for helpful feedback. The initial ideas for this work were generated while the authors were visiting the Computational Complexity of Statistical Inference program at the Simons Institute in Fall 2021.

5.2 Algorithmic thresholds from branching OGP

We begin this section by recalling some fundamental definitions and constructions from [HS25]. We then review the details of the branching overlap gap property introduced in [HS25], and in particular the link to hardness for overlap concentrated algorithms.

5.2.1 Correlation functions and overlap concentration

For any $p \in [0,1]$, we may construct two correlated copies $H_N^{(1)}, H_N^{(2)}$ of H_N as follows. Construct three i.i.d. copies $\widetilde{H}_N^{[0]}, \widetilde{H}_N^{[1]}, \widetilde{H}_N^{[2]}$ of $\widetilde{H}$ as in (5.3). For $i = 1, 2$ define

$$\begin{aligned} H_N^{(i)}(\boldsymbol{\sigma}) &= \langle \boldsymbol{h}, \boldsymbol{\sigma} \rangle + \widetilde{H}_N^{(i)}(\boldsymbol{\sigma}), \quad \text{where} \\ \widetilde{H}_N^{(i)}(\boldsymbol{\sigma}) &= \sqrt{p}\widetilde{H}_N^{[0]}(\boldsymbol{\sigma}) + \sqrt{1-p}\widetilde{H}_N^{[i]}(\boldsymbol{\sigma}). \end{aligned}$$

We say $H_N^{(1)}, H_N^{(2)}$ are p-correlated. Note that pairs of corresponding entries in $\boldsymbol{g}(H_N^{(1)})$ and $\boldsymbol{g}(H_N^{(2)})$ are Gaussian with covariance $\left[\begin{smallmatrix} 1 & p \\ p & 1 \end{smallmatrix}\right]$.

Given a function $\mathcal{A}_N : \mathscr{H}_N \to \mathcal{B}_N$ (always assumed to be measurable) define $\vec{\chi} : [0,1] \to \mathbb{R}^{\mathscr{S}}$ by

$$\vec{\chi}(p) = \mathbb{E}\vec{R}\left(\mathcal{A}(H_N^{(1)}), \mathcal{A}(H_N^{(2)})\right), \tag{5.27}$$

where $H_N^{(1)}, H_N^{(2)}$ are p-correlated copies of H_N. We say that $\vec{\chi}$ is the **correlation function** of $\mathcal{A}$. Let χ_s denote the s-coordinate of $\vec{\chi}$.

Proposition 5.2.1. *We have $\vec{\chi} \in \mathbb{I}(0,1)^{\mathscr{S}}$.*

Proof. Identically to [HS25, Proposition 3.1], Hermite expanding $R_s\left(\mathcal{A}(H_N^{(1)}), \mathcal{A}(H_N^{(2)})\right)$ shows that χ_s is continuous and non-decreasing. The same Hermite expansion shows χ_s is continuously differentiable. □

The other properties of correlation functions proved in [HS25, Proposition 3.1] also hold, namely that χ_s is convex and either strictly increasing or constant; however they are not needed in this paper.

We will determine the maximum energy attained by algorithms $\mathcal{A}_N : \mathscr{H}_N \to \mathcal{B}_N$ obeying the following overlap concentration property.

Definition 5.2.2. Let $\eta, \nu > 0$. An algorithm $\mathcal{A} = \mathcal{A}_N$ is (η, ν) overlap concentrated if for any $p \in [0,1]$ and p-correlated Hamiltonians $H_N^{(1)}, H_N^{(2)}$,

$$\P\left[\left\|\vec{R}\left(\mathcal{A}(H_N^{(1)}), \mathcal{A}(H_N^{(2)})\right) - \vec{\chi}(p)\right\|_\infty \geq \eta\right] \leq \nu. \tag{5.28}$$

Our main hardness result is the following bound on the performance of overlap concentrated algorithms.

Theorem 5.2.3. *Consider a multi-species spherical spin glass Hamiltonian H_N with parameters $(\xi, \check{h})$. Let* ALG *be given by (5.7). For any $\varepsilon > 0$ there are η, c, N_0 depending only on $\xi, \check{h}, \varepsilon$ such that the following holds for any $N \ge N_0$ and $\nu \in [0,1]$. For any (η,ν)-overlap concentrated $\mathcal{A}_N : \mathscr{H}_N \to \mathcal{B}_N$,*

$$\mathbb{P}\left[H_N(\mathcal{A}_N(H_N))/N \ge \mathsf{ALG} + \varepsilon\right] \le \exp(-cN) + \nu^c.$$

By Gaussian concentration of measure (see [HS25, Propositon 8.2]), any τ-Lipschitz algorithm is $(\eta, e^{-c(\eta,\tau)N})$-overlap concentrated for any $\eta > 0$ and appropriate $c(\eta,\tau) > 0$. Thus Theorem 5.2.3 implies Theorem 5.1.1.

5.2.2 Ultrametrically correlated Hamiltonians

Next we define the hierarchically correlated ensemble of Hamiltonians used to define the branching overlap gap property. Let $k \ge 2$, $D \ge 1$ be positive integers. For each $0 \le d \le D$, let $V_d = [k]^d$ denote the set of length d sequences of elements of $[k]$. The set V_0 consists of the empty tuple, which we denote $\emptyset$. Let $\mathbb{T}(k,D)$ denote the depth D tree rooted at $\emptyset$ with depth d vertex set V_d, where $u \in V_d$ is the parent of $v \in V_{d+1}$ if u is the length d initial substring of v. For nodes $u^1, u^2 \in \mathbb{T}(k,D)$, let

$$u^1 \wedge u^2 = \max\left\{d \in \mathbb{Z}_{\ge 0} : u^1_{d'} = u^2_{d'} \text{ for all } 1 \le d' \le d\right\},$$

where the set on the right-hand side always contains 0 vacuously. This is the depth of the least common ancestor of u^1 and u^2. Let $\mathbb{L}(k,D) = V_D$ denote the set of leaves of $\mathbb{T}(k,D)$. When k, D are clear from context, we denote $\mathbb{T}(k,D)$ and $\mathbb{L}(k,D)$ by $\mathbb{T}$ and $\mathbb{L}$. Finally, let $K = |\mathbb{L}| = k^D$.

Let the sequences $\underline{p} = (p_0, p_1, \dots, p_D) \in \mathbb{R}^{D+1}$ and $\underline{\vec{\phi}} = (\vec{\phi}_0, \vec{\phi}_1, \dots, \vec{\phi}_D) \in (\mathbb{R}^{\mathscr{S}})^{D+1}$ satisfy

$$0 = p_0 \le p_1 \le \cdots \le p_D = 1,$$
$$\vec{0} \preceq \vec{\phi}_0 \preceq \vec{\phi}_1 \preceq \cdots \preceq \vec{\phi}_D \preceq \vec{1}.$$

The sequence $\underline{p}$ controls the correlation structure of our ensemble of Hamiltonians while the sequence $\underline{\vec{\phi}}$ controls the overlap structure of their inputs. For each $u \in \mathbb{T}$, including interior nodes, let $\widetilde{H}_N^{[u]}$ be an independent copy of $\widetilde{H}_N$ generated by (5.3), and let

$$\widetilde{H}_N^{(u)} = \sum_{d=1}^{|u|} \sqrt{p_d - p_{d-1}}\, \widetilde{H}_N^{[(u_1 u_2 \dots u_d)]} \tag{5.29}$$

where $|u|$ is the length of u and $(u_1 u_2 \dots u_d)$ is the length-d prefix of u. For $u \in \mathbb{L}$, define

$$H_N^{(u)}(\boldsymbol{\sigma}) = \langle \boldsymbol{h}, \boldsymbol{\sigma} \rangle + \widetilde{H}_N^{(u)}(\boldsymbol{\sigma}).$$

This constructs a Hamiltonian ensemble $(H_N^{(u)})_{u\in\mathbb{L}}$ where each $H_N^{(u)}$ is marginally distributed as H_N and each pair of Hamiltonians $H_N^{(u^1)}, H_N^{(u^2)}$ is $p_{u^1\wedge u^2}$-correlated. We define a grand Hamiltonian on states

$$\underline{\boldsymbol{\sigma}} = (\boldsymbol{\sigma}(u))_{u\in\mathbb{L}} \in (\mathbb{R}^N)^{\mathbb{L}}.$$

by

$$\mathcal{H}_N^{k,D,\underline{p}}(\underline{\boldsymbol{\sigma}}) = \frac{1}{K} \sum_{u\in\mathbb{L}} H_N^{(u)}(\boldsymbol{\sigma}(u)). \tag{5.30}$$

We denote this by $\mathcal{H}_N$ when $k, D, \underline{p}$ are clear from context. Note that we have thus far not used the definition of $\widetilde{H}_N^{(u)}$ for interior nodes $u \in \mathbb{T} \setminus \mathbb{L}$; these Hamiltonians will be useful in our analysis of the branching OGP threshold in Section 5.3. The branching OGP is defined by a maximization of $\mathcal{H}_N$ over the overlap-constrained set

$$\mathcal{Q}^{k,D,\underline{\vec{\phi}}}(\eta) = \left\{\underline{\boldsymbol{\sigma}} \in \mathcal{B}_N^{\mathbb{L}} : \left\|\vec{R}(\boldsymbol{\sigma}(u^1), \boldsymbol{\sigma}(u^2)) - \vec{\phi}_{u^1\wedge u^2}\right\|_\infty \le \eta,\ \forall u^1, u^2 \in \mathbb{L}\right\}. \tag{5.31}$$

We denote this set $\mathcal{Q}(\eta)$ when $k, D, \underline{\vec{\phi}}$ are clear from context.

5.2.3 The branching OGP threshold

We will show that overlap concentrated algorithms cannot outperform a *branching OGP* energy BOGP defined as the ground state energy of the grand Hamiltonian (5.30) in the limit of "continuously branching" ultrametrics.

Definition 5.2.4 (Branching OGP energy)**.** The energy $\mathsf{BOGP} = \mathsf{BOGP}(\xi, \check{h})$ is the infimum of energies E such that the following holds. Choose sufficiently large D, followed by small η and then large k. For any $\vec{\chi} \in \mathbb{I}(0,1)^{\mathscr{S}}$ there exists $\underline{p}$ such that for $\underline{\vec{\phi}} = \vec{\chi}(\underline{p})$ element-wise (i.e. $\vec{\phi}_d = \vec{\chi}(p_d)$),

$$\limsup_{N\to\infty} \frac{1}{N} \mathbb{E} \sup_{\underline{\boldsymbol{\sigma}} \in \mathcal{Q}(\eta)} \mathcal{H}_N(\underline{\boldsymbol{\sigma}}) \le E. \tag{5.32}$$

More explicitly,

$$\mathsf{BOGP}(\xi, \check{h}) \equiv \lim_{D\to\infty} \lim_{\eta\to 0} \lim_{k\to\infty} \sup_{\vec{\chi} \in \mathbb{I}(0,1)^{\mathscr{S}}} \inf_{\substack{\underline{\vec{\phi}}, \underline{p}: \\ \underline{\vec{\phi}} = \vec{\chi}(\underline{p})}} \limsup_{N\to\infty} \frac{1}{N} \mathbb{E} \sup_{\underline{\boldsymbol{\sigma}} \in \mathcal{Q}^{k,D,\underline{\vec{\phi}}}(\eta)} \mathcal{H}_N^{k,D,\underline{p}}(\underline{\boldsymbol{\sigma}}). \tag{5.33}$$

Our previous work [HS25] implicitly considered the same quantity. Note that the limits in (D, k, η) are decreasing, so they could actually be taken in any order (and moreover the limiting value BOGP exists apriori). Additionally the role of the infimum over $(\underline{\vec{\phi}}, \underline{p})$ is quite simple: the only important thing is to ensure both sequences increase in uniformly small steps (see Definition 5.2.7).

Section 5.3 proves the following proposition identifying BOGP with the formula (5.7) for ALG.

Proposition 5.2.5. *For all $(\xi, \check{h})$, we have $\mathsf{BOGP} = \mathsf{ALG}$.*

Let us first prove Theorem 5.2.3 assuming Proposition 5.2.5. Let $\varepsilon > 0$ be arbitrary and k, D, η be given by Definition 5.2.4 for $E = \mathsf{ALG} + \varepsilon/4$. Let $\mathcal{A} = \mathcal{A}_N : \mathscr{H}_N \to \mathcal{B}_N$ be a (η, ν)-overlap concentrated algorithm with correlation function $\vec{\chi}$. Let $\underline{p}$ and $\underline{\vec{\phi}}$ be given by Definition 5.2.4 (depending on $\vec{\chi}$). Since $\mathsf{BOGP} = \mathsf{ALG}$ by Proposition 5.2.5, for sufficiently large N

$$\frac{1}{N} \mathbb{E} \sup_{\underline{\boldsymbol{\sigma}} \in \mathcal{Q}(\eta)} \mathcal{H}_N(\underline{\boldsymbol{\sigma}}) \le \mathsf{ALG} + \varepsilon/2.$$

Let

$$\alpha_N = \mathbb{P}\left[H_N(\mathcal{A}(H_N))/N \ge \mathsf{ALG} + \varepsilon\right].$$

Let $\boldsymbol{\sigma}(u) = \mathcal{A}(H_N^{(u)})$ and $\underline{\boldsymbol{\sigma}} = (\boldsymbol{\sigma}(u))_{u\in\mathbb{L}}$. Define the events

$$\begin{aligned} S_{\text{solve}} &= \left\{ H_N^{(u)}(\boldsymbol{\sigma}(u))/N \ge \mathsf{ALG} + \varepsilon \ \forall u \in \mathbb{L} \right\}, \\ S_{\text{overlap}} &= \left\{ \underline{\boldsymbol{\sigma}} \in \mathcal{Q}(\eta) \right\}, \\ S_{\text{ogp}} &= \left\{ \sup_{\underline{\boldsymbol{\sigma}} \in \mathcal{Q}(\eta)} \mathcal{H}_N(\underline{\boldsymbol{\sigma}})/N < \mathsf{ALG} + \varepsilon \right\}. \end{aligned} \tag{5.34}$$

Proposition 5.2.6. *The following inequalities hold.*

(a) $\mathbb{P}(S_{\text{solve}}) \ge \alpha_N^K$.

(b) $\mathbb{P}(S_{\text{overlap}}) \ge 1 - K^2\nu$.

(c) $\mathbb{P}(S_{\text{ogp}}) \ge 1 - 2\exp(-cN)$ *for suitable $c = c(\varepsilon) > 0$.*

Proof of (a). We apply Jensen's inequality D times as in [HS25, Proof of Proposition 3.6(a)]. Namely, for each $0 \le d \le D$ and $v \in \mathbb{T}$ with $|v| = d$, let $S_d(v)$ be the event that all k^{D-d} leaves u in the subtree rooted at v satisfy $H_N^{(u)}(\boldsymbol{\sigma}(u)) \ge \mathsf{ALG} + \varepsilon$. Let $P_d = \mathbb{P}[S_d(v)]$, which depends only on d. Note the events

$S_d(v1), \dots, S_d(vk)$ are conditionally IID given $\widetilde{H}_N^{(v)}$. Hence Jensen's inequality on the function $f(x) = x^k$ (which is convex on $[0,1]$) shows that for $|v| = d$ with $0 \le d < D$:

$$P_d = \mathbb{E}[\mathbb{P}[S_d(v1)|\widetilde{H}_N^{(v)}]^k] \ge \mathbb{E}[\mathbb{P}[S_d(v1)|\widetilde{H}_N^{(v)}]]^k = P_{d+1}^k.$$

It follows by iterating this bound that $P_0 \ge P_D^{k^D} = \alpha_N^K$ as desired. □

Proof of (b). For each $u^1, u^2 \in \mathbb{L}$, $\mathbb{E}\vec{R}(\boldsymbol{\sigma}(u^1), \boldsymbol{\sigma}(u^2)) = \vec{\chi}(p_{u^1 \wedge u^2}) = \vec{\phi}_{u^1 \wedge u^2}$. So,

$$\mathbb{P}\left[\left\|\vec{R}(\boldsymbol{\sigma}(u^1), \boldsymbol{\sigma}(u^2)) - \vec{\phi}_{u^1 \wedge u^2}\right\|_\infty \le \eta\right] \ge 1 - \nu.$$

The result follows by a union bound on u^1, u^2. □

Proof of (c). This follows by Proposition 5.2.5 and concentration of Lipschitz functions. For the latter, we apply the Borell-TIS inequality to $Y = \frac{1}{N}\sup_{\underline{\boldsymbol{\sigma}} \in \mathcal{Q}(\eta)} \mathcal{H}_N(\underline{\boldsymbol{\sigma}})$, as in [HS25, Proof of Proposition 3.6(d)]. □

Proof of Theorem 5.2.3. Note that $S_{\text{solve}} \cap S_{\text{overlap}} \cap S_{\text{ogp}} = \emptyset$. So, $\mathbb{P}(S_{\text{solve}}) + \mathbb{P}(S_{\text{overlap}}) + \mathbb{P}(S_{\text{ogp}}) \le 2$. The bounds in Proposition 5.2.6 imply

$$\alpha_N^K \le 2\exp(-cN) + K^2\nu$$

By adjusting the constant c,

$$\alpha_N \le \exp(-cN) + \nu^c.$$

□

5.2.4 An alternate definition for the BOGP threshold

The overlap-constrained input set $\mathcal{Q}(\eta)$ used to define BOGP was designed to capture the properties of $\underline{\boldsymbol{\sigma}} = (\mathcal{A}(H_N^{(u)}))_{u \in \mathbb{L}}$. In this set, overlap constraints are enforced *globally*, between each pair of states, and the constraints are *approximate*, within a tolerance $\eta > 0$.

In this subsection, we define a variant $\mathsf{BOGP}_{\text{loc},0}$ of BOGP, based on an input set $\mathcal{Q}_{\text{loc}}(0)$, in which overlap constraints are enforced *locally*, between only adjacent and sibling nodes in $\mathbb{T}$, and the constraints are *exact*. We also enforce that the sequences p_d, $\vec{\phi}_d$ increase in small steps. To define the local constraints, we introduce the extended states

$$\underline{\boldsymbol{\rho}} = (\boldsymbol{\rho}(u))_{u \in \mathbb{T}} \in \mathcal{B}_N^{\mathbb{T}}$$

whose indices now also include interior $u \in \mathbb{T}$. For $u, v \in \mathbb{T}$, let $u \sim v$ indicate that $u = v$, or one of u, v is the parent of the other, or u, v are siblings. Define

$$\begin{aligned}
\mathcal{Q}_{\text{loc}+}^{k,D,\underline{\vec{\phi}}}(\eta) &= \left\{\underline{\boldsymbol{\rho}} \in \mathcal{B}_N^{\mathbb{T}} : \left\|\vec{R}(\boldsymbol{\rho}(u), \boldsymbol{\rho}(v)) - \vec{\phi}_{u \wedge v}\right\|_\infty \le \eta, \ \forall u \sim v\right\} \\
\mathcal{Q}_{\text{loc}}^{k,D,\underline{\vec{\phi}}}(\eta) &= \left\{\underline{\boldsymbol{\sigma}} \in \mathcal{B}_N^{\mathbb{L}} : \exists \underline{\boldsymbol{\rho}} \in \mathcal{Q}_{\text{loc}+}^{k,D,\underline{\vec{\phi}}}(\eta) \text{ such that } (\boldsymbol{\rho}(u))_{u \in \mathbb{L}} = \underline{\boldsymbol{\sigma}}\right\}.
\end{aligned}$$

We similarly omit the superscript $k, D, \underline{\vec{\phi}}$ when this is clear from context. The following definition captures the property that p_d, $\vec{\phi}_d$ increase in small steps.

Definition 5.2.7. The pair of sequences $(\underline{p}, \underline{\vec{\phi}})$ is **δ-dense** if $p_d - p_{d-1} \le \delta$ and $\vec{\phi}_d - \vec{\phi}_{d-1} \preceq \delta\vec{1}$ for all d.

The following technical condition ensures continuous dependence of orthogonal bands on their centers.

Definition 5.2.8. The function $\vec{\chi} \in \mathbb{I}(0,1)^{\mathscr{S}}$ is **δ-separated** if $\vec{\chi}(0) \succeq \delta\vec{1}$.

Define

$$\mathsf{BOGP}_{\text{loc},0} = \lim_{D \to \infty} \lim_{k \to \infty} \sup_{\substack{\vec{\chi} \in \mathbb{I}(0,1)^{\mathscr{S}} \\ D^{-2}\text{-separated}}} \inf_{\substack{\underline{\vec{\phi}}, \underline{p} : \underline{\vec{\phi}} = \vec{\chi}(\underline{p}) \\ 6r/D\text{-dense}}} \limsup_{N \to \infty} \frac{1}{N} \mathbb{E} \sup_{\underline{\boldsymbol{\sigma}} \in \mathcal{Q}_{\text{loc}}(0)} \mathcal{H}_N(\underline{\boldsymbol{\sigma}}). \tag{5.35}$$

Note that the limit in D is no longer obviously decreasing, so the existence of this limit also needs to be proven.

The following proposition, which we prove in Appendix 5.A, shows that $\mathsf{BOGP}_{\mathrm{loc},0}$ is an equivalent characterization of BOGP. This characterization will be more convenient for the proof of Proposition 5.2.5 carried out in the next section. We note that in the proof we define several more variants of BOGP and show all are equal, and it also follows that the average in the definition (5.30) of $\mathcal{H}_N$ can be replaced by a minimum with no change. This illustrates some flexibility in using the branching OGP.

Proposition 5.2.9. *The limit* $\mathsf{BOGP}_{\mathrm{loc},0}$ *exists and* $\mathsf{BOGP} = \mathsf{BOGP}_{\mathrm{loc},0}$.

The main intuition for Proposition 5.2.9 is that a tree enforcing local orthogonality constraints contains a subtree obeying global orthogonality constraints. This subtree can be built "greedily" starting from the root, at each step choosing which children of the current vertex to include. The idea is that because the children of each vertex are orthogonal, only a small number of them can be correlated with any previously constructed vertex. Furthermore, approximate orthogonality can be improved to exact orthogonality by slightly adjusting the tree.

Finally we record two useful facts.

Lemma 5.2.10. *If* $\underline{\boldsymbol{\rho}} \in \mathcal{Q}_{\mathrm{loc},+}(0)$ *and* $\bar{\boldsymbol{\rho}} = \frac{1}{K}\sum_{u\in\mathbb{L}} \boldsymbol{\rho}(u)$, *then* $\frac{1}{\sqrt{N}}\|\boldsymbol{\rho}(\emptyset) - \bar{\boldsymbol{\rho}}\|_2 \le \sqrt{D/k}$.

Proof. Define $\underline{\boldsymbol{\tau}} \in (\mathbb{R}^N)^{\mathbb{T}}$ by $\boldsymbol{\tau}(u) = \boldsymbol{\rho}(u)$ for $u \in \mathbb{L}$ and otherwise recursively $\boldsymbol{\tau}(u) = \frac{1}{k}\sum_{i=1}^k \boldsymbol{\tau}(ui)$. By bilinearity of $\vec{R}$, for all $u \in \mathbb{T}\setminus\mathbb{L}$ with $|u| = d$,

$$\vec{R}\left(\boldsymbol{\rho}(u) - \frac{1}{k}\sum_{i=1}^k \boldsymbol{\rho}(ui), \boldsymbol{\rho}(u) - \frac{1}{k}\sum_{i=1}^k \boldsymbol{\rho}(ui)\right) = \frac{1}{k}(\vec{\phi}_{d+1} - \vec{\phi}_d),$$

so

$$\frac{1}{\sqrt{N}}\left\|\boldsymbol{\rho}(u) - \frac{1}{k}\sum_{i=1}^k \boldsymbol{\rho}(ui)\right\|_2 = \sqrt{\frac{q_{d+1}-q_d}{k}},$$

where $q_d = \langle \vec{\lambda}, \vec{\phi}_d\rangle$. It is easy to see by induction on d that

$$\begin{aligned}\frac{1}{\sqrt{N}}\|\boldsymbol{\rho}(u) - \boldsymbol{\tau}(u)\|_2 &\le \frac{1}{\sqrt{N}}\left\|\boldsymbol{\rho}(u) - \frac{1}{k}\sum_{i=1}^k \boldsymbol{\rho}(ui)\right\|_2 + \frac{1}{k}\sum_{i=1}^k \frac{1}{\sqrt{N}}\|\boldsymbol{\rho}(ui) - \boldsymbol{\tau}(ui)\|_2 \\ &\le \sum_{\ell=d}^{D-1}\sqrt{\frac{q_{\ell+1}-q_\ell}{k}}.\end{aligned}$$

Since $\bar{\boldsymbol{\rho}} = \boldsymbol{\tau}(\emptyset)$,

$$\frac{1}{\sqrt{N}}\|\boldsymbol{\rho}(\emptyset) - \bar{\boldsymbol{\rho}}\|_2 \le \sum_{d=0}^{D-1}\sqrt{\frac{q_{d+1}-q_d}{k}} \le \sqrt{\frac{D}{k}}$$

by Cauchy-Schwarz. □

Lemma 5.2.11. *For any* $S \subseteq \mathcal{B}_N^{\mathbb{L}}$, $\frac{1}{N}\sup_{\underline{\boldsymbol{\sigma}}\in S}\mathcal{H}_N(\underline{\boldsymbol{\sigma}})$ *is* $O(N^{-1/2})$*-subgaussian, in particular*

$$\mathbb{P}\left[\left|\sup_{\underline{\boldsymbol{\sigma}}\in S}\mathcal{H}_N(\underline{\boldsymbol{\sigma}}) - \mathbb{E}[\sup_{\underline{\boldsymbol{\sigma}}\in S}\mathcal{H}_N(\underline{\boldsymbol{\sigma}})]\right| \ge tN^{1/2}\right] \le Ce^{-t^2/C}$$

for a constant C *and all* $t \ge 0$.

Proof. We calculate identically to [HS25, Proof of Proposition 3.6(d)] that for any fixed $\underline{\boldsymbol{\sigma}} \in \mathcal{B}_N^{\mathbb{L}}$, $\mathsf{Var}\mathcal{H}_N(\underline{\boldsymbol{\sigma}}) = O(N)$. The result follows from the Borell-TIS inequality, whose statement and proof hold for noncentered Gaussian processes with no modification. □

5.2.5 Hardness for a class of algorithms imitating Subag's approach

Here following [HS25, Section 3.7], we outline why our arguments imply hardness for a class of non-Lipschitz algorithms, which come closer to the approach of [Sub21a]. Fix $D \geq \mathbb{Z}_+$ and $\delta = 1/D$, independent of N. Given H_N, let $W(\boldsymbol{x}; H_N) \subseteq \mathbb{R}^N$ be a (measurable in $(H_N, \boldsymbol{x})$ linear subspace of dimension $\lfloor \delta N \rfloor$. Starting from the origin $\boldsymbol{x}^0 \in \mathbb{R}^N$, we repeatedly choose a uniformly random unit vector $\boldsymbol{v}^i \in W(\boldsymbol{x}^i; H_N)$ and set

$$\boldsymbol{x}^{i+1} = \boldsymbol{x}^i + \boldsymbol{v}^i \sqrt{\delta N}, \quad 0 \leq i < m.$$

Then the (random) output $\boldsymbol{x}^D \in \mathbb{R}^N$ defines a δ-subspace random walk algorithm $\mathcal{A}(H_N; \omega)$, where ω is an independent source of randomness used to specify $\boldsymbol{v}^i \in W(\boldsymbol{x}^i; H_N)$ in each step. Then our methods imply the following result.

Theorem 5.2.12. *For any $\varepsilon > 0$ there eixst $\delta_0, c > 0$ such that for any $\delta \in (0, \delta_0$ and N large enough the following holds. For any δ-subspace random walk algorithm $\mathcal{A}$,*

$$\mathbb{P}\left[H_N(\mathcal{A}(H_N, \omega))/N \geq \mathsf{ALG} + \varepsilon \text{ and } \|R(\mathcal{A}(H_N, \omega)), \mathcal{A}(H_N, \omega)) - \vec{1}\|_\infty \leq \delta_0\right] \leq e^{-cN}. \tag{5.36}$$

Proof outline. Fix $\eta = \eta(D, k, \varepsilon, \delta) > 0$ small. Given H_N, we may generate a tree of outputs isomorphic to $\mathbb{T}(k, d)$ as follows. In each of the D steps, given $\boldsymbol{x}^{(u)}$ for $|u| = d$, choose k IID unit vectors within $W(\boldsymbol{x}^{(u)}$ to define $\boldsymbol{x}^{(u1)}, \ldots, \boldsymbol{x}^{(uk)}$. For k, D independent of N, it follows by simple concentration estimates that with probability $1 - e^{-cN}$, all distinct unit vectors chosen this way have overlap at most $\eta > 0$. Using Jensen's inequality as in Proposition 5.2.6a, it suffices to show that there is an exponentially small probability for the event in (5.36) to hold simultaneously across all k^D leaves.

We first use a pruning argument to homogenize the tree, so that self-overlaps are approximately constant at each level. We will find a subtree $\mathbb{T}'$ isomorphic to $\mathbb{T}(k', d)$, where $k' \geq \Omega(k^{1/D}/\eta^r)$, such that for all $u \in \mathbb{T}'$,

$$\|R(\boldsymbol{x}^u, \boldsymbol{x}^u) - \vec{y}^{|u|}\|_\infty \leq \eta/2D$$

holds for some (possibly random) sequence $(\vec{y}^1, \ldots, \vec{y}^D)$. Namely for each $u \in [k]^{D-1}$, by the pigeonhole principle we can find some $\vec{y}^D(u)$ and distinct children $ui_1 \ldots ui_{k'}$ such that $\|R(ui_j, ui_j) - \vec{y}^D(u)\|_\infty - \eta/2D$ for each $1 \leq j \leq k'$. Recursing similarly, for each $u \in [k]^d$ we can apply the pigeonhole principle to the k sequences $(R(v, v), \vec{y}^{d+1}(v), \ldots, \vec{y}^D(v))$ for each $v = uj$ for $1 \leq j \leq k$. This allows us to find some $(\vec{y}^{d+1}(u), \ldots, \vec{y}^D(u))$ and distinct children $ui_1 \ldots ui_{k'}$ such that $\|R(ui_j, ui_j) - \vec{y}^{d+1}(u)\|_\infty - \eta/2D$ for each $1 \leq j \leq k'$, and with $\|\vec{y}^{d'}(ui_j) - \vec{y}^{d'}\|_\infty \leq \frac{(D-d')\eta}{2D}$ for each $d < d' \leq D$. Finally once $u = \emptyset$, we obtain the desired subtree $\mathbb{T}'$.

Finally, by Borell–TIS as in Proposition 5.2.6c, it is exponentially unlikely for all leaves of $\mathbb{T}'$ to have energy at least $\mathsf{BOGP} + \varepsilon$. Since $\mathsf{ALG} = \mathsf{BOGP}$ by Proposition 5.2.5, the result follows. ☐

5.3 Branching OGP from uniform concentration

We now turn to the proof of Proposition 5.2.5. In light of Proposition 5.2.9, it suffices to prove $\mathsf{BOGP}_{\text{loc},0} = \mathsf{ALG}$. We begin with a very general argument that due to the "many orthogonal increments" property at each layer of the branching tree, it suffices to consider "greedy" embeddings in some sense. This argument is essentially elementary and relies on a "uniform concentration" argument introduced by Subag in [Sub24]. To give an informal description, let $\boldsymbol{\sigma}(u) \in \mathcal{B}_N$ have some fixed self-overlap $\vec{x} \in [0,1]^r$. We consider the largest possible average Hamiltonian value on $\boldsymbol{\sigma}(u1), \ldots, \boldsymbol{\sigma}(uk)$, subject to the conditions that $R(\boldsymbol{\sigma}(ui), \boldsymbol{\sigma}(ui)) = \vec{y}$ is fixed for each $1 \leq i \leq k$ and the orthogonality conditions from before. Uniform concentration shows the difference between this value and the value at $\boldsymbol{\sigma}$ itself concentrates around an (apriori N-dependent) function of $(\vec{x}, \vec{y})$, *uniformly in* $\boldsymbol{\sigma}(u)$. Iteratively applying this down $\mathbb{T}$ reveals that to maximize the objective defining $\mathsf{BOGP}_{\text{loc},0}$, the optimal method is to start from the root $\emptyset$ and iteratively choose the "greedy" children embedding $(\boldsymbol{\sigma}(ui))_{1 \leq i \leq k}$ in each step given $\boldsymbol{\sigma}(u)$, which maximizes $\sum_{i=1}^k H_N^{[ui]}(\boldsymbol{\sigma}(ui))$ subject to the overlap and orthogonality constraints.

5.3.1 Uniform concentration

For $\vec{x} \in [0,1]^{\mathscr{S}}$, define the product of spheres

$$\mathcal{S}_N(\vec{x}) = \left\{\boldsymbol{\sigma} \in \mathbb{R}^N : \vec{R}(\boldsymbol{\sigma},\boldsymbol{\sigma}) = \vec{x}\right\} = \left\{\boldsymbol{\sigma} \in \mathbb{R}^N : \|\boldsymbol{\sigma}_s\|_2^2 = \lambda_s x_s N \ \forall s \in \mathscr{S}\right\}.$$

For $\boldsymbol{\sigma}^0 \in \mathcal{S}_N(\vec{x})$ and $\vec{y} \succeq \vec{x}$, define

$$B(\boldsymbol{\sigma}^0, \vec{y}, k) = \left\{ \begin{array}{l} \underline{\boldsymbol{\sigma}} = (\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2, \dots, \boldsymbol{\sigma}^k) \in \mathcal{S}_N(\vec{y})^k : \\ \vec{R}(\boldsymbol{\sigma}^i - \boldsymbol{\sigma}^0, \boldsymbol{\sigma}^0) = \vec{R}(\boldsymbol{\sigma}^i - \boldsymbol{\sigma}^0, \boldsymbol{\sigma}^j - \boldsymbol{\sigma}^0) = \vec{0} \quad \forall i,j \in [k], i \neq j \end{array} \right\}. \tag{5.37}$$

Let $0 \le p_- < p_+ \le 1$. Generate $k+1$ i.i.d. copies $\widehat{H}_N^{[0]}, \widehat{H}_N^{[1]}, \dots, \widehat{H}_N^{[k]}$ of $\widetilde{H}_N$ as in (5.3). Set

$$\widehat{H}_N^{(0)}(\boldsymbol{\sigma}) = \sqrt{p_-}\widehat{H}_N^{[0]}(\boldsymbol{\sigma}) \quad \text{and} \tag{5.38}$$

$$\widehat{H}_N^{(i)}(\boldsymbol{\sigma}) = \sqrt{p_-}\widehat{H}_N^{[0]}(\boldsymbol{\sigma}) + \sqrt{p_+ - p_-}\widehat{H}_N^{[i]}(\boldsymbol{\sigma}), \quad 1 \le i \le k. \tag{5.39}$$

Define

$$F_{p_-,p_+}(\boldsymbol{\sigma}^0, \vec{y}, k) = \frac{1}{kN} \max_{\underline{\boldsymbol{\sigma}} \in B(\boldsymbol{\sigma}^0, \vec{y}, k)} \sum_{i=1}^k \left(\widehat{H}_N^{(i)}(\boldsymbol{\sigma}^i) - \widehat{H}_N^{(0)}(\boldsymbol{\sigma}^0)\right).$$

Lemma 5.3.1. *There exists C such that the following holds. Suppose that $\delta\vec{1} \preceq \vec{x} \preceq \vec{y} \preceq \vec{1}$ and $\boldsymbol{\sigma}^0, \boldsymbol{\rho}^0 \in \mathcal{S}_N(\vec{x})$ satisfy $\left\|\boldsymbol{\sigma}^0 - \boldsymbol{\rho}^0\right\|_2 \le \iota\sqrt{N}$. If $\widehat{H}_N^{[0]}, \dots, \widehat{H}_N^{[k]} \in K_N$ for the event K_N in Proposition 5.1.23, then*

$$|F_{p_-,p_+}(\boldsymbol{\sigma}^0, \vec{y}, k) - F_{p_-,p_+}(\boldsymbol{\rho}^0, \vec{y}, k)| \le \frac{C\iota}{\sqrt{\delta}}. \tag{5.40}$$

Proof. Let $T : \mathbb{R}^N \to \mathbb{R}^N$ be a product of rotation maps T_s in the r factors $\mathbb{R}^{\mathcal{I}_s}$ such that $T(\boldsymbol{\sigma}^0) = \boldsymbol{\rho}^0$. Then

$$T\left(B(\boldsymbol{\sigma}^0, \vec{y}, k)\right) = B(T(\boldsymbol{\sigma}^0), \vec{y}, k) = B(\boldsymbol{\rho}^0, \vec{y}, k).$$

In particular, we take T to be obtained using geodesic rotations from each $\boldsymbol{\sigma}_s^0$ to $\boldsymbol{\rho}_s^0$. Then given $\underline{\boldsymbol{\sigma}} \in B(\boldsymbol{\sigma}^0, \vec{y}, k)$, we set $\underline{\boldsymbol{\rho}} = (\boldsymbol{\rho}^1, \dots, \boldsymbol{\rho}^k) \in B(\boldsymbol{\rho}^0, \vec{y}, k)$ by defining $\boldsymbol{\rho}^i = T\boldsymbol{\sigma}^i$ for each i. Then for all $i \in [k]$ and $s \in \mathscr{S}$:

$$\frac{\left\|\boldsymbol{\rho}_s^i - \boldsymbol{\sigma}_s^i\right\|_2}{\|\boldsymbol{\sigma}_s^i\|_2} \le \frac{\left\|\boldsymbol{\rho}_s^0 - \boldsymbol{\sigma}_s^0\right\|_2}{\|\boldsymbol{\sigma}_s^0\|_2} \le \frac{\iota}{\sqrt{\delta}}.$$

(The first estimate just says that for each T_s, the ratio $\frac{\|T_s\boldsymbol{z}_s - \boldsymbol{z}_s\|_2}{\|\boldsymbol{z}_s\|_2}$ is maximized for non-zero $\boldsymbol{z}_s \in \mathbb{R}^{\mathcal{I}_s}$ by $\boldsymbol{z}_s = \boldsymbol{\sigma}_s$; this holds by definition of T_s.) Thus $\frac{1}{\sqrt{N}}\left\|\boldsymbol{\rho}^i - \boldsymbol{\sigma}^i\right\|_2 \le \iota/\sqrt{\delta}$. On the event $\widehat{H}_N^{[0]}, \dots, \widehat{H}_N^{[k]} \in K_N$, it follows that

$$\left|\widehat{H}_N^{(i)}(\boldsymbol{\sigma}^i) - \widehat{H}_N^{(i)}(\boldsymbol{\rho}^i)\right| \le \frac{C\iota}{\sqrt{\delta}}$$

for $1 \le i \le k$ and

$$\left|\widehat{H}_N^{(0)}(\boldsymbol{\sigma}^0) - \widehat{H}_N^{(0)}(\boldsymbol{\rho}^0)\right| \le C\iota,$$

which implies the conclusion (after adjusting C). □

Lemma 5.3.2. *There exist constants $c, C > 0$ such that for all $k \in \mathbb{N}$ and $\delta, \varepsilon > 0$ the following holds. For any $\vec{x}, \vec{y}$ satisfying $\delta\vec{1} \preceq \vec{x} \preceq \vec{y}$,*

$$\mathbb{P}\left(\sup_{\boldsymbol{\sigma}^0 \in \mathcal{S}_N(\vec{x})} |F_{p_-,p_+}(\boldsymbol{\sigma}^0, \vec{y}, k) - \mathbb{E}F_{p_-,p_+}(\boldsymbol{\sigma}^0, \vec{y}, k)| \le \varepsilon\right)$$
$$\ge 1 - \exp\left(C\log\left(\frac{1}{\delta\varepsilon}\right)N - ck\varepsilon^2 N\right) - e^{-cN}$$

Proof. Fix for now $\boldsymbol{\sigma}^0 \in \mathcal{S}_N(\vec{x})$ and $\underline{\boldsymbol{\sigma}} = (\boldsymbol{\sigma}^1, \dots, \boldsymbol{\sigma}^k) \in B(\boldsymbol{\sigma}^0, \vec{y}, k)$. Using the definition (5.37) in the final step, we find that for small $c > 0$,

$$
\begin{aligned}
&\mathbb{E}\left[\left(\sum_{i=1}^k (\widehat{H}_N^{(i)}(\boldsymbol{\sigma}^i) - \widehat{H}_N^{(0)}(\boldsymbol{\sigma}^0))\right)^2\right] \\
&= \mathbb{E}\left[\left(\sum_{i=1}^k \sqrt{p_-}(\widehat{H}_N^{[0]}(\boldsymbol{\sigma}^i) - \widehat{H}_N^{[0]}(\boldsymbol{\sigma}^0)) + \sqrt{p_+ - p_-}\widehat{H}_N^{[i]}(\boldsymbol{\sigma}^i)\right)^2\right] \\
&= p_- \sum_{i,j=1}^k \mathbb{E}\left[(\widehat{H}_N^{[0]}(\boldsymbol{\sigma}^i) - \widehat{H}_N^{[0]}(\boldsymbol{\sigma}^0))(\widehat{H}_N^{[0]}(\boldsymbol{\sigma}^j) - \widehat{H}_N^{[0]}(\boldsymbol{\sigma}^0))\right] + (p_+ - p_-)\sum_{i=1}^k \mathbb{E}\left[\widehat{H}_N^{[i]}(\boldsymbol{\sigma}^i)^2\right] \\
&= p_- \sum_{i,j=1}^k \xi(\vec{R}(\boldsymbol{\sigma}^i, \boldsymbol{\sigma}^j)) - \xi(\vec{R}(\boldsymbol{\sigma}^i, \boldsymbol{\sigma}^0)) - \xi(\vec{R}(\boldsymbol{\sigma}^0, \boldsymbol{\sigma}^j)) + \xi(\vec{R}(\boldsymbol{\sigma}^0, \boldsymbol{\sigma}^0)) + (p_+ - p_-)\sum_{i=1}^k \xi(\vec{R}(\boldsymbol{\sigma}^i, \boldsymbol{\sigma}^i)) \\
&\le \frac{k}{8c}.
\end{aligned}
$$

By the Borell-TIS inequality, for each fixed $\boldsymbol{\sigma}^0 \in \mathcal{S}_N(\vec{x})$

$$
\mathbb{P}\left[|F_{p_-,p_+}(\boldsymbol{\sigma}^0, \vec{y}, k) - \mathbb{E}F_{p_-,p_+}(\boldsymbol{\sigma}^0, \vec{y}, k)| \le \varepsilon/2\right] \ge 1 - 2\exp\left(-ck\varepsilon^2 N\right). \tag{5.41}
$$

Choose $\iota = \Theta(\varepsilon\sqrt{\delta})$ so that the right-hand side of (5.40) is bounded by $\varepsilon/2$, and let $\mathcal{N}$ be an $\iota\sqrt{N}$-net of $\mathcal{S}_N(\vec{x})$ with size $|\mathcal{N}| \le (\varepsilon\sqrt{\delta})^{-CN} \le (\varepsilon\delta)^{-CN}$. Define the events

$$
\begin{aligned}
S_{\text{conc}} &= \left\{\, |F_{p_-,p_+}(\boldsymbol{\rho}^0, \vec{y}, k) - \mathbb{E}F_{p_-,p_+}(\boldsymbol{\rho}^0, \vec{y}, k)| \le \varepsilon/2 \;\; \forall\, \boldsymbol{\rho}^0 \in \mathcal{N} \right\}, \\
S_{\text{lip}} &= \left\{\, \widehat{H}_N^{[0]}, \dots, \widehat{H}_N^{[k]} \in K_N \right\},
\end{aligned}
$$

where K_N is defined in Proposition 5.1.23. By a union bound (after adjusting c, C),

$$
\mathbb{P}\left(S_{\text{conc}} \cap S_{\text{lip}}\right) \ge 1 - \exp\left(C\log\left(\frac{1}{\delta\varepsilon}\right)N - ck\varepsilon^2 N\right) - e^{-cN}. \tag{5.42}
$$

Suppose $S_{\text{conc}} \cap S_{\text{lip}}$ holds. For any $\boldsymbol{\sigma}^0 \in \mathcal{S}_N(\vec{x})$, there exists $\boldsymbol{\rho}^0 \in \mathcal{N}$ such that $\|\boldsymbol{\sigma}^0 - \boldsymbol{\rho}^0\|_2 \le \iota\sqrt{N}$, and so

$$
\begin{aligned}
&|F_{p_-,p_+}(\boldsymbol{\sigma}^0, \vec{y}, k) - \mathbb{E}F_{p_-,p_+}(\boldsymbol{\sigma}^0, \vec{y}, k)| \\
&\le |F_{p_-,p_+}(\boldsymbol{\sigma}^0, \vec{y}, k) - F_{p_-,p_+}(\boldsymbol{\rho}^0, \vec{y}, k)| + |F_{p_-,p_+}(\boldsymbol{\rho}^0, \vec{y}, k) - \mathbb{E}F_{p_-,p_+}(\boldsymbol{\rho}^0, \vec{y}, k)| \le \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.
\end{aligned}
$$

Here the last estimate follows by Lemma 5.3.1. □

For now, let $k, D, \vec{\phi}, \underline{p}$ (recall Definition 5.2.4) be arbitrary. In Proposition 5.3.3 below, we obtain an estimate for $\frac{1}{N}\mathbb{E}\max_{\underline{\boldsymbol{\sigma}} \in \mathcal{Q}_{\text{loc}}(0)} \mathcal{H}_N(\underline{\boldsymbol{\sigma}})$ by applying Lemma 5.3.2 repeatedly at each internal vertex $u \in \mathbb{T}\backslash\mathbb{L}$. This maximum will take the form of an abstract sum of energy increments. In the next subsection we will take a continuum limit of this bound, which will yield the variational formula (5.7) for ALG and prove Proposition 5.2.5.

Spherical symmetry implies that $\mathbb{E}F_{p_-,p_+}(\boldsymbol{\sigma}, \vec{y}, k)$ depends on $\boldsymbol{\sigma}$ only through $\vec{R}(\boldsymbol{\sigma}, \boldsymbol{\sigma})$. Hence for $\vec{\phi}_- = \vec{R}(\boldsymbol{\sigma}, \boldsymbol{\sigma})$ we may define

$$
f(\vec{\phi}_-, \vec{\phi}_+; p_-, p_+; k) = \mathbb{E}F_{p_-,p_+}\left(\boldsymbol{\sigma}, \vec{\phi}_+, k\right). \tag{5.43}
$$

Proposition 5.3.3. *Fix $D \in \mathbb{N}$ and $\varepsilon, \delta > 0$. Suppose that $\vec{\phi}_0 \succeq \delta\vec{1}$. There exists $k_0 = k_0(D, \varepsilon, \delta)$ such that for all $k \ge k_0$, there exists $c = c(D, \varepsilon, \delta, k)$ such that*

$$
\mathbb{P}\left[\left|\frac{1}{N}\sup_{\underline{\boldsymbol{\sigma}} \in \mathcal{Q}_{\text{loc}}(0)} \mathcal{H}_N(\underline{\boldsymbol{\sigma}}) - \left(\sum_{s \in \mathscr{S}} h_s\lambda_s\sqrt{\phi_0^s} + \sum_{d=0}^{D-1} f\left(\vec{\phi}_d, \vec{\phi}_{d+1}; p_d, p_{d+1}; k\right)\right)\right| \le 2D\varepsilon\right] \ge 1 - e^{-cN}.
$$

Proof. Let C, c be as in Lemma 5.3.2, and k_0 large enough that

$$C \log\left(\frac{1}{\delta\varepsilon}\right) - ck_0\varepsilon^2 \le -c, \tag{5.44}$$

$$\|\check{h}\|_\infty / \sqrt{k_0} \le \varepsilon. \tag{5.45}$$

Recall the construction of $\widetilde{H}_N^{(u)}$ from (5.29). For any $u \in V_d$, $0 \le d \le D-1$, let $\mathcal{E}_u$ denote the event in Lemma 5.3.2, with $(p_-, p_+) = (p_d, p_{d+1})$, $(\vec{x}, \vec{y}) = (\vec{\phi}_d, \vec{\phi}_{d+1})$, and

$$(\widehat{H}_N^{(0)}, \widehat{H}_N^{(1)}, \dots, \widehat{H}_N^{(k)}) = \left(\widetilde{H}_N^{(u)}, \widetilde{H}_N^{(u1)}, \dots, \widetilde{H}_N^{(uk)}\right). \tag{5.46}$$

Let $\mathcal{E} = \bigcap_{u \in \mathbb{T}\setminus\mathbb{L}} \mathcal{E}_u$. Lemma 5.3.2 and equation (5.44) imply $\mathbb{P}(\mathcal{E}_u) \ge 1 - 2e^{-cN}$ for all $u \in \mathbb{L}$. By a union bound, $\mathbb{P}(\mathcal{E}) \ge 1 - e^{-cN}$ (after adjusting c).

Denote by $F^u_{p_d,p_{d+1}}$ the function $F_{p_d,p_{d+1}}$ defined with Hamiltonians (5.46). Let $\underline{\boldsymbol{\sigma}} \in \mathcal{Q}_{\mathrm{loc}}(0)$, so there exists $\underline{\boldsymbol{\rho}} \in \mathcal{Q}_{\mathrm{loc}+}(0)$ with $(\boldsymbol{\rho}(u))_{u\in\mathbb{L}} = \underline{\boldsymbol{\sigma}}$. On the event $\mathcal{E}$,

$$\begin{aligned}
\frac{1}{N}\mathcal{H}_N(\underline{\boldsymbol{\sigma}}) - \frac{1}{KN}\sum_{v\in\mathbb{L}}\langle \boldsymbol{h}, \boldsymbol{\sigma}(u)\rangle &= \frac{1}{KN}\sum_{u\in\mathbb{L}} \widetilde{H}_N^{(u)}(\boldsymbol{\sigma}(u)) \\
&= \sum_{d=0}^{D-1} \frac{1}{k^d} \sum_{u\in V_d} \frac{1}{kN} \sum_{i=1}^{k} \left(\widetilde{H}_N^{(ui)}(\boldsymbol{\rho}(ui)) - \widetilde{H}_N^{(u)}(\boldsymbol{\rho}(u))\right) \\
&\le \sum_{d=0}^{D-1} \frac{1}{k^d} \sum_{u\in V_d} F^u_{p_d,p_{d+1}}\left(\boldsymbol{\rho}(u), \vec{\phi}_{d+1}, k\right) \\
&\overset{Lem.\ 5.3.2}{\le} D\varepsilon + \sum_{d=0}^{D-1} f(\vec{\phi}_d, \vec{\phi}_{d+1}; p_d, p_{d+1}; k).
\end{aligned}$$

In the telescoping sum, we used that $\widetilde{H}_N^{(\emptyset)}$ is the zero function. By Lemma 5.2.10 and equation (5.45),

$$\begin{aligned}
\left|\frac{1}{KN}\sum_{u\in\mathbb{L}}\langle \boldsymbol{h}, \boldsymbol{\sigma}(u)\rangle - \frac{1}{N}\langle \boldsymbol{h}, \boldsymbol{\rho}(\emptyset)\rangle\right| &\le \frac{1}{\sqrt{N}}\|\boldsymbol{h}\|_2 \cdot \frac{1}{\sqrt{N}}\left\|\boldsymbol{\rho}(\emptyset) - \frac{1}{K}\sum_{u\in\mathbb{L}}\boldsymbol{\sigma}(u)\right\|_2 \\
&\le \|\check{h}\|_\infty \sqrt{\frac{D}{k}} \le D\varepsilon.
\end{aligned}$$

Finally,

$$\frac{1}{N}\langle \boldsymbol{h}, \boldsymbol{\rho}(\emptyset)\rangle = \frac{1}{N}\sum_{s\in\mathscr{S}} h_s \|\boldsymbol{\rho}(\emptyset)_s\|_1 \le \frac{1}{N}\sum_{s\in\mathscr{S}} h_s \sqrt{|\mathcal{I}_s|}\|\boldsymbol{\rho}(\emptyset)_s\|_2 = \sum_{s\in\mathscr{S}} h_s \lambda_s \sqrt{\phi_0^s}. \tag{5.47}$$

This completes the proof of the upper bound for $\frac{1}{N}\sup_{\underline{\boldsymbol{\sigma}}\in\mathcal{Q}_{\mathrm{loc}}(0)} \mathcal{H}_N(\underline{\boldsymbol{\sigma}})$. Finally, observe that equality holds above (up to the same $2D\varepsilon$ error) if we choose $\boldsymbol{\rho}(\emptyset) = \sqrt{\vec{\phi}_0} \diamond \mathbf{1}$ and then recursively choose $(\boldsymbol{\rho}(ui))_{i\in[k]}$ given $\boldsymbol{\rho}(u)$ so that, for $|u| = d$,

$$\frac{1}{Nk}\sum_{i=1}^{k}\left(\widetilde{H}_N^{(ui)}(\boldsymbol{\rho}(ui)) - \widetilde{H}_N^{(u)}(\boldsymbol{\rho}(u))\right) = F^u_{p_d,p_{d+1}}(\boldsymbol{\rho}(u), \vec{\phi}_{d+1}, k). \qquad \square$$

5.3.2 The algorithmic functional

Our next objective is to estimate the terms $f(\vec{\phi}_-, \vec{\phi}_+; p_-, p_+; k)$ appearing in Proposition 5.3.3. The key point is that when the differences $\vec{\phi}_+ - \vec{\phi}_-$ and $p_+ - p_-$ are small, which is ensured by δ-denseness of $(\underline{p}, \underline{\vec{\phi}})$, this estimate only requires Taylor approximating the relevant Hamiltonians to second order. We take advantage of this using the following lemma, which (for $k = 1$) gives the ground state energy $\mathbf{GS}(W, \vec{v}, 1)$ of a quadratic multi-species spin glass with Gaussian external field. For general k, this lemma gives the limiting ground state energy $\mathbf{GS}(W,\vec{v},k)$ of a k-replica Hamiltonian (5.48) with shared quadratic component $W \diamond \boldsymbol{G}$ and independent external fields $\vec{v}\diamond \boldsymbol{g}^i$, whose inputs (5.49) are k pairwise orthogonal elements of $\mathcal{S}_N(\vec{1})$. Note that $\mathbf{GS}(W,\vec{v},k) \le \mathbf{GS}(W,\vec{v},1)$ by definition. In fact equality holds, i.e. there exist orthogonal $\boldsymbol{\sigma}^1,\dots,\boldsymbol{\sigma}^k$ such that each $\boldsymbol{\sigma}^i$ approximately maximizes $H^i_{N,k}(\boldsymbol{\sigma}^i)$. We prove this lemma in Appendix 5.B by combining a known formula from [BBvH23] for the case $(k,\vec{v}) = (1,\vec{0})$ with an elementary recursive argument along subspaces.

Lemma 5.3.4. *Let $W = (w_{s,s'})_{s,s\in\mathscr{S}} \in \mathbb{R}^{\mathscr{S}\times\mathscr{S}}_{\ge 0}$ be symmetric and $\vec{v} = (v_s)_{s\in\mathscr{S}} \in \mathbb{R}^{\mathscr{S}}_{\ge 0}$. Let $k\in\mathbb{N}$ and sample independent $\boldsymbol{g}^1,\dots,\boldsymbol{g}^k \in \mathbb{R}^N$ and $\boldsymbol{G}\in\mathbb{R}^{N\times N}$ with i.i.d. standard Gaussian entries. Consider the k-replica Hamiltonian*

$$H_{N,k}(\underline{\boldsymbol{\sigma}}) = \frac{1}{k}\sum_{i=1}^k H^i_{N,k}(\boldsymbol{\sigma}^i), \qquad H^i_{N,k}(\boldsymbol{\sigma}^i) = \langle \vec{v}\diamond \boldsymbol{g}^i, \boldsymbol{\sigma}^i\rangle + \frac{1}{\sqrt{N}}\langle W\diamond \boldsymbol{G}, (\boldsymbol{\sigma}^i)^{\otimes 2}\rangle \tag{5.48}$$

on the input space of orthogonal replicas

$$\mathcal{S}^{k,\perp}_N = \left\{\underline{\boldsymbol{\sigma}} = (\boldsymbol{\sigma}^1,\dots,\boldsymbol{\sigma}^k) \in \mathcal{S}_N(\vec{1}) : \vec{R}(\boldsymbol{\sigma}^i,\boldsymbol{\sigma}^j) = \vec{0}\ \forall i\neq j\right\}. \tag{5.49}$$

Define the k-replica ground state energy

$$\mathrm{GS}_N(W,\vec{v},k) \equiv \frac{1}{N}\max_{\underline{\boldsymbol{\sigma}}\in\mathcal{S}^{k,\perp}_N} H_{N,k}(\underline{\boldsymbol{\sigma}}). \tag{5.50}$$

Then $\mathbf{GS}(W,\vec{v},k) \equiv \lim_{N\to\infty}\mathbb{E}\mathrm{GS}_N(W,\vec{v},k)$ exists, does not depend on k, and is given by

$$\mathbf{GS}(W,\vec{v},k) = \sum_{s\in\mathscr{S}}\lambda_s\sqrt{v_s^2 + 2\sum_{s'\in\mathscr{S}}\lambda_s w_{s,s'}^2}.$$

Proposition 5.3.5. *Suppose $0\le p_-\le p_+\le 1$, $\vec{0}\preceq\vec{\phi}_-\preceq\vec{\phi}_+\preceq\vec{1}$ and*

$$p_+ - p_- \le \delta, \qquad \vec{\phi}_+ - \vec{\phi}_- \preceq \delta\vec{1}. \tag{5.51}$$

Then,

$$\begin{aligned} f(\vec{\phi}_-,\vec{\phi}_+;p_-,p_+;k) = \sum_{s\in\mathscr{S}}\lambda_s\sqrt{(\phi^s_+ - \phi^s_-)\left((p_+-p_-)\xi^s(\vec{\phi}_-) + p_-\sum_{s'\in\mathscr{S}}\partial_{x_{s'}}\xi^s(\vec{\phi}_-)(\phi^s_+ - \phi^s_-)\right)} \\ + O\big(\delta^{3/2} + (\delta/k)^{1/2}\big) + o_N(1), \end{aligned}$$

where $o_N(1)$ denotes a term tending to 0 as $N\to\infty$.

Proof. Fix $\boldsymbol{\sigma}^0$ such that $\vec{R}(\boldsymbol{\sigma}^0,\boldsymbol{\sigma}^0) = \vec{\phi}_-$. Let $\underline{\boldsymbol{\sigma}} = (\boldsymbol{\sigma}^1,\dots,\boldsymbol{\sigma}^k)\in B(\boldsymbol{\sigma}^0,\vec{\phi}_+,k)$. Let $\Delta\vec{\phi} = \vec{\phi}_+ - \vec{\phi}_-$ and $\underline{\boldsymbol{x}} = (\boldsymbol{x}^1,\dots,\boldsymbol{x}^k)$ for $\boldsymbol{x}^i = (\Delta\vec{\phi})^{-1/2}\diamond(\boldsymbol{\sigma}^i - \boldsymbol{\sigma}^0)$. Define

$$\begin{aligned} \mathcal{S}_\bullet &= \left\{\boldsymbol{y}\in\mathcal{S}_N(\vec{1}) : \vec{R}(\boldsymbol{y},\boldsymbol{\sigma}^0) = \vec{0}\right\}, \\ \mathcal{S}^{k,\perp}_\bullet &= \left\{\underline{\boldsymbol{y}} = (\boldsymbol{y}^1,\dots,\boldsymbol{y}^k)\in\mathcal{S}^k_\bullet : \vec{R}(\boldsymbol{y}^i,\boldsymbol{y}^j) = \vec{0}\ \forall i\neq j\right\}. \end{aligned}$$

Note that $\underline{\boldsymbol{x}}\in\mathcal{S}^{k,\perp}_\bullet$. Recall that $\widehat{H}^{[0]}_N,\dots,\widehat{H}^{[k]}_N$ are i.i.d. copies of $\widetilde{H}_N$, and that $\widehat{H}^{(0)}_N,\dots,\widehat{H}^{(k)}_N$ are defined by (5.38), (5.39). Let

$$\overline{H}^i_N(\boldsymbol{x}^i) = \widehat{H}^{[i]}_N(\boldsymbol{\sigma}^i) - \widehat{H}^{[i]}_N(\boldsymbol{\sigma}^0) = \widehat{H}^{[i]}_N\left(\boldsymbol{\sigma}^0 + \sqrt{\Delta\vec{\phi}}\diamond\boldsymbol{x}^i\right) - \widehat{H}^{[i]}_N(\boldsymbol{\sigma}^0).$$

Then

$$
\begin{aligned}
f(\vec{\phi}_-,\vec{\phi}_+;p_-,p_+;k) &= \frac{1}{kN}\mathbb{E}\max_{\underline{\boldsymbol{\sigma}}\in B(\boldsymbol{\sigma}^0,\vec{\phi}_+,k)}\sum_{i=1}^k\left(\widehat{H}_N^{(i)}(\boldsymbol{\sigma}^i)-\widehat{H}_N^{(0)}(\boldsymbol{\sigma}^0)\right)\\
&= \frac{1}{kN}\mathbb{E}\max_{\underline{\boldsymbol{\sigma}}\in B(\boldsymbol{\sigma}^0,\vec{\phi}_+,k)}\sum_{i=1}^k\Bigg(\sqrt{p_-}\left(\widehat{H}_N^{[0]}(\boldsymbol{\sigma}^i)-\widehat{H}_N^{[0]}(\boldsymbol{\sigma}^0)\right)\\
&\qquad+\sqrt{p_+-p_-}\left(\widehat{H}_N^{[i]}(\boldsymbol{\sigma}^i)-\widehat{H}_N^{[i]}(\boldsymbol{\sigma}^0)\right)+\sqrt{p_+-p_-}\,\widehat{H}_N^{[i]}(\boldsymbol{\sigma}^0)\Bigg)\\
&= \frac{1}{kN}\mathbb{E}\max_{\underline{\boldsymbol{x}}\in\mathcal{S}_\bullet^{k,\perp}}\sum_{i=1}^k\left(\sqrt{p_-}\,\overline{H}_N^0(\boldsymbol{x}^i)+\sqrt{p_+-p_-}\,\overline{H}_N^i(\boldsymbol{x}^i)\right)
\end{aligned}
\tag{5.52}
$$

where we note that $\mathbb{E}\widehat{H}_N^{[i]}(\boldsymbol{\sigma}^0)=0$. Let $\overline{H}_N^{i,\mathrm{tay}}$ denote the degree 2 Taylor expansion of $\overline{H}_N^i$ around $\mathbf{0}$. By Proposition 5.1.23 (recalling (5.51)),

$$
\mathbb{E}\sup_{\boldsymbol{x}\in\mathcal{S}_\bullet}|\overline{H}_N^i(\boldsymbol{x})-\overline{H}_N^{i,\mathrm{tay}}(\boldsymbol{x})|=O(N\delta^{3/2}).
$$

So, for all $0\le i\le k$, we have as processes on $\mathcal{S}_\bullet$

$$
\overline{H}_N^i(\boldsymbol{x})=_d\langle\vec{v}\diamond\boldsymbol{g}^i,\boldsymbol{x}\rangle+\langle W\diamond\boldsymbol{G}^i,\boldsymbol{x}^{\otimes 2}\rangle+O_{\mathbb{P}}(N\delta^{3/2}),
\tag{5.53}
$$

where $O_{\mathbb{P}}(N\delta^{3/2})$ denotes a $\mathcal{S}_\bullet$-valued process $X(\boldsymbol{x})$ with $\mathbb{E}\sup_{\boldsymbol{x}\in\mathcal{S}_\bullet}|X(\boldsymbol{x})|=O(N\delta^{3/2})$ and $\vec{v}=(v_s)_{s\in\mathscr{S}}$ and $W=(w_{s,s'})_{s,s'\in\mathscr{S}}$ are given by

$$
v_s=\sqrt{\xi^s(\vec{\phi}_-)(\Delta\vec{\phi})^s},\qquad w_{s,s'}=\frac{1}{\sqrt{2}}\sqrt{\lambda_{s'}^{-1}\partial_{x_{s'}}\xi^s(\vec{\phi}_-)(\Delta\vec{\phi})^s(\Delta\vec{\phi})^{s'}}.
$$

Next we observe some simplifications. Because $\Delta\vec{\phi}\preceq\delta\vec{1}$, we have $v_s=O(\delta^{1/2})$, $w_{s,s'}=O(\delta)$ uniformly over s,s'. The linear contribution to $\overline{H}_N^0$ in (5.53) is small because

$$
\frac{1}{kN}\sum_{i=1}^k\langle\vec{v}\diamond\boldsymbol{g}^0,\boldsymbol{x}^i\rangle\le\frac{1}{kN}\left\|\vec{v}\diamond\boldsymbol{g}^0\right\|_2\left\|\sum_{i=1}^k\boldsymbol{x}^i\right\|_2=O_{\mathbb{P}}((\delta/k)^{1/2})
$$

by orthogonality of the $\boldsymbol{x}^i$. Because $p_+-p_-\le\delta$, the quadratic contributions to $\overline{H}_N^i$ for $i\ge 1$ are also small:

$$
\frac{\sqrt{p_+-p_-}}{N}\langle W\diamond\boldsymbol{G}^i,(\boldsymbol{x}^i)^{\otimes 2}\rangle=O_{\mathbb{P}}(\delta^{3/2}).
$$

Combining these estimates with (5.52) and (5.53), we find

$$
\begin{aligned}
f(\vec{\phi}_-,\vec{\phi}_+;p_-,p_+;k) &= \frac{1}{kN}\mathbb{E}\max_{\underline{\boldsymbol{x}}\in\mathcal{S}_\bullet^{k,\perp}}\sum_{i=1}^k\sqrt{p_+-p_-}\left\langle\vec{v}\diamond\boldsymbol{g}^i,\boldsymbol{x}^i\right\rangle+\sqrt{p_-}\left\langle W\diamond\boldsymbol{G}^0,(\boldsymbol{x}^i)^{\otimes 2}\right\rangle\\
&\qquad+O((\delta/k)^{1/2}+\delta^{3/2}).
\end{aligned}
$$

By Lemma 5.3.4 (applied in dimension $N-r$ due to the linear constraint $\vec{R}(\boldsymbol{x}^i,\boldsymbol{\sigma}^0)=\vec{0}$ in $\mathcal{S}_\bullet$), this remaining expectation is given up to $o_N(1)$ error by

$$
\begin{aligned}
&\sum_{s\in\mathscr{S}}\lambda_s\sqrt{(p_+-p_-)v_s^2+2\sum_{s'\in\mathscr{S}}\lambda_{s'}p_-w_{s,s'}^2}\\
&=\sum_{s\in\mathscr{S}}\lambda_s\sqrt{(\Delta\vec{\phi})^s\left((p_+-p_-)\xi^s(\vec{\phi}_-)+p_-\sum_{s'\in\mathscr{S}}\partial_{x_{s'}}\xi^s(\vec{\phi}_-)(\Delta\vec{\phi})^{s'}\right)}.
\end{aligned}
$$

This implies the result. $\square$

We now evaluate $\mathsf{BOGP}_{\mathrm{loc},0}$ by taking a continuous limit of Propositions 5.3.3 and 5.3.5. Fix D, k and $\delta = 6r/D$, and let $(\underline{p}, \underline{\vec{\phi}})$ be δ-dense. We parametrize time by $q_d = \langle \vec{\lambda}, \vec{\phi}_d \rangle$, so in particular $q_0 = \langle \vec{\lambda}, \vec{\phi}_0 \rangle$. Let the functions $\widetilde{p} : [q_0, 1] \to [0,1]$ and $\widetilde{\Phi} : [q_0, 1] \to [0,1]^{\mathscr{S}}$ satisfy

$$\widetilde{p}(q_d) = p_d, \qquad \widetilde{\Phi}(q_d) = \vec{\phi}_d. \tag{5.54}$$

and be linear on each interval $[q_d, q_{d+1}]$. These are piecewise linear approximations of inputs (p, Φ) to the algorithmic functional $\mathbb{A}$. Define

$$A_d^s = \sqrt{(\phi_{d+1}^s - \phi_d^s)\left((p_{d+1} - p_d)\xi^s(\vec{\phi}_d) + p_d \sum_{s' \in \mathscr{S}} \partial_{x_{s'}} \xi^s(\vec{\phi}_d)(\phi_{d+1}^{s'} - \phi_d^{s'})\right)}. \tag{5.55}$$

This term appears in the estimate of $f\left(\vec{\phi}_d, \vec{\phi}_{d+1}; p_{d+1}, p_d; k\right)$ obtained from Proposition 5.3.5.

Lemma 5.3.6. *We have*

$$\left| \sum_{d=0}^{D-1} A_d^s - \int_{q_0}^{q_D} \sqrt{\widetilde{\Phi}_s'(q)(\widetilde{p} \times \xi^s \circ \widetilde{\Phi})'(q)} \, \mathsf{d}q \right| \le C D^{-1/2}$$

for a constant $C > 0$ independent of $D, \underline{p}, \underline{\vec{\phi}}$.

Proof. Until the end, we focus on estimating the difference

$$\Delta_d^s \equiv \left| A_d^s - \int_{q_d}^{q_{d+1}} \sqrt{\widetilde{\Phi}_s'(q)(\widetilde{p} \times \xi^s \circ \widetilde{\Phi})'(q)} \, \mathsf{d}q \right|.$$

Note the general inequality

$$\begin{aligned} \int_{q_d}^{q_{d+1}} \sqrt{a(q)} \cdot |\sqrt{b(q)} - \sqrt{c(q)}| \mathsf{d}q &\le \left(\int_{q_d}^{q_{d+1}} a(q) \mathsf{d}q \right)^{1/2} \cdot \left(\int_{q_d}^{q_{d+1}} \left(\sqrt{b(q)} - \sqrt{c(q)}\right)^2 \mathsf{d}q \right)^{1/2} \\ &\le \left(\int_{q_d}^{q_{d+1}} a(q) \mathsf{d}q \right)^{1/2} \cdot \left(\int_{q_d}^{q_{d+1}} |b(q) - c(q)| \mathsf{d}q \right)^{1/2}. \end{aligned} \tag{5.56}$$

Thus

$$\begin{aligned} \Delta_d^s &= \left| \int_{q_d}^{q_{d+1}} \sqrt{\widetilde{\Phi}_s'(q)} \left(\sqrt{(\widetilde{p} \times \xi^s \circ \widetilde{\Phi})'(q)} - \sqrt{\frac{(p_{d+1} - p_d)\xi^s(\vec{\phi}_d) + p_d \sum_{s' \in \mathscr{S}} \partial_{x_{s'}} \xi^s(\vec{\phi}_d)(\phi_{d+1}^{s'} - \phi_d^{s'})}{q_{d+1} - q_d}} \right) \mathsf{d}q \right| \\ &\le \sqrt{(\phi_{d+1}^s - \phi_d^s) \int_{q_d}^{q_{d+1}} \left| (\widetilde{p} \times \xi^s \circ \widetilde{\Phi})'(q) - \frac{(p_{d+1} - p_d)\xi^s(\vec{\phi}_d) + p_d \sum_{s' \in \mathscr{S}} \partial_{x_{s'}} \xi^s(\vec{\phi}_d)(\phi_{d+1}^{s'} - \phi_d^{s'})}{q_{d+1} - q_d} \right| \mathsf{d}q}. \end{aligned}$$

In the first step we used that $\widetilde{\Phi}'(q) = (\vec{\phi}_{d+1} - \vec{\phi}_d)/(q_{d+1} - q_d)$ by definition, and in the second we used (5.56). Let $(\widetilde{p} \times \xi^s \circ \widetilde{\Phi})'(q_d)$ and $(\widetilde{p} \times \xi^s \circ \widetilde{\Phi})'(q_{d+1})$ denote the right and left derivatives at these points, respectively. The definitions of $\widetilde{p}'$ and $\widetilde{\Phi}'$ imply

$$\frac{(p_{d+1} - p_d)\xi^s(\vec{\phi}_d) + p_d \sum_{s' \in \mathscr{S}} \partial_{x_{s'}} \xi^s(\vec{\phi}_d)(\phi_{d+1}^{s'} - \phi_d^{s'})}{q_{d+1} - q_d} = (\widetilde{p} \times \xi^s \circ \widetilde{\Phi})'(q_d),$$

so in fact

$$\begin{aligned} \Delta_d^s &\le \sqrt{(\phi_{d+1}^s - \phi_d^s) \int_{q_d}^{q_{d+1}} \left| (\widetilde{p} \times \xi^s \circ \widetilde{\Phi})'(q) - (\widetilde{p} \times \xi^s \circ \widetilde{\Phi})'(q_d) \right| \mathsf{d}q} \\ &\le \sqrt{(\phi_{d+1}^s - \phi_d^s)(q_{d+1} - q_d) \left((\widetilde{p} \times \xi^s \circ \widetilde{\Phi})'(q_{d+1}) - (\widetilde{p} \times \xi^s \circ \widetilde{\Phi})'(q_d) \right)}. \end{aligned} \tag{5.57}$$

Let $\nabla\vec{\phi}_d = (\vec{\phi}_{d+1} - \vec{\phi}_d)/(q_{d+1} - q_d)$ be the constant value of $\nabla\widetilde{\Phi}$ on $[q_d, q_{d+1}]$. Then

$$\begin{aligned}(\widetilde{p} \times \xi^s \circ \widetilde{\Phi})'(q_{d+1}) - (\widetilde{p} \times \xi^s \circ \widetilde{\Phi})'(q_d) = \left(\frac{p_{d+1} - p_d}{q_{d+1} - q_d}\right) \Big(\xi^s(\vec{\phi}_{d+1}) - \xi^s(\vec{\phi}_d)\Big) \\ + p_{d+1}\langle \nabla\xi^s(\vec{\phi}_{d+1}), \nabla\vec{\phi}_d\rangle - p_d\langle \nabla\xi^s(\vec{\phi}_d), \nabla\vec{\phi}_d\rangle.\end{aligned}$$

We thus obtain

$$\begin{aligned}&(q_{d+1} - q_d)\Big((\widetilde{p} \times \xi^s \circ \widetilde{\Phi})'(q_{d+1}) - (\widetilde{p} \times \xi^s \circ \widetilde{\Phi})'(q_d)\Big) \\ &= (p_{d+1} - p_d)\Big(\xi^s(\vec{\phi}_{d+1}) - \xi^s(\vec{\phi}_d)\Big) + (q_{d+1} - q_d)(p_{d+1} - p_d)\langle \nabla\xi^s(\vec{\phi}_{d+1}), \nabla\vec{\phi}_d\rangle \\ &\quad + (q_{d+1} - q_d)p_d\langle \nabla\xi^s(\vec{\phi}_{d+1}) - \nabla\xi^s(\vec{\phi}_d), \nabla\vec{\phi}_d\rangle \\ &\le O\left((p_{d+1} - p_d)\Big\|\vec{\phi}_{d+1} - \vec{\phi}_d\Big\|_2 + \Big\|\vec{\phi}_{d+1} - \vec{\phi}_d\Big\|_2^2\right) = O(\delta^2).\end{aligned}$$

Combining with (5.57) gives the estimate $\Delta_d^s = O(\delta)\sqrt{\phi_{d+1}^s - \phi_d^s}$. Summing this over $0 \le d \le D-1$ gives the final estimate

$$\begin{aligned}\left|\sum_{d=0}^{D-1} A_d^s - \int_{q_0}^{q_D} \sqrt{\widetilde{\Phi}_s'(q)(\widetilde{p} \times \xi^s \circ \widetilde{\Phi})'(q)}\ \mathsf{d}q\right| &\le \sum_{d=0}^{D-1} \Delta_d^s \le O(\delta)\sum_{d=0}^{D-1}\sqrt{\phi_{d+1}^s - \phi_d^s} \\ &\le O(\delta\sqrt{D}) = O(D^{-1/2}).\end{aligned}$$

by Cauchy-Schwarz. □

We next show that discretizing any C^1 functions (p, Φ) preserves the value of $\mathbb{A}$.

Lemma 5.3.7. *Suppose $q_0 \in [0,1]$, $p \in \mathbb{I}(q_0, 1)$, and $\Phi \in \mathsf{Adm}(q_0, 1)$. Consider any $\underline{q} = (q_0, \dots, q_D)$ with $q_0 < \cdots < q_D = 1$, such that the $(\underline{p}, \underline{\vec{\phi}})$ defined by $p_d = p(q_d)$ and $\vec{\phi}_d = \Phi(q_d)$ is $6r/D$-dense. Then, for all $s \in \mathscr{S}$,*

$$\left|\int_{q_0}^1 \sqrt{\Phi_s'(q)(p \times \xi^s \circ \Phi)'(q)}\ \mathsf{d}q - \int_{q_0}^1 \sqrt{\widetilde{\Phi}_s'(q)(\widetilde{p} \times \xi^s \circ \widetilde{\Phi})'(q)}\ \mathsf{d}q\right| = o_D(1),$$

where $\widetilde{p}, \widetilde{\Phi}$ are the piecewise linear interpolations defined by (5.54) and $o_D(1)$ is a term tending to 0 as $D \to \infty$ (for fixed (p, Φ)).

Proof. The functions Φ, Φ', p, p' are uniformly continuous because they are continuous on $[q_0, 1]$. So,

$$\|\Phi - \widetilde{\Phi}\|_\infty, \|\Phi' - \widetilde{\Phi}'\|_\infty, \|p - \widetilde{p}\|_\infty, \|p' - \widetilde{p}'\|_\infty = o_D(1).$$

Bounded convergence implies the result. □

Proof of Proposition 5.2.5. By Proposition 5.2.9 it suffices to prove that $\mathsf{BOGP}_{\mathrm{loc},0} = \mathsf{ALG}$. We will separately show $\mathsf{BOGP}_{\mathrm{loc},0} \le \mathsf{ALG}$ and $\mathsf{BOGP}_{\mathrm{loc},0} \ge \mathsf{ALG}$.

We first show $\mathsf{BOGP}_{\mathrm{loc},0} \le \mathsf{ALG}$. Let $\iota > 0$. Let D be sufficiently large, $\varepsilon = D^{-2}$ and $\delta = 6r/D$, and k be sufficiently large depending on D such that the following holds. First, $k \ge k_0(D, \varepsilon, \delta)$ for k_0 defined in Proposition 5.3.3. Second, for some $1/D^2$-separated $\vec{\chi} \in \mathbb{I}(0,1)^{\mathscr{S}}$ and all δ-dense $(\underline{p}, \underline{\vec{\phi}})$ with $\underline{\vec{\phi}} = \vec{\chi}(\underline{p})$, we have

$$\frac{1}{N}\mathbb{E}\sup_{\boldsymbol{\sigma} \in \mathcal{Q}_{\mathrm{loc}}(0)} \mathcal{H}_N(\boldsymbol{\sigma}) \ge \mathsf{BOGP}_{\mathrm{loc},0} - \iota/2.$$

Let $q_d = \langle \vec{\lambda}, \vec{\phi}_d\rangle$. Let $\widetilde{p}, \widetilde{\Phi}$ be the piecewise linear interpolations defined by (5.54) and A_d^s be defined by

(5.55). Then

$$\begin{aligned}
&\left|\frac{1}{N}\sup_{\underline{\boldsymbol{\sigma}}\in\mathcal{Q}_{\rm loc}(0)}\mathcal{H}_N(\underline{\boldsymbol{\sigma}})-\sum_{s\in\mathscr{S}}\left(h_s\lambda_s\sqrt{\widetilde{\Phi}_s(q_0)}+\lambda_s\int_{q_0}^{q_D}\sqrt{\widetilde{\Phi}_s'(q)(p\times\xi^s\circ\Phi)(q)}\,\mathsf{d}q\right)\right|\\
&\le\left|\frac{1}{N}\sup_{\underline{\boldsymbol{\sigma}}\in\mathcal{Q}_{\rm loc}(0)}\mathcal{H}_N(\underline{\boldsymbol{\sigma}})-\sum_{s\in\mathscr{S}}\left(h_s\lambda_s\sqrt{\phi_0^s}+\sum_{d=0}^{D-1}f\left(\vec{\phi}_d,\vec{\phi}_{d+1};p_d,p_{d+1};k\right)\right)\right|\\
&\quad+\sum_{d=0}^{D-1}\left|f\left(\vec{\phi}_d,\vec{\phi}_{d+1};p_d,p_{d+1};k\right)-\sum_{s\in\mathscr{S}}\lambda_sA_d^s\right|+\sum_{s\in\mathscr{S}}\lambda_s\left|\sum_{d=0}^{D-1}A_d^s-\int_{q_0}^{q_D}\sqrt{\widetilde{\Phi}_s'(q)(p\times\xi^s\circ\Phi)(q)}\,\mathsf{d}q\right|.
\end{aligned}$$

By Propositions 5.3.3, 5.3.5 and Lemma 5.3.6, on an event with probability $1-e^{-cN}$ this is bounded by

$$2D\varepsilon+O(D\delta^{3/2}+D(\delta/k)^{1/2})+O(D^{-1/2})+o_N(1)=O(D^{-1/2}+(D/k)^{1/2})+o_N(1)\le\iota/4,$$

for sufficiently large N,D,k. Because $\frac{1}{N}\sup_{\underline{\boldsymbol{\sigma}}\in\mathcal{Q}_{\rm loc}(0)}\mathcal{H}_N(\underline{\boldsymbol{\sigma}})$ is subgaussian with fluctuations $O(N^{-1/2})$ by Lemma 5.2.11, the contributions of the complement of this event are $o_N(1)$, and so

$$\left|\frac{1}{N}\mathbb{E}\sup_{\underline{\boldsymbol{\sigma}}\in\mathcal{Q}_{\rm loc}(0)}\mathcal{H}_N(\underline{\boldsymbol{\sigma}})-\sum_{s\in\mathscr{S}}\left(h_s\lambda_s\sqrt{\widetilde{\Phi}_s(q_0)}+\lambda_s\int_{q_0}^{q_D}\sqrt{\widetilde{\Phi}_s'(q)(p\times\xi^s\circ\Phi)(q)}\,\mathsf{d}q\right)\right|\le\iota/4. \tag{5.58}$$

Let $p\in\mathbb{I}(q_0,1)$ and $\Phi\in\mathsf{Adm}(q_0,1)$ approximate the piecewise linear functions $(\widetilde{p},\widetilde{\Phi})$ on $[q_0,q_D]$, in the sense that

$$\left|\sum_{s\in\mathscr{S}}\lambda_s\int_{q_0}^{q_D}\left(\sqrt{\widetilde{\Phi}_s'(q)(\widetilde{p}\times\xi^s\circ\widetilde{\Phi})'(q)}-\sqrt{\Phi_s'(q)(p\times\xi^s\circ\Phi)'(q)}\right)\,\mathsf{d}q\right|\le\iota/4. \tag{5.59}$$

It is clear that such p,Φ exist. Thus

$$\mathsf{BOGP}_{\rm loc,0}\le\mathbb{A}(p,\Phi;q_0)-\iota\le\mathsf{ALG}-\iota.$$

Since ι was arbitrary, we conclude $\mathsf{BOGP}_{\rm loc,0}\le\mathsf{ALG}$.

Next, we will show $\mathsf{BOGP}_{\rm loc,0}\ge\mathsf{ALG}$. Let $\iota>0$, and let D be sufficiently large and k be sufficiently large depending on D. There exist $q_0\in[0,1]$, $p\in\mathbb{I}(q_0,1)$, and $\Phi\in\mathsf{Adm}(q_0,1)$ such that

$$\mathbb{A}(p,\Phi;q_0)\ge\mathsf{ALG}-\iota/2.$$

By replacing Φ with $(1-D^{-2})\Phi+D^{-2}\vec{1}$ we may assume $\Phi(q_0)\succeq\vec{1}/D^2$, as this replacement affects the left-hand side by $o_D(1)$. Similarly, by replacing $p(q)$ with $(1-D^{-1})p(q)+D^{-1}q$, we may assume p is strictly increasing. We choose $\vec{\chi}=\Phi\circ p^{-1}$, which is $1/D^2$-separated. Consider any $\underline{q}=(q_0,q_1,\dots,q_D)$ with $q_0<q_1<\cdots<q_D=1$ such that for $p_d=p(q_d)$, $\vec{\phi}_d=\Phi(q_d)$, the pair $(\underline{p},\underline{\vec{\phi}})$ is $6r/D$-dense. Similarly to above, we have (5.58) for sufficiently large N,D,k. By Lemma 5.3.7, (5.59) holds for D sufficiently large. This implies

$$\mathbb{A}(p,\Phi;q_0)\le\mathsf{BOGP}_{\rm loc,0}+\iota/2,$$

and so $\mathsf{ALG}\le\mathsf{BOGP}_{\rm loc,0}+\iota$. Because ι was arbitrary, we have $\mathsf{ALG}\le\mathsf{BOGP}_{\rm loc,0}$. □

5.4 Optimization of the algorithmic variational principle

In this section we will prove Propositions 5.1.9 and 5.1.11 and Theorem 5.1.12. Throughout this section we assume Assumption 5.1.3 except where stated.

To ensure a priori existence of a maximizer in (5.7), we work in the following compact space which removes the constraint that p and Φ are continuously differentiable.

Definition 5.4.1. The space $\mathcal{M}$ consists of all triples (p,Φ,q_0) such that:

- $q_0\in[0,1]$.

- $p : [q_0, 1] \to [0, 1]$ is non-decreasing and right-continuous (we write $p \in \widehat{\mathbb{I}}(q_0, 1)$).
- $\Phi = (\Phi_s)_{s\in\mathscr{S}}$ consists of r non-decreasing functions $\Phi_s : [q_0, 1] \to [0, 1]$ satisfying admissibility (5.5) (we write $\Phi \in \widehat{\mathsf{Adm}}(q_0, 1)$).

Because we assume almost no regularity for elements of $\mathcal{M}$, we formally define the integral in (5.6) as follows. Since $(p\times\xi^s\circ\Phi)$ is a bounded non-decreasing function, it has a positive-measure-valued distributional derivative

$$(p \times \xi^s \circ \Phi)'(q)\, \mathsf{d}q = f(q)\ \mathsf{d}q + \mathsf{d}\mu(q) \tag{5.60}$$

where $f \in L^1([q_0, 1])$ and μ is an atomic-plus-singular measure supported in $[q_0, 1]$. Moreover, (5.5) implies Φ_s is λ_s^{-1}-Lipschitz, hence has distributional derivative $\Phi_s' \in L^\infty([q_0, 1])$.

Definition 5.4.2. For $(p, \Phi, q_0) \in \mathcal{M}$, define

$$\widehat{\mathsf{ALG}} \equiv \sup_{(p,\Phi,q_0)\in\mathcal{M}} \mathbb{A}(p, \Phi; q_0). \tag{5.61}$$

where the second term of $\mathbb{A}$ is given (with f as in (5.60)) by:

$$\int_{q_0}^{1} \sqrt{\Phi_s'(q)(p \times \xi^s \circ \Phi)'(q)}\ \mathsf{d}q = \int_{q_0}^{1} \sqrt{\Phi_s'(q)f(q)}\ \mathsf{d}q. \tag{5.62}$$

Informally, the reason that μ from (5.60) disappears is that in a suitable approximating limit, "the square-root of a Dirac mass has L^1 norm zero". Formally, one can take (5.62) to be the definition of the left-hand side.

It will follow from our results in this section that for non-degenerate ξ, all maximizers to the extended variational problem are continuously differentiable on $[q_0, 1]$. The equality $\mathsf{ALG} = \widehat{\mathsf{ALG}}$ follows in general since both are continuous in $(\xi, \check{h})$.

Remark 5.4.3. A related (for the most part, simpler) variational problem was considered in [DZ95]. There, after showing existence and other basic properties, the general result [Ces12, Theorem 5.1] was used to derive an ordinary differential equation [DZ95, Theorem 4] for the optimal Φ. The same general result applies in our setting, and essentially yields Proposition 5.4.17, assuming certain functions f_s defined below in (5.66) are absolutely continuous for all $s \in \mathscr{S}$. However the only way we could establish absolute continuity of f_s was by going through the full proof of Proposition 5.4.17.

5.4.1 Linear algebraic and analytic preliminaries

We first prove Corollary 5.4.5 below, an equivalent characterization of (super, strict sub)-solvability.

Proposition 5.4.4. *Let $M \in \mathbb{R}^{\mathscr{S}\times\mathscr{S}}$ be diagonally signed. Then*

$$\Lambda(M) = \sup_{\vec{v}\in\mathbb{R}^{\mathscr{S}}_{>0}} \min_{s\in\mathscr{S}} \frac{(M\vec{v})_s}{v_s} \tag{5.63}$$

equals the smallest eigenvalue $\lambda_{\min}(M)$ of M.

Proof. Let $\vec{w}$ be a (unit) minimal eigenvector of M. Note that

$$\vec{w}^\top M\vec{w} = \sum_{s,s'\in\mathscr{S}} M_{s,s'} w_s w_{s'} \geq \sum_{s,s'\in\mathscr{S}} M_{s,s'} |w_s||w_{s'}|.$$

Since $\vec{w}$ minimizes $\vec{w}^\top M\vec{w}$, all entries of $\vec{w}$ are the same sign. We may thus assume $\vec{w} \in \mathbb{R}^{\mathscr{S}}_{\geq 0}$. Moreover, if $w_s = 0$ for any s, then $(M\vec{w})_s < 0$ so $\vec{w}$ is not an eigenvector; thus $\vec{w} \in \mathbb{R}^{\mathscr{S}}_{>0}$. Because $M\vec{w} = \lambda_{\min}(M)\vec{w}$, clearly $\Lambda(M) \geq \lambda_{\min}(M)$. For any other $\vec{v} \in \mathbb{R}^{\mathscr{S}}_{>0}$,

$$\min_{s\in\mathscr{S}} \frac{(M\vec{v})_s}{v_s} \leq \langle \vec{w}, \vec{v}\rangle^{-1} \sum_{s\in\mathscr{S}} w_s v_s \cdot \frac{(M\vec{v})_s}{v_s} = \frac{\langle \vec{w}, M\vec{v}\rangle}{\langle \vec{w}, \vec{v}\rangle} = \frac{\langle M\vec{w}, \vec{v}\rangle}{\langle \vec{w}, \vec{v}\rangle} = \lambda_{\min}(M),$$

so $\Lambda(M) \leq \lambda_{\min}(M)$. Thus $\Lambda(M) = \lambda_{\min}(M)$. $\square$

Corollary 5.4.5. *For $\vec{x} \in (0,1]^{\mathscr{S}} \cup \{\vec{0}\}$ define*

$$M^*(\vec{x}) = \operatorname{diag}\left((\xi^s(\vec{x}) + h_s^2)_{s\in\mathscr{S}}\right) - \left(x_s \partial_{x_{s'}} \xi^s(\vec{x})\right)_{s,s'\in\mathscr{S}}.$$

Then $\vec{x}$ is super-solvable (resp. solvable, strictly sub-solvable) if and only if $\Lambda(M^*(\vec{x})) \geq 0$ (resp. $= 0$, < 0).

Proof. Suppose first $\vec{x} \in (0,1]^{\mathscr{S}}$. By Proposition 5.4.4, $\vec{x}$ is super-solvable (resp. solvable, strictly sub-solvable) if and only if $\Lambda(M^*_{\mathsf{sym}}(\vec{x})) \geq 0$ (resp. $= 0$, < 0). Note that

$$M^*(\vec{x}) = \operatorname{diag}\left((\lambda_s x_s)_{s\in\mathscr{S}}\right) M^*_{\mathsf{sym}}(\vec{x}), \tag{5.64}$$

so $\Lambda(M^*(\vec{x}))$ has the same sign as $\Lambda(M^*_{\mathsf{sym}}(\vec{x}))$, as desired. If $\vec{x} = \vec{0}$, then clearly $\Lambda(M^*(\vec{x})) \geq 0$ with equality at $\check{h} = \vec{0}$, which agrees with the convention from Definition 5.1.5. □

The following proposition is clear.

Proposition 5.4.6. *Let Λ be as in (5.63) and let $M \in \mathbb{R}^{r\times r}_{\geq 0}$ (not necessarily diagonally signed). Then $\Lambda(M)$ is non-negative and locally bounded. Moreover if for some $c \in \mathbb{R}$ we have $M'_{s,s'} \geq M_{s,s'} + c \cdot 1_{s=s'}$ for all s, s', then $\Lambda(M') \geq \Lambda(M) + c$.*

Many perturbation arguments used to establish regularity rely on the following basic fact.

Proposition 5.4.7 ([Rud87, Theorem 7.7])**.** *For $f \in L^1([0,1])$, almost all $x \in [0,1]$ are **Lebesgue points**:*

$$\lim_{\varepsilon\to 0} \frac{1}{2\varepsilon} \int_{x-\varepsilon}^{x+\varepsilon} |f(y) - f(x)| \, \mathsf{d}y = 0.$$

The next fact ensures that Lipschitz ordinary differential equations are well-posed (even if they are only required to hold almost everywhere).

Proposition 5.4.8 ([Rou13, Theorem 1.45, Part (ii)])**.** *Suppose $Y_1, Y_2 : [0,1] \to \mathbb{R}^d$ are each absolutely continuous with $Y_1(0) = Y_2(0)$ and solve the ODE $Y_i'(q) = F(Y_i(q))$ at almost all q for $F : \mathbb{R}^d \to \mathbb{R}^d$ Lipschitz. Then Y_1, Y_2 agree and solve the ODE for all q.*

5.4.2 A priori regularity of maximizers

We first show that for the optimization problem (5.7), admissibility (5.5) is just a convenient choice of normalization. This makes variational arguments more convenient because we do not need to worry about preserving admissibility of Φ under perturbations. Let $\widetilde{\mathbb{I}}(q_0,1) \subseteq \widehat{\mathbb{I}}(q_0,1)$ be the set of non-decreasing and Lipschitz functions $f : [q_0,1] \to [0,1]$ with no explicit bound on the Lipschitz constant and with $f(1) = 1$. Note that the algorithmic functional $\mathbb{A}$ (5.6) remains well-defined for $\Phi \in \widetilde{\mathbb{I}}(q_0,1)^{\mathscr{S}}$.

Lemma 5.4.9. *We have that*

$$\widehat{\mathsf{ALG}} = \sup_{q_0\in[0,1]} \sup_{\substack{p\in\widehat{\mathbb{I}}(q_0,1)\\ \Phi\in\widetilde{\mathbb{I}}(q_0,1)^{\mathscr{S}}}} \mathbb{A}(p,\Phi;q_0). \tag{5.65}$$

Proof. Let $\widehat{\mathsf{ALG}}'$ be the right-hand side of (5.65). We will show that $\widehat{\mathsf{ALG}} \geq \widehat{\mathsf{ALG}}'$ (the opposite implication being trivial).

Consider any $q_0 \in [0,1]$, $p \in \widehat{\mathbb{I}}(q_0,1)$, and $\Phi \in \widetilde{\mathbb{I}}(q_0,1)^{\mathscr{S}}$. For small $\delta > 0$, consider

$$\Phi_\delta(q) = \delta q \vec{1} + (1-\delta)\Phi(q)$$

and let $\alpha(q) = \langle \vec{\lambda}, \Phi_\delta(q)\rangle$, so $\alpha'(q) \geq \delta$. Thus α^{-1} exists and is δ^{-1}-Lipschitz. Consider $(\widetilde{p}, \widetilde{\Phi}, \widetilde{q}_0)$ given by

$$\widetilde{p}(q) = p(\alpha^{-1}(q)), \quad \widetilde{\Phi}(q) = \Phi(\alpha^{-1}(q)), \quad \widetilde{q}_0 = \alpha(q_0).$$

By construction, $\widetilde{\Phi} \in \widehat{\mathsf{Adm}}(\widetilde{q}_0, 1)$. By the chain rule, $\mathbb{A}(\widetilde{p}, \widetilde{\Phi}; \widetilde{q}_0) = \mathbb{A}(p, \Phi_\delta; q_0)$. Thus

$$\widehat{\mathsf{ALG}} \geq \limsup_{\delta\downarrow 0} \mathbb{A}(\widetilde{p}, \widetilde{\Phi}, \widetilde{q}_0) = \limsup_{\delta\downarrow 0} \mathbb{A}(p, \Phi_\delta; q_0) \geq \mathbb{A}(p, \Phi; q_0).$$

Since p, Φ, q_0 were arbitrary the conclusion follows. □

A routine compactness argument given in Appendix 5.C.1 yields the following.

Proposition 5.4.10. *There exists a maximizer $(p, \Phi, q_0) \in \mathcal{M}$ for $\mathbb{A}$ and $\mathbb{A}(p, \Phi; q_0) < \infty$.*

From now on, we let $(p, \Phi, q_0) \in \mathcal{M}$ denote any maximizer and study the behavior of (p, Φ, q_0). While almost no regularity on (p, Φ) is assumed, it is possible to establish a priori regularity using variational arguments. We defer the proofs of the following two propositions to Appendix 5.C.2. Proposition 5.4.11 implies that the discussion following (5.7) is not necessary to define $\mathbb{A}(p, \Phi; q_0)$.

Proposition 5.4.11. *The functions p, Φ are continuously differentiable on $[q_0+\varepsilon, 1]$ for any $\varepsilon > 0$. Moreover, there exists $L > 0$ (possibly depending on $(p, \Phi; q_0)$ as well as ξ) such that $L^{-1}\vec{1} \preceq \Phi'(q) \preceq L\vec{1}$ for almost all $q \in (q_0, 1]$.*

Proposition 5.4.12. *The function p satisfies $p(q) > 0$ for all $q > q_0$, $p(1) = 1$, and $p(q_0) = 0$ if $q_0 > 0$.*

Throughout the next subsection we will use $\varepsilon > 0$ as in Proposition 5.4.11. Later we slightly improve the result of Proposition 5.4.11 to continuity on $[q_0, 1]$ using more detailed properties of the maximizers.

5.4.3 Identification of root-finding and tree-descending phases

In this subsection we will prove the following result. Recall that the Sobolev space $W^{2,\infty}([q_0 + \varepsilon, 1])$ consists of C^1 functions with Lipschitz derivative on the interval.

Proposition 5.4.13. *The restrictions of p and Φ_s, for all $s \in \mathscr{S}$, lie in the space $W^{2,\infty}([q_0 + \varepsilon, 1])$ for any $\varepsilon > 0$. There exists $q_1 \in [q_0, 1]$ such that the following holds.*

(a) On $[q_0, q_1]$, $p' > 0$ almost everywhere and the quantities $\frac{\Phi_s'(q)}{(p\times\xi^s\circ\Phi)'(q)}$ are constant. Moreover $p(q_1) = 1$.

(b) On $[q_1, 1]$, the ODE (5.15) is satisfied for all $s, s' \in \mathscr{S}$ almost everywhere and $p = 1$.

We begin with a result on diagonally dominant matrices. Variants especially with $\varepsilon = 0$ have been used many times, see e.g. [Tau49]. Related linear algebraic statements will appear later in Lemmas 5.4.18 and 5.4.20 as, roughly speaking, r-dimensional analogs of monotonicity.

Lemma 5.4.14. *Let $A = (a_{i,j})_{i,j\in[r]} \in \mathbb{R}^{r\times r}$ satisfy $a_{i,i} > 0$ and $a_{i,j} < 0$ for all $i \neq j$.*

(a) If $\sum_{j=1}^r a_{i,j} = 0$ for all $i \in [r]$, then all solutions $\vec{v} \in \mathbb{R}^r$ to $A\vec{v} \preceq \varepsilon\vec{1}$ satisfy $|v_i - v_j| \le \varepsilon/a_{\min}$ for all i, j, where $a_{\min} = \min_{i\neq j} |a_{i,j}|$.

(b) If $\sum_{j=1}^r a_{i,j} \ge d_{\min} > 0$ for all $i \in [r]$, then all solutions $\vec{v} \in \mathbb{R}^r$ to $\|A\vec{v}\|_\infty \le \varepsilon$ satisfy $\|v_i\|_\infty \le \varepsilon/d_{\min}$.

Proof of Lemma 5.4.14. Assume without loss of generality that $v_1 \ge v_s$ for all s. If $\sum_{j=1}^r a_{i,j} = 0$ for all $i \in [r]$, then

$$\varepsilon \ge (A\vec{v})_1 = a_{1,1}v_1 + \sum_{j=2}^r a_{1,j}v_i = \sum_{j=2}^r |a_{1,j}|(v_1 - v_j) \ge a_{\min}(v_1 - v_i)$$

for all $i \ge 2$. Thus $v_1 - v_i \le \varepsilon/a_{\min}$, proving the first part. For the second part, we will first show $v_1 \le \varepsilon/d_{\min}$. If $v_1 < 0$ there is nothing to prove, and otherwise

$$\varepsilon \ge (A\vec{v})_1 = a_{1,1}v_1 + \sum_{j=2}^r a_{1,j}v_j \ge \left(a_{1,1} - \sum_{j=2}^r a_{1,j}\right) v_1 \ge d_{\min}v_1.$$

So $v_1 \le \varepsilon/d_{\min}$, as claimed. Finally, note that if $\|A\vec{v}\|_\infty \le \varepsilon$, the same is true for $-\vec{v}$. By the same argument we find the largest entry of $-\vec{v}$ is at most $\varepsilon/d_{\min}$. This implies the second part. $\square$

Corollary 5.4.15. *Let $A = (a_{i,j})_{i,j\in[r]} \in \mathbb{R}^{r\times r}$ satisfy $a_{i,i} > 0$ and $a_{i,j} < 0$ for all $i \neq j$. If $\sum_{j=1}^r a_{i,j} > 0$ for all $i \in [r]$, then the only solution to $A\vec{v} = \vec{0}$ is $\vec{v} = \vec{0}$.*

Proof. Apply Lemma 5.4.14(b) with $\varepsilon = 0$. $\square$

To establish additional regularity we use the following fact on distributional derivatives.

Lemma 5.4.16 (See e.g. [Zie12, Theorem 2.2.1])**.** *If $A, B \in L^\infty([q_0, 1])$ satisfy*

$$\int_{q_0}^{1} A(q)\psi(q) + B(q)\psi'(q) \; \mathsf{d}q = 0$$

for all $\psi \in C_c^\infty((q_0, 1); \mathbb{R})$, then there exists $C \in \mathbb{R}$ such that for all $q \in [q_0, 1]$,

$$B(q) = \int_{q_0}^{q} A(t) \; \mathsf{d}t + C.$$

We will make use of the functions

$$f_s(q) = \sqrt{\frac{\Phi_s'(q)}{(p \times \xi^s \circ \Phi)'(q)}} \tag{5.66}$$

Note that Propositions 5.4.11 and 5.4.12 imply f_s is continuous on $[q_0 + \varepsilon, 1]$.

Proposition 5.4.17. *The functions f_s are Lipschitz on $[q_0 + \varepsilon, 1]$. Thus (recall Proposition 5.4.11) the functions*

$$\Psi_s(q) = f_s'(q)/\Phi_s'(q) \tag{5.67}$$

are measurable and locally bounded on $(q_0, 1]$. Moreover for almost all $q \in (q_0, 1]$, the following holds:

$$\Psi_1(q) = \cdots = \Psi_r(q), \tag{5.68}$$

and furthermore this common value is 0 if $p'(q) > 0$.

Proof. Let $\psi \in C_c^\infty((q_0, 1); \mathbb{R})$. Consider the perturbation

$$\widetilde{\Phi}_1(q) = \Phi_1(q) + \delta\psi(q),$$

and let $\widetilde{\Phi}_s(q) = \Phi_s(q)$ for $s \neq 1$. By Proposition 5.4.11, $\widetilde{\Phi}$ remains coordinate-wise non-decreasing and Lipschitz for small positive and negative δ. Although $\widetilde{\Phi} \notin \widehat{\mathsf{Adm}}(q_0, 1)$, recalling Lemma 5.4.9 we nonetheless have $\mathbb{A}(p, \widetilde{\Phi}; q_0) \leq \mathbb{A}(p, \Phi; q_0)$. Thus,

$$F_1 \equiv \frac{\mathsf{d}}{\mathsf{d}\delta} \mathbb{A}(p, \widetilde{\Phi}; q_0)\Big|_{\delta=0} = 0.$$

We now calculate F_1. Note that

$$\frac{\mathsf{d}}{\mathsf{d}\delta}(p \times \xi^s \circ \widetilde{\Phi})'(q)\Big|_{\delta=0} = (p\psi \times \partial_{x_1}\xi^s \circ \Phi)'(q) = \frac{\lambda_1}{\lambda_s}(p\psi \times \partial_{x_s}\xi^1 \circ \Phi)'(q). \tag{5.69}$$

So,

$$\begin{aligned} 0 = \frac{2}{\lambda_1} F_1 &= \int_{q_0}^{1} f_1(q)^{-1}\psi'(q) \; \mathsf{d}q + \sum_{s \in \mathscr{S}} \int_{q_0}^{1} f_s(q)(p\psi \times \partial_{x_s}\xi^1 \circ \Phi)'(q) \; \mathsf{d}q \\ &= \int_{q_0}^{1} A_1(q)\psi(q) + B_1(q)\psi'(q) \; \mathsf{d}q \end{aligned}$$

where

$$A_1(q) \equiv \sum_{s \in \mathscr{S}} f_s(q)(p \times \partial_{x_s}\xi^1 \circ \Phi)'(q), \quad B_1(q) \equiv f_1(q)^{-1} + \sum_{s \in \mathscr{S}} f_s(q)(p \times \partial_{x_s}\xi^1 \circ \Phi)(q).$$

By Proposition 5.4.11, for all $\varepsilon > 0$ $A_1(q)$ and $B_1(q)$ are bounded for $q \in [q_0+\varepsilon, 1]$. Lemma 5.4.16 implies that $B_1(q)$ is absolutely continuous and $B_1'(q) = A_1(q)$ for all $q \in (q_0, 1]$. In fact by Proposition 5.4.11, A_1 is bounded and continuous on $[q_0+\varepsilon, 1]$, so $B_1 \in C^1([q_0+\varepsilon, 1])$ (for all $\varepsilon > 0$).

Fix $q \in (q_0, 1]$. For $\iota \in \mathbb{R}$ with $|\iota|$ small, let $\Delta_s^\iota = f_s(q+\iota) - f_s(q)$. By Proposition 5.4.11 all f_s are continuous, so $\Delta_s^\iota = o(1)$; here are below we use $o(\cdot)$ for limits as $\iota \to 0$. Thus,

$$\begin{aligned} B_1(q+\iota) - B_1(q) = &\frac{1}{f_1(q)+\Delta_1^\iota} - \frac{1}{f_1(q)} + \sum_{s \in \mathscr{S}} \Delta_s^\iota \cdot (p \times \partial_{x_s} \xi^1 \circ \Phi)(q) \\ &+ \sum_{s\in\mathscr{S}} f_s(q+\iota)\left((p \times \partial_{x_s}\xi^1 \circ \Phi)(q+\iota) - (p \times \partial_{x_s}\xi^1\circ\Phi)(q)\right). \end{aligned}$$

Since $(p \times \partial_{x_s}\xi^1 \circ \Phi)$ is differentiable and f_s is continuous,

$$\begin{aligned} \sum_{s\in\mathscr{S}} f_s(q+\iota)\cdot\left((p \times \partial_{x_s}\xi^1 \circ \Phi)(q+\iota) - (p \times \partial_{x_s}\xi^1\circ\Phi)(q)\right) &= \iota \sum_{s\in\mathscr{S}} f_s(q)(p \times \partial_{x_s}\xi^1\circ\Phi)'(q) + o(|\iota|) \\ &= \iota A_1(q) + o(|\iota|). \end{aligned}$$

Moreover,

$$\begin{aligned} \frac{1}{f_1(q)+\Delta_1^\iota} - \frac{1}{f_1(q)} &= -\frac{\Delta_1^\iota}{f_1(q)(f_1(q)+\Delta_1^\iota)} = \frac{(\Delta_1^\iota)^2}{f_1(q)^2(f_1(q)+\Delta_1^\iota)} - \frac{\Delta_1^\iota}{f_1(q)^2} \\ &= \frac{(\Delta_1^\iota)^2}{f_1(q)^2(f_1(q)+\Delta_1^\iota)} - \frac{\Delta_1^\iota}{\Phi_1'(q)}\left(p'(q)(\xi^1\circ\Phi)(q) + \sum_{s\in\mathscr{S}}(p\times\partial_{x_s}\xi^1\circ\Phi)(q)\Phi_s'(q)\right). \end{aligned}$$

We also have $B_1(q+\iota) - B_1(q) = A_1\iota + o(|\iota|)$ (recall A_1 is continuous). Thus

$$\sum_{s\in\mathscr{S}}(p\times\partial_{x_s}\xi^1\circ\Phi)(q)\Phi_s'(q)\left[\frac{\Delta_1^\iota}{\Phi_1'(q)} - \frac{\Delta_s^\iota}{\Phi_s'(q)}\right] + p'(q)(\xi^1\circ\Phi)(q)\frac{\Delta_1^\iota}{\Phi_1'(q)} - \frac{(\Delta_1^\iota)^2}{f_1(q)^2(f_1(q)+\Delta_1^\iota)} = o(|\iota|). \quad (5.70)$$

We get similar equations from perturbing any Φ_s instead of Φ_1. If $p'(q) > 0$, then we can write the last two terms on the left-hand side of (5.70) as

$$\frac{\Delta_1^\iota}{\Phi_1'(q)}\left(p'(q)(\xi^1\circ\Phi)(q) - \frac{\Delta_1^\iota\Phi_1'(q)}{f_1(q)^2(f_1(q)+\Delta_1^\iota)}\right) = \frac{\Delta_1^\iota}{\Phi_1'(q)}\left(p'(q)(\xi^1\circ\Phi)(q) + o(1)\right).$$

Then, (5.70) and its analogs form a linear system in variables $x_s \equiv \Delta_s^\iota/\Phi_s'(q)$ with all row sums positive. (E.g. in (5.70), the first term gives zero coefficient sum so the total coefficient sum is just $\left(p'(q)(\xi^1\circ\Phi)(q) + o(1)\right) > 0$.) Moreover the diagonal coefficients of this system are e.g.

$$a_{1,1} = \left(p'(q)(\xi^1\circ\Phi)(q) + o(1)\right) + \sum_{s\in\mathscr{S}}(p\times\partial_{x_s}\xi^1\circ\Phi)(q)\Phi_s'(q) > 0$$

while the off-diagonal coefficients are e.g.

$$a_{1,s} = -(p\times\partial_{x_s}\xi^1\circ\Phi)(q)\Phi_s'(q) < 0.$$

Applying Lemma 5.4.14(b), we obtain

$$|\Delta_s^\iota/\Phi_s'(q)| = o(|\iota|)$$

for all $s \in \mathscr{S}$. Taking $\iota \to 0$ we conclude that $f_s'(q)$ is well-defined and equals 0. This implies the conclusion for $p'(q) > 0$.

Otherwise $p'(q) = 0$, and (5.70) implies that

$$\sum_{s\in\mathscr{S}}(p\times\partial_{x_s}\xi^1\circ\Phi)(q)\Phi_s'(q)\left[\frac{\Delta_1^\iota}{\Phi_1'(q)} - \frac{\Delta_s^\iota}{\Phi_s'(q)}\right] \geq -o(|\iota|)$$

and analogously with any $s \in \mathscr{S}$ in place of 1. This is a linear system of inequalities in variables $-\Delta_s^\iota/\Phi_s'(q)$, so Lemma 5.4.14(a) implies that

$$\left| \frac{\Delta_s^\iota}{\Phi_s'(q)} - \frac{\Delta_{s'}^\iota}{\Phi_{s'}'(q)} \right| \le o(|\iota|) \tag{5.71}$$

for all $s, s' \in \mathscr{S}$. The result now follows if we find a constant $C = C(\varepsilon)$ such that

$$\max_{s \in \mathscr{S}} |\Delta_s^\iota/\Phi_s'(q)| \le C|\iota| \tag{5.72}$$

for all sufficiently small ι, and $q \in [q_0 + \varepsilon, 1]$. Indeed, this would imply by Proposition 5.4.11 that f_s is Lipschitz on $[q_0 + \varepsilon, 1]$. It would then follow that $\Psi \in L^\infty([q_0 + \varepsilon, 1])$, and we would conclude from (5.71) that $\Psi_1 = \cdots = \Psi_r$ almost everywhere.

Since p and $\partial_{x_s}\xi^1$ are differentiable and $p'(q) = 0$, we have $p(q + \iota) = p(q) + o(|\iota|)$ and $\partial_{x_s}\xi^1(q + \iota) = \partial_{x_s}\xi^1(q) + O(|\iota|)$. Suppose first that $\iota > 0$. Using $p'(q) = 0$ and $p'(q + \iota) \ge 0$, we find

$$\begin{aligned} \Delta_1^\iota &\le \sqrt{\frac{\Phi_1'(q+\iota)}{p(q+\iota)(\xi^1 \circ \Phi)'(q+\iota)}} - \sqrt{\frac{\Phi_1'(q)}{p(q)(\xi^1 \circ \Phi)'(q)}} \\ &= \frac{1}{\sqrt{p(q)}} \left(\sqrt{\frac{\Phi_1'(q+\iota)}{\sum_{s\in\mathscr{S}} (\partial_{x_s}\xi^1 \circ \Phi)(q)\Phi_s'(q+\iota)}} - \sqrt{\frac{\Phi_1'(q)}{\sum_{s\in\mathscr{S}} (\partial_{x_s}\xi^1 \circ \Phi)(q)\Phi_s'(q)}} \right) + O(\iota). \end{aligned} \tag{5.73}$$

The hidden constants are uniform on any interval $[q_0 + \varepsilon, 1]$. Analogous bounds hold for Δ_s^ι. We claim that we cannot have

$$\frac{\Phi_s'(q+\iota)}{\sum_{s'\in\mathscr{S}} (\partial_{x_{s'}}\xi^s \circ \Phi)(q)\Phi_{s'}'(q+\iota)} > \frac{\Phi_s'(q)}{\sum_{s'\in\mathscr{S}} (\partial_{x_{s'}}\xi^s \circ \Phi)(q)\Phi_{s'}'(q)} \tag{5.74}$$

for all $s \in \mathscr{S}$. Indeed, suppose this holds and let

$$\begin{aligned} b_s &= \frac{\sum_{s'\in\mathscr{S}} (\partial_{x_{s'}}\xi^s \circ \Phi)(q)\Phi_{s'}'(q)}{\Phi_s'(q)}, \\ b_s' &= \frac{\sum_{s'\in\mathscr{S}} (\partial_{x_{s'}}\xi^s \circ \Phi)(q)\Phi_{s'}'(q+\iota)}{\Phi_s'(q+\iota)}, \end{aligned}$$

so $b_s' < b_s$. The linear system given by

$$b_s'\Phi_s'(q+\iota)x_s - \left(\sum_{s'\in\mathscr{S}} (\partial_{x_{s'}}\xi^s \circ \Phi)(q)\Phi_{s'}'(q+\iota)x_{s'} \right) = 0$$

for all $s \in \mathscr{S}$ has solution $\vec{x} = \vec{1}$, and thus has row sums zero. The linear system given by

$$b_s\Phi_s'(q+\iota)x_s - \left(\sum_{s'\in\mathscr{S}} (\partial_{x_{s'}}\xi^s \circ \Phi)(q)\Phi_{s'}'(q+\iota)x_{s'} \right) = 0$$

has solution $x_s = \Phi_s'(q)/\Phi_s'(q+\iota)$. However, by Corollary 5.4.15 its only solution is $\vec{x} = \vec{0}$, contradiction. Thus (5.74) does not hold for all $s \in \mathscr{S}$. Assume without loss of generality (5.74) does not hold for $s = 1$. Then, $\Delta_1^\iota \le O(\iota)$. In conjunction with (5.71), this implies $\max_{s\in\mathscr{S}} \Delta_s^\iota/\Phi_s'(q) \le C\iota$.

For the matching lower bound, first consider the case $p'(q+\iota) = 0$. In this case, the inequality in (5.73) is an equality. We similarly cannot have

$$\frac{\Phi_s'(q+\iota)}{\sum_{s'\in\mathscr{S}} (\partial_{x_{s'}}\xi^s \circ \Phi)(q)\Phi_{s'}'(q+\iota)} < \frac{\Phi_s'(q)}{\sum_{s'\in\mathscr{S}} (\partial_{x_{s'}}\xi^s \circ \Phi)(q)\Phi_{s'}'(q)}$$

for all $s \in \mathscr{S}$, so the same argument implies $\min_{s\in\mathscr{S}} \Delta_s^\iota/\Phi_s'(q) \ge -C\iota$, which implies (5.72). Otherwise assume $p'(q+\iota) > 0$. Let $\iota_1 \in (0, \iota/2)$ be small enough that

$$p'(q') \ge \frac{1}{2}p'(q+\iota) \quad \text{for all } q' \in [q + \iota - \iota_1, q + \iota] \tag{5.75}$$

which exists by continuity of p'. Let $\psi \in C_c^\infty((q_0,1);\mathbb{R})$ satisfy that $|\psi'| \le 1$ and ψ' is supported on $[q,q+\iota_1] \cup [q+\iota-\iota_1, q+\iota]$, positive on $[q,q+\iota_1]$, and negative on $[q+\iota-\iota_1,q+\iota]$. (Note that ψ' integrates to zero because ψ has bounded support, and that ψ is clearly nonnegative.) Let $\iota_2 = \psi(q+\iota_1)$. Consider the perturbation $\widetilde{p} = p + \delta\psi$, which is non-decreasing for small $\delta > 0$ by (5.75). Let $o_{\iota_1}(1)$ denote a term tending to 0 as $\iota_1 \to 0$. We compute that

$$\begin{aligned}
F &\equiv \frac{\mathsf{d}}{\mathsf{d}\delta}\mathbb{A}(\widetilde{p},\Phi;q_0)\Big|_{\delta=0} \\
&= \sum_{s\in\mathscr{S}} \lambda_s \int_{q_0}^1 f_s(q)(\psi \times \xi^s \circ \Phi)'(q) \\
&\ge \sum_{s\in\mathscr{S}} \lambda_s \int_{q_0}^1 \psi'(q) f_s(q)(\xi^s\circ\Phi)(q) && \text{(positivity of } \psi) \\
&= \sum_{s\in\mathscr{S}} \lambda_s \cdot \iota_2 \left(f_s(q)(\xi^s\circ\Phi)(q) - f_s(q+\iota)(\xi^s\circ\Phi)(q+\iota) + o_{\iota_1}(1)\right) && \text{(continuity of } f_s, \xi^s\circ\Phi) \\
&= \iota_2 \left(-\sum_{s\in\mathscr{S}} \lambda_s \Delta_s^\iota (\xi^s\circ\Phi)(q) + o_{\iota_1}(1) + O(\iota)\right) && \text{(continuity of } \xi^s\circ\Phi) \\
&= \iota_2 \left(-\Delta_1^\iota \sum_{s\in\mathscr{S}} \frac{\Phi_s'(q)}{\Phi_1'(q)} \cdot \lambda_s (\xi^s\circ\Phi)(q) + o_{\iota_1}(1) + O(\iota)\right). && \text{(by (5.71))}
\end{aligned}$$

Since (p,Φ,q_0) is a maximizer, $F \le 0$. This implies $\Delta_1^\iota \ge -C\iota$, and by (5.71), $\min \Delta_1^\iota \ge -C\iota$. This proves (5.72) for $\iota > 0$. The proof for $\iota < 0$ is analogous. □

Lemma 5.4.18. *Let $A = (a_{i,j}) \in \mathbb{R}_{>0}^{r\times r}$, $\vec{a},\vec{b} \in \mathbb{R}_{>0}^r$, $\vec{c} \in \mathbb{R}_{>0}^r$, and $c \in \mathbb{R}_{>0}$. Let $A_{\min}, a_{\min}, b_{\min}$ denote the minimal entries of $A, \vec{a}, \vec{b}$, and $a_{\max}$ denote the maximal entry of $\vec{a}$. Suppose the linear system*

$$A\vec{x} + \vec{a}y = \vec{c} \odot \vec{x}, \quad \langle \vec{b}, \vec{x}\rangle = c$$

has solution $(y,\vec{x}) = (y_0, \vec{1})$. If $\vec{c}\,' \in \mathbb{R}_{>0}^r$ satisfies $\|\vec{c} - \vec{c}\,'\|_\infty \le \varepsilon$, then any solution $y \in \mathbb{R}_{\ge 0}$, $\vec{x} \in \mathbb{R}_{\ge 0}^r$ to

$$A\vec{x} + \vec{a}y = \vec{c}\,' \odot \vec{x}, \quad \langle \vec{b}, \vec{x}\rangle = c$$

satisfies

$$|y - y_0| \le \frac{\varepsilon c}{a_{\min} b_{\min}}, \quad \|\vec{x} - \vec{1}\|_\infty \le \frac{2a_{\max}}{a_{\min}} \cdot \frac{\varepsilon c}{A_{\min} b_{\min}}.$$

Proof. Without loss of generality let x_1, x_2 be the largest and smallest entries of $\vec{x}$. As $\langle \vec{b}, \vec{1}\rangle = \langle \vec{b}, \vec{x}\rangle = c$,

$$\frac{c}{b_{\min}} \ge x_1 \ge 1 \ge x_2.$$

Then

$$\begin{aligned}
0 &= a_1 y + \sum_{i=1}^r a_{1,i} x_i - c_1' x_1 \\
&\le a_1 y + \left(\sum_{i=1}^r a_{1,i} - c_1'\right) x_1 - A_{\min}(x_1 - x_2) \\
&= a_1 y + (c_1 - c_1' - a_1 y_0)\, x_1 - A_{\min}(x_1 - x_2) \\
&\le \varepsilon x_1 - a_1 y_0 (x_1 - 1) + a_1(y - y_0) - A_{\min}(x_1 - x_2) \\
&\le \frac{\varepsilon c}{b_{\min}} + a_1(y - y_0) - A_{\min}(x_1 - x_2).
\end{aligned}$$

Analogously

$$\begin{aligned}
0 &= a_2 y + \sum_{i=1}^{r} a_{2,i} x_i - c_2' x_2 \\
&\geq a_2 y + \left(\sum_{i=1}^{r} a_{2,i} - c_2'\right) x_2 + A_{\min}(x_1 - x_2) \\
&= a_2 y + (c_2 - c_2' - a_2 y_0)\, x_2 + A_{\min}(x_1 - x_2) \\
&\geq -\varepsilon x_2 - a_2 y_0 (x_2 - 1) + a_2(y - y_0) + A_{\min}(x_1 - x_2) \\
&\geq -\frac{\varepsilon c}{b_{\min}} + a_2(y - y_0) + A_{\min}(x_1 - x_2).
\end{aligned}$$

Since $x_1 - x_2 \geq 0$, this implies

$$y - y_0 \geq -\frac{\varepsilon c}{a_1 b_{\min}} \geq -\frac{\varepsilon c}{a_{\min} b_{\min}}, \quad y - y_0 \leq \frac{\varepsilon c}{a_2 b_{\min}} \leq \frac{\varepsilon c}{a_{\min} b_{\min}},$$

which proves the first conclusion. Thus,

$$A_{\min}(x_1 - x_2) \leq \left(\frac{\varepsilon c}{b_{\min}} + a_1(y - y_0)\right) \leq \frac{2a_{\max}}{a_{\min}} \cdot \frac{\varepsilon c}{b_{\min}}.$$

Since $x_1 \geq 1 \geq x_2$, we have $\|\vec{x} - \vec{1}\|_\infty \leq x_1 - x_2$ which implies the second conclusion. □

Proposition 5.4.19. *The functions p' and Φ' are Lipschitz on $[q_0 + \varepsilon, 1]$ for all $\varepsilon > 0$. Thus p'' and Φ'' are well-defined as bounded measurable functions on $[q_0 + \varepsilon, 1]$.*

Proof. By Proposition 5.4.17, f_s is Lipschitz on $[q_0 + \varepsilon, 1]$. Since it is also bounded on $[q_0 + \varepsilon, 1]$ by Proposition 5.4.11, f_s^{-2} is Lipschitz as well. Thus, for $q \in [q_0 + \varepsilon, 1]$, $C = C(q)$, and sufficiently small $\iota \in \mathbb{R}$,

$$\begin{aligned}
O(\iota) &\geq |f_1(q + \iota)^{-2} - f_1(q)^{-2}| \\
&= \left| \frac{p'(q+\iota)(\xi^1 \circ \Phi)(q+\iota) + p(q+\iota) \sum_{s \in \mathscr{S}} (\partial_{x_s} \xi^1 \circ \Phi)(q+\iota) \Phi_s'(q+\iota)}{\Phi_1'(q+\iota)} \right. \\
&\qquad \left. - \frac{p'(q)(\xi^1 \circ \Phi)(q) + p(q) \sum_{s \in \mathscr{S}} (\partial_{x_s} \xi^1 \circ \Phi)(q) \Phi_s'(q)}{\Phi_1'(q)} \right| \\
&= |C_1' - C_1 + O(\iota)|
\end{aligned}$$

for

$$\begin{aligned}
C_1 &= \frac{p'(q)(\xi^1 \circ \Phi)(q) + p(q) \sum_{s \in \mathscr{S}} (\partial_{x_s} \xi^1 \circ \Phi)(q) \Phi_s'(q)}{\Phi_1'(q)}, \\
C_1' &= \frac{p'(q+\iota)(\xi^1 \circ \Phi)(q) + p(q) \sum_{s \in \mathscr{S}} (\partial_{x_s} \xi^1 \circ \Phi)(q) \Phi_s'(q+\iota)}{\Phi_1'(q+\iota)}.
\end{aligned}$$

Thus $|C_1 - C_1'| \leq O(\iota)$. Similarly, $|C_s - C_s'| \leq O(\iota)$ for analogously defined C_s, C_s'. Note that the system given by

$$1 = \sum_{s \in \mathscr{S}} \lambda_s \Phi_s'(q) x_s \tag{5.76}$$

$$C_1 \Phi_1'(q) x_1 = (\xi^1 \circ \Phi)(q) y + p(q) \sum_{s \in \mathscr{S}} (\partial_{x_s} \xi^1 \circ \Phi)(q) \Phi_s'(q) x_s \tag{5.77}$$

and analogous equations to (5.77) with $s \in \mathscr{S}$ in place of 1 has solution $y = p'(q)$, $x_1 = \cdots = x_r = 1$. Moreover, the system given by (5.76),

$$C_1' \Phi_1'(q) x_1 = (\xi^1 \circ \Phi)(q) y + p(q) \sum_{s \in \mathscr{S}} (\partial_{x_s} \xi^1 \circ \Phi)(q) \Phi_s'(q) x_s \tag{5.78}$$

and analogous equations to (5.78) with $s \in \mathscr{S}$ in place of 1 has solution $y = p'(q+\iota)$, $x_s = \Phi_s'(q+\iota)/\Phi_s'(q)$. Since $|C_s - C_s'| \le O(\iota)$ for all s, we may apply Lemma 5.4.18 with $\vec{c} = \vec{C}, \vec{c}' = \vec{C}'$, y taking the place of $p'(q)$ or $p'(q+\iota)$, and A corresponding to the last term of (5.77) or (5.78). The result is that

$$|p'(q+\iota) - p'(q)|, \left|\frac{\Phi_s'(q+\iota)}{\Phi_s'(q)} - 1\right| \le O(\iota).$$

(The required constants $A_{\min}, a_{\min}, b_{\min}, a_{\max}$ are bounded thanks to Propositions 5.4.11 and 5.4.12.)

Since Φ_s' is bounded below by Proposition 5.4.11, we conclude that p', Φ' are Lipschitz in a neighborhood of $q \in (q_0, 1]$. This Lipschitz constant is uniform on any $[q_0+\varepsilon, 1]$, thus p', Φ' are Lipschitz on these sets. $\square$

Lemma 5.4.20. *Suppose $A = (a_{i,j}) \in \mathbb{R}_{>0}^{r\times r}$ and $\vec{b} \in \mathbb{R}_{>0}^r$. Let $A_{\max}, A_{\min}$ be the largest and smallest entries of A. Suppose the linear system $A\vec{x} = \vec{b} \odot \vec{x}$ admits the solution $\vec{x} = \vec{1}$. If $\vec{b}' \preceq \vec{b} + \varepsilon\vec{1}$, $A' \ge A$ entry-wise, and the system $A'\vec{x} = \vec{b}' \odot \vec{x}$ admits a nontrivial solution $\vec{x} \in \mathbb{R}_{\ge 0}^r$, then all entries of $A' - A$ are at most $\varepsilon \cdot \frac{rA_{\max}+A_{\min}+\varepsilon}{A_{\min}}$.*

Proof. Assume without loss of generality that x_1 is the smallest entry of $\vec{x}$. Let $\Delta_i = b_i' - b_i$ and $\Delta_{i,j} = a_{i,j}' - a_{i,j}$, so $\Delta_i \le \varepsilon$, $\Delta_{i,j} \ge 0$. We have

$$0 = (b_1 + \Delta_1)x_1 - \sum_{i=1}^r a_{i,j}' x_i \le \Delta_1 x_1 - \sum_{i=1}^r a_{i,j}(x_i - x_1).$$

Thus $a_{i,j}(x_i - x_1) \le \Delta_1 x_1$ for all i. If $x_1 = 0$, this implies $\vec{x} = \vec{0}$, contradiction. Thus $x_1 > 0$ and we may scale $\vec{x}$ such that $x_1 = 1$. This implies

$$1 \le x_i \le 1 + \frac{\Delta_1}{a_{i,j}} \le 1 + \frac{\varepsilon}{A_{\min}}$$

for all i. The equation $b_j' x_j = (A'\vec{x})_j$ implies

$$\begin{aligned}\sum_{i=1}^r \Delta_{j,i} x_i = b_j x_j + \Delta_j x_j - \sum_{i=1}^r a_{j,i} x_i &\le \left(1 + \frac{\varepsilon}{A_{\min}}\right)\sum_{i=1}^r a_{j,i} + \varepsilon\left(1 + \frac{\varepsilon}{A_{\min}}\right) - \sum_{i=1}^r a_{j,i} \\ &\le \varepsilon \cdot \frac{rA_{\max} + A_{\min} + \varepsilon}{A_{\min}}.\end{aligned}$$

Since $x_i \ge 1$ for all i, this implies the result. $\square$

Let $S \subseteq (q_0, 1)$ be the set of q for which (5.68) holds, and for $q \in S$ let $\Psi(q)$ be the common value of the $\Psi_s(q)$. Let $S_1 = \{q \in S : p'(q) > 0\}$ and $S_2 = S \setminus S_1$.

Proposition 5.4.21. *Almost everywhere in S_2, $\Psi(q) < 0$.*

Proof. Suppose for the sake of contradiction that $\Psi(q) \ge 0$ holds for a positive-measure set $T \subseteq S_2$. Let $U \subseteq [q_0, 1]$ be the set of q which are Lebesgue points of $f_s'(q)$ for all $s \in \mathscr{S}$. Since these functions are measurable and integrable on $[q_0+\varepsilon, 1]$ for all $\varepsilon > 0$, U is almost all of $[q_0, 1]$. So $T \cap U$ has positive measure. Let $q \in T \cap U$. Thus

$$\lim_{\iota \to 0^+} \frac{f_1(q+\iota) - f_1(q)}{\iota} = f_1'(q) = \Phi_1'(q)\Psi(q),$$

which implies that for small $\iota > 0$,

$$f_1(q+\iota) = f_1(q) + \Phi_1'(q)\Psi(q)\iota + o(\iota) \ge f_1(q) - o(\iota).$$

Define

$$C_1 = \frac{p(q)(\xi^1 \circ \Phi)'(q)}{\Phi_1'(q)} = f_1(q)^{-2},$$

$$C_1' = \frac{p(q+\iota)(\xi^1 \circ \Phi)'(q+\iota)}{\Phi_1'(q+\iota)} \le f_1(q+\iota)^{-2}.$$

Thus $C_1' \leq C_1 + o(\iota)$. For analogously defined C_s, C_s' we have $C_s' \leq C_s + o(\iota)$. Note that the system given by

$$C_1 \Phi_1'(q) x_1 = \sum_{s \in \mathscr{S}} p(q)(\partial_{x_s} \xi^1 \circ \Phi)(q) \Phi_s'(q) x_q$$

and analogous equations with $s \in \mathscr{S}$ in place of 1 has solution $\vec{x} = \vec{1}$, while the system

$$C_1' \Phi_1'(q) x_1 = \sum_{s \in \mathscr{S}} p(q+\iota)(\partial_{x_s} \xi^1 \circ \Phi)(q+\iota) \Phi_s'(q) x_q$$

and analogous equations with $s \in \mathscr{S}$ in place of 1 has solution $x_s = \Phi_s'(q+\iota)/\Phi_s'(q)$. By Lemma 5.4.20 this implies that for all $s, s' \in \mathscr{S}$,

$$p(q+\iota)(\partial_{x_s} \xi^{s'} \circ \Phi)(q+\iota) \leq p(q)(\partial_{x_s} \xi^{s'} \circ \Phi)(q) + o(\iota).$$

However, since ξ is non-degenerate, $(\partial_{x_s} \xi^{s'} \circ \Phi)(q+\iota) \geq (\partial_{x_s} \xi^{s'} \circ \Phi)(q) + \Omega(\iota)$ for some s, s'. This is a contradiction. $\square$

Lemma 5.4.22. *There exists $q_1 \in [q_0, 1]$ such that, up to modification by a measure zero set, $S_1 = [q_0, q_1]$ and $S_2 = [q_1, 1]$.*

Proof. We will show that there do not exist positive measure subsets $I \subseteq S_1$, $J \subseteq S_2$ with $\sup J \leq \inf I$. Suppose for contradiction that such subsets exist. Define $q^* = \sup J$, $m = \int_I p'(q) \, \mathsf{d}q$, and

$$\psi(q) = \begin{cases} m(\int_{[q_0,q] \cap J} \mathsf{d}q)/(\int_J \mathsf{d}q) & q \leq q^*, \\ m - \int_{[q^*,q] \cap I} p'(q) \, \mathsf{d}q & q > q^*. \end{cases}$$

Note that ψ is absolutely continuous, nonnegative-valued, and positive-valued almost everywhere in J. Moreover $\psi(q_0) = \psi(1) = 0$, and for small $\delta > 0$, the perturbation

$$\widetilde{p}(q) = p(q) + \delta \psi(q) \tag{5.79}$$

remains non-decreasing. Note that

$$\frac{\mathsf{d}}{\mathsf{d}\delta} (p \times \xi^s \circ \Phi)'(q) = (\psi \times \xi^s \circ \Phi)'(q).$$

Thus, integrating by parts,

$$\begin{aligned} F \equiv 2 \frac{\mathsf{d}}{\mathsf{d}\delta} \mathbb{A}(\widetilde{p}, \Phi; q_0) \Big|_{\delta=0} &= \sum_{s \in \mathscr{S}} \int_{q_0}^1 \sqrt{\frac{\Phi_s'(q)}{(p \times \xi^s \circ \Phi)'(q)}} (\psi \times \xi^s \circ \Phi)'(q) \, \mathsf{d}q \\ &= -\sum_{s \in \mathscr{S}} \int_{q_0}^1 \psi(q)(\xi^s \circ \Phi)(q) \Phi_s'(q) \Psi_s(q) \, \mathsf{d}q \\ &= -\sum_{s \in \mathscr{S}} \int_{S_2} \psi(q)(\xi^s \circ \Phi)(q) \Phi_s'(q) \Psi(q) \, \mathsf{d}q. \end{aligned}$$

By Proposition 5.4.21, $\Psi(q) < 0$ almost everywhere in S_2. Therefore $F > 0$ and the perturbation (5.79) improves the value of $\mathbb{A}(p, \Phi; q_0)$, a contradiction.

Finally, define measures

$$\mu([q_0, q]) = \int_{[q_0,q] \cap S_1} \mathsf{d}q, \qquad \nu([q_0, q]) = \int_{[q_0,q] \cap S_2} \mathsf{d}q.$$

The non-existence of I, J implies that $\max \mathsf{supp}(\mu) \leq \min \mathsf{supp}(\nu)$. Since $S_1 \cup S_2$ is almost all of $[q_0, 1]$ the result follows. $\square$

Proof of Proposition 5.4.13. That $p, \Phi_s \in W^{2,\infty}([q_0+\varepsilon, 1])$ follows from Proposition 5.4.19. By Lemma 5.4.22, $p' > 0$ almost everywhere on $[q_0, q_1]$. By Proposition 5.4.17, $\Psi_s = 0$ almost everywhere on $[q_0, q_1]$. Since f_s is Lipschitz, for all $q \in [q_0, q_1]$ we have

$$f_s(q) - f_s(q_0) = \int_{q_0}^{q} f_s'(q) \, \mathsf{d}q = \int_{q_0}^{q} \Phi_s'(q)\Psi_s(q) \, \mathsf{d}q = 0.$$

Thus $f_s(q)^{-2} = \frac{(p \times \xi^s \circ \Phi)'(q)}{\Phi_s'(q)}$ is constant on $[q_0, q_1]$. By Lemma 5.4.22 we have $p' = 0$ almost everywhere on $[q_1, 1]$, hence everywhere by Proposition 5.4.19. And by Proposition 5.4.12 we have $p(1) = 1$. Thus, for all $q \in [q_1, 1]$,

$$p(1) - p(q) = \int_{q}^{1} p'(q) \, \mathsf{d}q = 0,$$

so $p(q) = 1$ for all $q \in [q_1, 1]$. Finally, by Proposition 5.4.17 and Lemma 5.4.22, (5.15) is satisfied for all s, s' almost everywhere on $[q_1, 1]$. □

Given Proposition 5.4.13, it remains to study the behavior of (p, Φ) separately on $[q_0, q_1]$ and $[q_1, 1]$ and establish the root-finding and tree-descending descriptions in Propositions 5.1.9 and 5.1.11. We have seen that (p, Φ) are described by explicit differential equations on $[q_0, q_1]$ and $[q_1, 1]$, and it will be important to understand both. We will refer to them as the type I and II equations respectively in Subsections 5.4.5 and 5.4.6.

5.4.4 Behavior in the root-finding phase 1: super-solvability of $\Phi(q_1)$

Let q_0, q_1 be given by Proposition 5.4.13, and let L_s be the constant value of $(p \times \xi^s \circ \Phi)'(q)/\Phi_s'(q)$ on $[q_0, q_1]$, which exists by Proposition 5.4.13. The goal of this subsection is to prove that $\Phi(q_1)$ is super-solvable.

Lemma 5.4.23. *We have $\Phi_s(q_0) = 0$ if and only if $h_s = 0$.*

Proof. Assume without loss of generality that $s = 1$. First, suppose $h_1 = 0$ and $\Phi_1(q_0) > 0$. By admissibility, $q_0 > 0$. Consider the perturbation $\widetilde{q}_0 = q_0 - \delta$,

$$\widetilde{p}(q) = \begin{cases} q - \widetilde{q}_0 & q \in [\widetilde{q}_0, q_0] \\ \delta + (1-\delta)p(q) & q \in [q_0, 1] \end{cases} \quad \widetilde{\Phi}_s(q) = \begin{cases} \frac{q-\widetilde{q}_0}{\delta}\Phi_s(q_0) & q \in [\widetilde{q}_0, q_0], s = 1 \\ \Phi_s(q_0) & q \in [\widetilde{q}_0, q_0], s \neq 1 \\ \Phi_s(q) & q \in [q_0, 1] \end{cases}$$

for all $s \in \mathscr{S}$. Then,

$$\lambda_s \int_{\widetilde{q}_0}^{q_0} \sqrt{\widetilde{\Phi}_s'(q)(\widetilde{p} \times \xi^s \circ \widetilde{\Phi})'(q)} \, \mathsf{d}q \geq \begin{cases} \Omega(\delta^{1/2}) & s = 1 \\ 0 & s \neq 1 \end{cases}$$

while for all $s \in \mathscr{S}$,

$$\lambda_s \int_{q_0}^{1} \sqrt{\widetilde{\Phi}_s'(q)(\widetilde{p} \times \xi^s \circ \widetilde{\Phi})'(q)} \, \mathsf{d}q \geq \lambda_s \int_{q_0}^{1} \sqrt{\widetilde{\Phi}_s'(q)(\widetilde{p} \times \xi^s \circ \widetilde{\Phi})'(q)} \, \mathsf{d}q - O(\delta)$$
$$h_s\lambda_s\sqrt{\widetilde{\Phi}_s(\widetilde{q}_0)} = h_s\lambda_s\sqrt{\Phi_s(q_0)}.$$

Thus for small $\delta > 0$ the perturbation improves the value of $\mathbb{A}$, contradiction.

Conversely, suppose $h_1 > 0$ and $\Phi_1(q_0) = 0$. Consider the perturbation $(\widetilde{p}, \widetilde{\Phi}, \widetilde{q}_0)$ where $\widetilde{q}_0 = q_0 + \delta$ and $\widetilde{p}, \widetilde{\Phi}$ are p, Φ restricted to $[q_0 + \delta, 1]$. Note that $\widetilde{\Phi}_1(q_0) \geq \Omega(\delta)$ by Proposition 5.4.11. Thus

$$h_1\lambda_1\sqrt{\widetilde{\Phi}_1(q_0)} - h_1\lambda_1\sqrt{\Phi_1(q_0)} \geq \Omega(\delta^{1/2}),$$
$$h_s\lambda_s\sqrt{\widetilde{\Phi}_s(q_0)} - h_s\lambda_s\sqrt{\Phi_s(q_0)} \geq 0 \quad \forall s \neq 1.$$

Furthermore, for all $s \in \mathscr{S}$,

$$\begin{aligned}
&\lambda_s \int_{\widetilde{q}_0}^{1} \sqrt{\widetilde{\Phi}'_s(q)(\widetilde{p} \times \xi^s \circ \widetilde{\Phi})'(q)}\, \mathsf{d}q - \lambda_s \int_{q_0}^{1} \sqrt{\Phi'_s(q)(p \times \xi^s \circ \Phi)'(q)}\, \mathsf{d}q \\
&= \lambda_s \int_{q_0}^{q_0+\delta} \sqrt{\Phi'_s(q)(p \times \xi^s \circ \Phi)'(q)}\, \mathsf{d}q = O(\delta).
\end{aligned}$$

Thus for small $\delta > 0$ the perturbation improves the value of $\mathbb{A}$, contradiction. $\square$

Corollary 5.4.24. *If $\check{h} \neq \vec{0}$, then $0 < q_0 < q_1$ and $\Phi(q_1) \in (0,1]^{\mathscr{S}}$.*

Proof. Lemma 5.4.23 implies $0 < q_0$, so Proposition 5.4.12 implies $p(q_0) = 0$. Since $p(q_1) = 1$ by Proposition 5.4.13, we have $q_0 < q_1$. Proposition 5.4.11 gives $\Phi'(q) \succeq L^{-1}\vec{1}$ for $q \in [q_0, q_1]$, so all coordinates of $\Phi(q_1)$ are positive. $\square$

Lemma 5.4.25. *If $\check{h} = \vec{0}$, then $q_0 = q_1 = 0$ (and $\Phi(q_1) = \vec{0}$).*

Proof. By Lemma 5.4.23, $\Phi(q_0) = \vec{0}$ so $q_0 = 0$. Suppose that $q_1 > 0$. Then, for all $q \in [0, q_1]$, we have $L_s\Phi'_s(q) = (p \times \xi^s \circ \Phi)'(q)$, and by integrating $L_s\Phi_s(q) = p(q)(\xi^s \circ \Phi)(q)$. By Assumption 5.1.3, we can write $\xi^s(\vec{x}) = \sum_{s' \in \mathscr{S}} P_{s,s'}(\vec{x})x_{s'}$ where each $P_{s,s'}$ is a polynomial with nonnegative coefficients and positive constant and linear terms. Thus the functions $P_{s,s'} \circ \Phi$ are all strictly increasing. Let $0 < q < q' < q_1$. The linear system

$$L_s\Phi_s(q)x_s = \sum_{s' \in \mathscr{S}} p(q)(P_{s,s'} \circ \Phi)(q)\Phi_{s'}(q)x_s \quad \forall s \in \mathscr{S}$$

has solution $\vec{x} = \vec{1}$, while the linear system

$$L_s\Phi_s(q)x_s = \sum_{s' \in \mathscr{S}} p(q')(P_{s,s'} \circ \Phi)(q')\Phi_{s'}(q)x_s \quad \forall s \in \mathscr{S}$$

has solution $x_s = \Phi_s(q')/\Phi_s(q)$. Monotonicity of $P_{s,s'} \circ \Phi$ implies $p(q')(P_{s,s'} \circ \Phi)(q') \geq p(q)(P_{s,s'} \circ \Phi)(q)$, so Lemma 5.4.20 (with $\varepsilon = 0$) implies that $p(q')(P_{s,s'} \circ \Phi)(q') = p(q)(P_{s,s'} \circ \Phi)(q)$ for all s, s'. This contradicts that the $P_{s,s'} \circ \Phi$ are strictly increasing. $\square$

Lemma 5.4.26. *If $h_s > 0$, then $L_s = \frac{h_s^2}{\Phi_s(q_0)}$.*

Proof. Assume without loss of generality that $s = 1$. Consider the following perturbation $\widetilde{\Phi}$ of Φ. For all $s \neq 1$, $\widetilde{\Phi}_s = \Phi_s$, and $\widetilde{\Phi}_1(q) = \Phi_1(q) + \delta\psi(q)$ where $\psi \in C^\infty([q_0, 1])$ with $\psi(q_0) = 1$ and $\psi = 0$ on $[q_1, 1]$. This perturbation is not admissible, but we nonetheless have $\mathbb{A}(p, \widetilde{\Phi}; q_0) \leq \mathbb{A}(p, \Phi; q_0)$ by Lemma 5.4.9.

Recall the calculation (5.69). Integrating by parts,

$$\begin{aligned}
F_1 &\equiv 2\lambda_1^{-1} \frac{\mathsf{d}}{\mathsf{d}\delta} \mathbb{A}(p, \widetilde{\Phi}; q_0)\Big|_{\delta=0} \\
&= \frac{h_1}{\sqrt{\Phi_1(q_0)}} + \int_{q_0}^{1} L_1^{1/2}\psi'(q)\, \mathsf{d}q + \sum_{s \in \mathscr{S}} \int_{q_0}^{1} L_s^{1/2}(p\psi \times \partial_{x_s}\xi^1 \circ \Phi)'(q) = \frac{h_1}{\sqrt{\Phi_1(q_0)}} - L_1^{1/2}.
\end{aligned}$$

Recall that $\Phi'_1(q)$ is uniformly lower bounded by Proposition 5.4.11 and $\Phi_1(q_0) > 0$ by Lemma 5.4.23. So, this perturbation is valid for small positive and negative δ. Thus $F_1 = 0$ which implies the result. $\square$

Proposition 5.4.27. *If $\check{h} \neq \vec{0}$, then for all s,*

$$L_s = \frac{(\xi^s \circ \Phi)(q_1) + h_s^2}{\Phi_s(q_1)}, \tag{5.80}$$

which is well-defined by Corollary 5.4.24. Thus, (p, Φ) satisfies (5.11) for all $s \in \mathscr{S}$, $q \in [q_0, q_1]$ with $\vec{x} = \Phi(q_1)$.

Proof. Note that $\Phi_s(q_1) > 0$ for all s by Corollary 5.4.24 and Proposition 5.4.11. Integrating the equation $(p \times \xi^s \circ \Phi)'(q) = L_s \Phi_s'(q)$ on $[q_0 + \varepsilon, q]$ and using continuity of p and Φ and that $p(q_0) = 0$, we find

$$p(q)(\xi^s \circ \Phi)(q) = L_s(\Phi_s(q) - \Phi_s(q_0)). \tag{5.81}$$

Since $p(q_1) = 1$ by Proposition 5.4.13, we have

$$(\xi^s \circ \Phi)(q_1) = L_s(\Phi_s(q_1) - \Phi_s(q_0)). \tag{5.82}$$

If $h_s = 0$, by Lemma 5.4.23 $\Phi_s(q_0) = 0$, so $L_s = (\xi^s \circ \Phi)(q_1)/\Phi_s(q_1)$ as desired. Otherwise, by Lemma 5.4.26, $L_s = h_s^2/(\lambda_s \Phi_s(q_0))$. Plugging this into (5.82) implies

$$\Phi_s(q_0)\left((\xi^s \circ \Phi)(q_1) + h_s^2\right) = h_s^2 \Phi_s(q_1). \tag{5.83}$$

Thus

$$L_s = \frac{(\xi^s \circ \Phi)(q_1)}{\Phi_s(q_1) - \Phi_s(q_0)} = \frac{(\xi^s \circ \Phi)(q_1) + h_s^2}{\Phi_s(q_1)}$$

as desired. □

Corollary 5.4.28. *For $(p, \Phi; q_0)$ maximizing $\mathbb{A}$, we have*

$$\mathbb{A}(p, \Phi; q_0) = \sum_{s \in \mathscr{S}} \lambda_s \left[\sqrt{\Phi_s(q_1)(\xi^s(\Phi(q_1)) + h_s^2)} + \int_{q_1}^1 \sqrt{\Phi_s'(q)(\xi^s \circ \Phi)'(q)}\ \mathsf{d}q\right]. \tag{5.84}$$

Proof. If $\check{h} = \vec{0}$, then $q_1 = 0$ by Lemma 5.4.25. Thus, $p = 1$ on $[0, 1]$ by Proposition 5.4.13. Thus $(p \times \xi^s \circ \Phi)' = (\xi^s \circ \Phi)'$ and the result is clear. Otherwise $\check{h} \neq \vec{0}$, and Corollary 5.4.24 implies $q_1 > q_0$.

If $h_s = 0$, then by Lemma 5.4.23, $\Phi_s(q_0) = 0$. So,

$$\begin{aligned}
h_s \lambda_s \sqrt{\Phi_s(q_0)} + \lambda_s \int_{q_0}^{q_1} \sqrt{\Phi_s'(q)(p \times \xi^s \circ \Phi)'(q)}\ \mathsf{d}q &= \lambda_s \int_{q_0}^{q_1} \Phi_s'(q) \sqrt{L_s}\ \mathsf{d}q \\
&= \lambda_s \Phi_s(q_1) \sqrt{L_s} = \lambda_s \sqrt{\Phi_s(q_1)(\xi^s \circ \Phi)(q_1)},
\end{aligned}$$

as desired. The last step uses Proposition 5.4.27. If $h_s > 0$, then by Lemma 5.4.26 and Proposition 5.4.27,

$$\begin{aligned}
h_s \lambda_s \sqrt{\Phi_s(q_0)} + \lambda_s \int_{q_0}^{q_1} \sqrt{\Phi_s'(q)(p \times \xi^s \circ \Phi)'(q)}\ \mathsf{d}q &= \lambda_s \left[\Phi_s(q_0)\sqrt{L_s} + \int_{q_0}^{q_1} \Phi_s'(q)\sqrt{L_s}\ \mathsf{d}q\right] \\
&= \lambda_s \Phi_s(q_1) \sqrt{L_s} \\
&= \lambda_s \sqrt{\Phi_s(q_1)\left((\xi^s \circ \Phi)(q_1) + h_s^2\right)}.
\end{aligned}$$ □

The following variant of this calculation determines the energy attained by $(p, \Phi; q_0)$ partway through the root-finding phase, and is used in Remark 5.1.16.

Corollary 5.4.29. *If $(p, \Phi; q_0)$ maximizes $\mathbb{A}$ and $q \in [q_0, q_1]$, then*

$$\sum_{s \in \mathscr{S}} \lambda_s \left[h_s \sqrt{\Phi_s(q_0)} + \int_{q_0}^{q} \sqrt{\Phi_s'(t)(p \times \xi^s \circ \Phi)'(t)}\ \mathsf{d}t\right] = \sum_{s \in \mathscr{S}} \lambda_s \sqrt{\Phi_s(q)(p(q)(\xi^s \circ \Phi)(q) + h_s^2)}.$$

Proof. If $h_s = 0$, then by Lemma 5.4.23, $\Phi_s(q_0) = 0$. Then (5.81) implies $L_s = p(q)(\xi^s \circ \Phi)(q)/\Phi_s(q)$. So

$$h_s \sqrt{\Phi_s(q_0)} + \int_{q_0}^{q} \sqrt{\Phi_s'(t)(p \times \xi^s \circ \Phi)'(t)}\ \mathsf{d}t = (\Phi_s(q) - \Phi_s(q_0))\sqrt{L_s} = \sqrt{\Phi_s(q) p(q)(\xi^s \circ \Phi)(q)}.$$

If $h_s > 0$, (5.81) implies and Lemma 5.4.26 imply

$$p(q)(\xi^s \circ \Phi)(q) = \frac{h_s^2}{\Phi_s(q_0)}(\Phi_s(q) - \Phi_s(q_0)),$$

which rearranges to

$$\frac{h_s^2 \Phi_s(q)}{\Phi_s(q_0)} = p(q)(\xi^s \circ \Phi)(q) + h_s^2.$$

Then

$$\begin{aligned} h_s \sqrt{\Phi_s(q_0)} + \int_{q_0}^{q} \sqrt{\Phi_s'(t)(p \times \xi^s \circ \Phi)'(t)} \, \mathsf{d}t &= h_s \sqrt{\Phi_s(q_0)} + (\Phi_s(q) - \Phi_s(q_0)) \sqrt{\frac{h_s^2}{\Phi_s(q_0)}} \\ &= \frac{h_s \Phi_s(q)}{\sqrt{\Phi_s(q_0)}} = \sqrt{\Phi_s(q)(p(q)(\xi^s \circ \Phi)(q) + h_s^2)}. \end{aligned}$$

Summing over $s \in \mathscr{S}$ completes the proof. $\square$

Lemma 5.4.30. *If $q_1 = 1$, then $\Phi(q_1) = \vec{1}$ is super-solvable. If $q_1 < 1$, then $\Phi(q_1)$ is solvable.*

Proof. First suppose $q_1 = 1$. Admissibility and the fact that $\Phi(1) \in [0,1]^{\mathscr{S}}$ implies $\Phi(q_1) = \vec{1}$. We have $p(q_1) = 1$ by Proposition 5.4.12 and also $p'(q_1) \geq 0$. By Proposition 5.4.27,

$$\frac{(\xi^s \circ \Phi)(q_1) + h_s^2}{\Phi_s(q_1)} = \frac{(p \times \xi^s \circ \Phi)'(q_1)}{\Phi_s'(q_1)} \geq \frac{\sum_{s' \in \mathscr{S}} (\partial_{x_{s'}} \xi^s \circ \Phi)(q_1) \Phi_{s'}'(q_1)}{\Phi_s'(q_1)}. \tag{5.85}$$

This implies via Corollary 5.4.5 (with Φ' in the role of $\vec{v}$) that $\Phi(q_1)$ is super-solvable.

Now suppose $q_1 < 1$. If $\check{h} = \vec{0}$ the result follows from Lemma 5.4.25, so assume $\check{h} \neq \vec{0}$. Because $p(q) = 1$ on $[q_1, 1]$ and p' is continuous (Proposition 5.4.11), $p'(q_1) = 0$. So, the inequality in (5.85) is an equality. Thus $\Phi'(q_1)$ is in the null space of $M^*(\Phi(q_1))$, and thus (by (5.64)) of $M^*_{\mathsf{sym}}(\Phi(q_1))$. So $M^*_{\mathsf{sym}}(\Phi(q_1))$ is singular and $\Phi(q_1)$ is solvable. $\square$

5.4.5 Behavior in the root-Finding phase 2: well-posedness

In this subsection we prove Proposition 5.1.9 and give a detailed characterization of (p, Φ) on $[q_0, q_1]$ in Proposition 5.4.37. Recalling Propositions 5.4.27 and 5.4.30, we consider a path (p, Φ) defined by the **type I equation**

$$\begin{gathered} \frac{(p \times \xi^s \circ \Phi)'(q)}{\Phi_s'(q)} = L_s = \frac{(\xi^s \circ \Phi)(q_1) + h_s^2}{\Phi_s(q_1)}, \quad \forall s \in \mathscr{S} \\ \Phi_s'(q) \geq 0, \quad \langle \vec{\lambda}, \Phi'(q) \rangle = 1 \end{gathered} \tag{5.86}$$

with super-solvable initial condition $\Phi(q_1)$ and $p(q_1) = 1$. We start by verifying the first part of Proposition 5.1.9, namely that $\check{h} \neq \vec{0}$ if and only if there exists a super-solvable point $\vec{x} \in [0,1]^{\mathscr{S}}$ with $\langle \vec{\lambda}, \vec{x} \rangle > 0$.

Proof of Proposition 5.1.9 (first claim). First, assume $\check{h} \neq \vec{0}$. We will show that all $\vec{x} \in [\delta/2, \delta]^{\mathscr{S}}$ are super-solvable for $\delta > 0$ sufficiently small. Assume without loss of generality that $h_1 > 0$. Note that for all $s \in \mathscr{S}$,

$$(M^*(\vec{x})\vec{x})_s = x_s \left(h_s^2 + \xi^s(\vec{x}) - \sum_{s' \in \mathscr{S}} x_{s'} \partial_{x_{s'}} \xi^s(\vec{x}) \right) = x_s \left(h_s^2 - O(\delta^2) \right).$$

Moreover, $(M^*(\vec{x})\vec{e}_1)_1 \leq h_1^2 + O(\delta)$, while for $s \neq 1$,

$$(M^*(\vec{x})\vec{e}_1)_s = -x_s \partial_{x_1} \xi^s(\vec{x}).$$

Thus, for $\vec{v} = \vec{x} - \frac{1}{2} x_1 \vec{e}_1$, we have

$$(M^*(\vec{x})\vec{v})_1 \geq x_1 \left(\frac{1}{2} h_1^2 - O(\delta) \right) \geq 0$$

and for $s \neq 1$,

$$(M^*(\vec{x})\vec{v})_s \geq x_s \left(x_1 \partial_{x_1} \xi^s(\vec{x}) - O(\delta^2) \right) \geq 0.$$

This implies by Corollary 5.4.5 that $\vec{x}$ is super-solvable.

If $\check{h} = \vec{0}$, fix any $\vec{x} \in (0,1]^{\mathscr{S}}$. Note that

$$\vec{x}^{\top} M^*_{\mathsf{sym}}(\vec{x})\vec{x} = \sum_{s\in\mathscr{S}} x_s \partial_{x_s}\xi(\vec{x}) - \sum_{s,s'\in\mathscr{S}} x_s x_{s'} \partial_{x_s,x_{s'}}\xi(\vec{x}) < 0,$$

as any monomial of $\xi(\vec{x})$ with total degree $p \geq 2$ appears with multiplicity p in the first sum and $p(p-1) \geq p$ in the second, with strict inequality for any $p > 2$. Thus $\vec{x}$ is strictly sub-solvable. $\square$

Proposition 5.4.31. *Define for $(p(q), p'(q), \Phi(q)) \in [0,1]^{\mathscr{S}} \times [0,1] \times \mathbb{R}$ the $\mathscr{S} \times \mathscr{S}$ matrix $M(p(q), p'(q), \Phi(q))$ with entries*

$$M(p(q), p'(q), \Phi(q))_{s,s'} = \frac{p(q)\partial_{x_{s'}}\xi^s\left(\Phi(q)\right) + \lambda_{s'}p'(q)\xi^s(\Phi(q))}{L_s}, \qquad s, s' \in \mathscr{S}.$$

If (p, Φ) solves (5.86) then $\Lambda(M(p, p', \Phi)) = 1$ with Perron-Frobenius eigenvector $\Phi'(q)$.

Proof. It suffices to expand the left-hand side of the top line of (5.86):

$$p'(q)\xi^s(\Phi(q))\left(\sum_{s'\in\mathscr{S}} \lambda_{s'}\Phi'_{s'}(q)\right) + p(q)\sum_{s'\in\mathscr{S}} \Phi'_{s'}(q)\partial_{x_{s'}}\xi^s\left(\Phi(q)\right) = L_s\Phi'_s(q), \quad \forall s \in \mathscr{S}. \tag{5.87}$$

Rearranging shows that $M(p, p', \Phi)\Phi'(q) = \Phi'(q)$, and it is clear that M has non-negative entries. $\square$

We now show the ODE (5.86) is well-posed.

Lemma 5.4.32. *Fix $q \in [0,1]$ and let $Y(q) = (p(q), \Phi(q))$ and $L_s > 0$ be arbitrary. The equation (5.86) for any fixed q is equivalent to*

$$Y'(q) = F(Y(q))$$

for a locally Lipschitz function $F : [0,1] \times \left([0,1]^r \backslash \vec{0}\right) \to \mathbb{R}^{r+1}$.

Proof. Let M be as in Proposition 5.4.31. Because ξ is non-degenerate, Propositions 5.4.6 and 5.4.31 imply existence of $c > 0$ such that

$$M(p, x+y, \Phi) \geq M(p, x, \Phi) + cy$$

holds entrywise for all $x, y \geq 0$, as long as $\Phi(q) \in \mathbb{R}^r_{\geq 0} \backslash [0, \varepsilon]^r$. Therefore a unique value $p'(q)$ solving (5.86) exists. Moreover M is locally Lipschitz in (p, Φ), so if

$$M(p, p', \Phi) = M(\widetilde{\Phi}, \widetilde{p}, \widetilde{p}')$$

then

$$|p' - \widetilde{p}'| \leq O(\|\Phi - \widetilde{\Phi}\|_{L^\infty} + |p - \widetilde{p}|).$$

(With implicit constant depending on ε as introduced above.) This shows that p' has locally Lipschitz dependence on $Y = (p, \Phi)$. It remains to show Φ', defined by the resulting solution to (5.87), also has locally Lipschitz dependence on Y. This follows by Proposition 5.4.33 below. (Note that all entries of M are of the same order up to constants for $\Phi(q) \in \mathbb{R}^r_{\geq 0} \backslash [0, \varepsilon]^r$ by non-degeneracy of ξ.) $\square$

Proposition 5.4.33 ([Yeo18, Lemma 27])**.** *Let $\mathcal{M} \subseteq \mathbb{R}^{r\times r}_{\geq 0}$ be a compact set of square matrices all of whose Perron-Frobenius eigenvalues have multiplicity 1. Let $M, \widetilde{M} \in \mathcal{M}$ have entrywise positive Perron-Frobenius eigenvectors $v, \widetilde{v}$, normalized so that $\|v\|_1 = \|\widetilde{v}\| = 1$. Then*

$$\|v - \widetilde{v}\|_1 \leq O_{\mathcal{M}}(\|M - \widetilde{M}\|_1).$$

In particular, this holds for $\mathcal{M} = [c, C]^{r\times r}$ for any $0 < c < C < \infty$.

Lemma 5.4.32 shows that for any right endpoint $(p(q_1), \Phi(q_1))$, it is possible to solve (5.86) backwards in time until q_* when $\Phi(q)$ reaches the boundary of $\mathbb{R}^r_{\geq 0}$, or at which $p(q)$ reaches 0. We now show that the latter occurs first.

Lemma 5.4.34. *There exists $c > 0$ such that for any super-solvable point $\Phi(q_1)$, the solution to the type* I *equation* (5.86) *on* $[q_*, q_1]$ *satisfies*

$$\Phi_s(q) \geq cp(q)q.$$

Moreover $p(q_*) = 0$ and Lemma 5.4.23 holds for q_*, i.e. $h_s > 0$ if and only if $\Phi_s(q_*) > 0$.

Proof. Observe that in (5.86), we have

$$L_s \geq K_s \equiv \frac{\xi^s(\Phi(q_1))}{\Phi_s(q_1)}.$$

Therefore on $q \in [q_\varepsilon, q_1]$, the left-hand equation in (5.86) implies

$$\frac{(p \times \xi^s \circ \Phi)(q)}{\Phi_s(q)} \leq K_s.$$

Recall that ξ^s is non-degenerate, and so admissibility and $\Phi \succeq 0$ implies $\xi^s(\Phi(q)) = \Theta(q)$. Hence for some $c > 0$ and all $s \in \mathscr{S}$,

$$\Phi_s(q) \geq \Omega(p(q)q/K_s) \geq cp(q)q.$$

This concludes the proof of the first statement, which implies that $p(q_*) = 0$.

For the second, note that strict inequality holds in the first step if $h_s > 0$, and so p must reach 0 before Φ_s does. On the other hand if $h_s = 0$, then it is easy to see from (5.86) that p cannot reach zero strictly sooner than Φ_s, hence the numerator and denominator on the left-hand side in (5.86) both reach zero at time q_*. □

Lemma 5.4.35. *If $\Phi(q_1)$ is super-solvable, then the p solving* (5.87) *is non-decreasing and concave on $[q_*, q_1]$. Moreover $p, \Phi_s \in C^1([q_*, q_1])$.*

Proof. We claim that p' is decreasing. The key point is that with M as in Proposition 5.4.31,

$$M(p, p', \Phi) < M(\widetilde{\Phi}, \widetilde{p}, \widetilde{p}')$$

if $\Phi \preceq \widetilde{\Phi}$, $p \leq \widetilde{p}$ and $p' < \widetilde{p}'$. Indeed this is immediate by Proposition 5.4.6. It follows that p' must increase backward in time, i.e. $p'(q)$ is a decreasing function. Since $p'(q_1) \geq 0$ by super-solvability, this completes the proof. □

Proof of Proposition 5.1.9, parts (a,b). Existence and uniqueness of the root-finding trajectory follows from Lemma 5.4.32 and Proposition 5.4.8. Lemma 5.4.34 ensures that the solution exists until p reaches 0. Concavity of p was just shown in Lemma 5.4.35. This proves part (a). Part (b) follows from Lemma 5.4.23 or 5.4.34. □

Lemma 5.4.36. *If $\vec{1}$ is super-solvable, then $q_1 = 1$ and $\Phi(q_1) = \vec{1}$. Otherwise $q_1 < 1$.*

Proof. If $\vec{1}$ is strictly sub-solvable, Lemma 5.4.30 implies that $q_1 < 1$. Suppose $\vec{1}$ is super-solvable. Let (p^*, Φ^*, q_0^*) be the root-finding trajectory with endpoint $\vec{1}$, which exists by Proposition 5.1.9. By Corollary 5.4.28,

$$\mathbb{A}(p^*, \Phi^*; q_0^*) = \sum_{s \in \mathscr{S}} \lambda_s \sqrt{\xi^s(\vec{1}) + h_s^2}.$$

Suppose for contradiction that there is a different maximizer (p, Φ, q_0) of $\mathbb{A}$ with $\mathbb{A}(p, \Phi; q_0) \geq \mathbb{A}(p^*, \Phi^*; q_0^*)$. The maximizer (p, Φ, q_0) has its own value q_1, and we must have $q_1 < 1$ since for this to be a different maximizer. Note that for each $s \in \mathscr{S}$,

$$\sqrt{\xi^s(\vec{1}) + h_s^2} - \sqrt{\Phi_s(q_1)(\xi^s(\Phi(q_1)) + h_s^2)} = \int_{q_1}^1 \frac{\mathsf{d}}{\mathsf{d}q} \sqrt{\Phi_s(q)((\xi^s \circ \Phi)(q) + h_s^2)}\, \mathsf{d}q$$

$$= \frac{1}{2} \int_{q_1}^1 \left(\Phi_s'(q) \sqrt{\frac{(\xi^s \circ \Phi)(q) + h_s^2}{\Phi_s(q)}} + (\xi^s \circ \Phi)'(q) \sqrt{\frac{\Phi_s(q)}{(\xi^s \circ \Phi)(q) + h_s^2}} \right) \mathsf{d}q.$$

By Corollary 5.4.28,

$$F \equiv \mathbb{A}(p^*, \Phi^*; q_0^*) - \mathbb{A}(p, \Phi; q_0)$$
$$= \sum_{s\in\mathscr{S}} \frac{\lambda_s}{2} \int_{q_1}^{1} (\xi^s \circ \Phi)'(q) \sqrt{\frac{\Phi_s(q)}{(\xi^s \circ \Phi)(q) + h_s^2}} \left(\sqrt{\frac{\Phi_s'(q)}{(\xi^s \circ \Phi)'(q)} \cdot \frac{(\xi^s \circ \Phi)(q) + h_s^2}{\Phi_s(q)}} - 1 \right)^2 \mathsf{d}q \geq 0.$$

Since $\mathbb{A}(p, \Phi; q_0) \geq \mathbb{A}(p^*, \Phi^*; q_0^*)$, we have $F = 0$. So, for all $s \in \mathscr{S}$, and almost all $q \in (q_1, 1]$

$$\frac{(\xi^s \circ \Phi)'(q)}{(\xi^s \circ \Phi)(q) + h_s^2} = \frac{\Phi_s'(q)}{\Phi_s(q)} \qquad \Rightarrow \qquad \frac{\mathsf{d}}{\mathsf{d}q} \log\left((\xi^s \circ \Phi)(q) + h_s^2\right) = \frac{\mathsf{d}}{\mathsf{d}q} \log \Phi_s(q).$$

Both sides of this equation are continuous on $(q_1, 1]$, so in fact it holds for all $q \in (q_1, 1]$. Thus there exist constants C_s such that

$$(\xi^s \circ \Phi)(q) + h_s^2 = C_s \Phi_s(q).$$

Thus, for $q_1 < q < q + \iota \leq 1$, we have

$$C_s \Phi_s'(q) = (\xi^s \circ \Phi)'(q) = \sum_{s'\in\mathscr{S}} (\partial_{x_{s'}} \xi^s \circ \Phi)(q) \Phi_{s'}'(q) \quad \forall s \in \mathscr{S},$$
$$C_s \Phi_s'(q + \iota) = (\xi^s \circ \Phi)'(q + \iota) = \sum_{s'\in\mathscr{S}} (\partial_{x_{s'}} \xi^s \circ \Phi)(q + \iota) \Phi_{s'}'(q + \iota) \quad \forall s \in \mathscr{S},$$

We treat these equations as linear systems in $\Phi'(q)$ and $\Phi'(q+\iota)$. Since both linear systems have nonnegative solutions and $(\partial_{x_{s'}} \xi^s \circ \Phi)(q + \iota) \geq (\partial_{x_{s'}} \xi^s \circ \Phi)(q)$ for all s, s', Lemma 5.4.20 (with $\varepsilon = 0$) implies that $(\partial_{x_{s'}} \xi^s \circ \Phi)(q + \iota) = (\partial_{x_{s'}} \xi^s \circ \Phi)(q)$ for all s, s'. This contradicts that ξ is non-degenerate and completes the proof. □

Proposition 5.4.37. *The following assertions hold.*

(a) If $\vec{1}$ is super-solvable, then $0 < q_0 < q_1 = 1$ (and thus $\Phi(q_1) = \vec{1}$).

(b) If $\vec{1}$ is sub-solvable and $\check{h} \neq \vec{0}$, then $0 < q_0 < q_1 < 1$ and $\Phi(q_1) \in (0, 1]^{\mathscr{S}}$.

(c) If $\check{h} = \vec{0}$, then $\vec{1}$ is sub-solvable and $0 = q_0 = q_1$ (and thus $\Phi(q_1) = \vec{0}$).

In cases (b, c), $\Phi(q_1)$ is solvable. In cases (a, b) (and vacuously in case (c)) (p, Φ) restricted to $[q_0, q_1]$ is the root-finding trajectory with endpoint $\Phi(q_1)$.

Proof of Proposition 5.4.37. If $\vec{1}$ is super-solvable, Lemma 5.4.36 implies $q_1 = 1$. Comparing Corollary 5.4.24 and Lemma 5.4.25 gives $q_0 > 0$. If $\vec{1}$ is sub-solvable and $\check{h} \neq \vec{0}$, Lemma 5.4.36 implies $q_1 < 1$ while Corollary 5.4.24 implies $0 < q_0 < q_1$ and $\Phi(q_1) \in (0, 1]^{\mathscr{S}}$. If $\check{h} = \vec{0}$, Lemma 5.4.25 implies $0 = q_0 = q_1$. This proves assertions (a, b, c).

In cases (b, c), since $q_1 < 1$, Lemma 5.4.30 implies $\Phi(q_1)$ is solvable. In cases (a, b), Proposition 5.4.27 implies (p, Φ) restricted to $[q_0, q_1]$ is the root-finding trajectory with endpoint $\Phi(q_1)$. □

5.4.6 Behavior in the tree-descending phase

The next lemma, proved in Appendix 5.C.3, shows the tree-descending ODE is also well-posed.

Lemma 5.4.38. *Fix $\varepsilon > 0$. For $\Phi(q) \in \mathbb{R}_{\geq 0}^{\mathscr{S}}$ and $\Phi'(q) \in A_{\geq 0}(q)$, the **type** II **equation***

$$\Psi_s(q) = \Psi_{s'}(q) \quad \forall s, s' \in \mathscr{S};$$
$$\langle \vec{\lambda}, \Phi''(q) \rangle = 0$$

is equivalent (for each fixed q) to

$$\Phi''(q) = F(\Phi(q), \Phi'(q))$$

for a locally Lipschitz function $F : \mathbb{R}_{\geq 0}^{\mathscr{S}} \times A_{\geq 0}^{\mathscr{S}} \to \mathbb{R}^{\mathscr{S}}$. Moreover,

$$|\Phi_s''(q)| \leq O(|\Phi_s'(q)|), \quad \forall s \in \mathscr{S}. \tag{5.88}$$

with a uniform constant for bounded $\Phi'(q)$.

Lemma 5.4.39. *The type* II *equation has a unique solution on* $q \in [q_1, 1]$ *for any initial condition* $(\Phi(q_1), \Phi'(q_1)) \in \mathbb{R}_{\geq 0}^{\mathscr{S}} \times A_{\geq 0}$. *This solution satisfies* $\Phi'(q) \succeq 0$ *for all* q.

Proof. The result now follows from Proposition 5.4.8, since (5.88) implies that $\Phi_s'(q)$ stays non-negative for all s, and stays strictly positive if $\Phi_s'(q_1) > 0$. □

Proof of Proposition 5.1.11. Given the above, it only remains to show existence and uniqueness of $\vec{v}$. Consider the matrix

$$M(\vec{x})_{s,s'} = \frac{\partial_{x_{s'}} \xi^s(\vec{x})}{\xi^s(\vec{x}) + h_s^2}.$$

Then M has strictly positive entries by non-degeneracy. The equation $M^*_{\mathsf{sym}}(\vec{x})\vec{v} = \vec{0}$ is equivalent to $M^*(\vec{x})\vec{v} = \vec{0}$ by (5.64), which is in turn equivalent to

$$M(\vec{x})\vec{v} = \vec{v}.$$

Since $\vec{x} \neq \vec{0}$, non-degeneracy of ξ implies that $\xi^s(\vec{x}) > 0$ so there is no division by 0. Hence any such $\vec{v}$ is uniquely determined as the Perron-Frobenius eigenvector of M. Conversely it is easy to see that if M has Perron-Frobenius eigenvector **not** equal to 1 then M^* would not be solvable, which ensures that $\vec{v}$ as above exists. □

Corollary 5.4.40. $p, \Phi_s \in C^1([q_0, 1])$ *and their restrictions to* $[q_1, 1]$ *are* C^2.

Proof. From Proposition 5.4.11, for the first statement it suffices to verify continuity of p', Φ_s' at q_0. If $\check{h} \neq \vec{0}$ this follows by Lemmas 5.4.32 and 5.4.35. If $\check{h} = \vec{0}$ this and the second conclusion both follow from Lemmas 5.4.25 and 5.4.39. □

The statement of Theorem 5.1.12 is a combination of many of the results established in this section.

Proof of Theorem 5.1.12. Existence of a maximizer $(p, \Phi; q_0)$ was shown in Proposition 5.4.10, and such p, Φ are continuously differentiable on $[q_0, 1]$ by Corollary 5.4.40. The value q_1 was identified in Lemma 5.4.22. The behavior on $S_1 = [q_0, q_1]$ and $S_2 = [q_1, 1]$ comes directly from the well-posedness of the corresponding ODEs as shown in Lemmas 5.4.32 and 5.4.39. The formula (5.19) was proved in Corollary 5.4.28. The last assertions follow from Proposition 5.4.37. □

We finally prove a slight generalization of Proposition 5.1.22. Recall that $\Delta^r \subseteq \mathbb{R}^r_{\geq 0}$ denotes the simplex of admissible Φ' vectors. For any initial point $\vec{x}$ and time-increment $t > 0$, solving the type II equation yields a map $F_{\vec{x},t} : \Delta^r \to \Delta^r$ given by

$$F_{\vec{x},t}(\vec{v}) = (\Phi(q+t) - \vec{x})/t \tag{5.89}$$

where Φ solves the type II equation with initial condition $\Phi(q) = \vec{x}$, $\Phi'(q) = \vec{v}$.

We remark that in the case $\vec{x} = 0$ of Proposition 5.1.22, surjectivity also follows simply by taking $(p, \Phi; q_0)$ maximizing a version of $\mathbb{A}$ rescaled to have an arbitrary endpoint.

Corollary 5.4.41. *Assume* ξ *is non-degenerate. For* $C > 0$, *there exists* $\varepsilon = \varepsilon(C)$ *such that the map* $F_{\vec{x},t}$ *defined in* (5.89) *is injective for* $t \in [0, \varepsilon]$ *and* $\|\vec{x}\|_1 \leq C$. *Moreover* $F_{\vec{x},t}$ *is always surjective.*

Proof. An easy Grönwall argument using (5.88) implies that for $0 \leq t \leq \varepsilon$,

$$\langle \Phi(q+t) - \widetilde{\Phi}(q+t), \Phi'(q) - \widetilde{\Phi}'(q) \rangle > 0$$

for any pair $(\Phi, \widetilde{\Phi})$ of solutions to the type II equation with $\Phi(q) = \widetilde{\Phi}(q)$ and $\Phi'(q) \neq \widetilde{\Phi}'(q)$. This implies injectivity. Surjectivity follows from Lemma 5.4.42 since (5.88) implies that if $\vec{v}_s = 0$ then $F_{\vec{x},t}(\vec{v})_s = 0$. □

Lemma 5.4.42 ([JR76, Lemma 2.1] or [Kar09, Lemma 1])**.** *Let* F *be a continuous map from* Δ^r *to itself such that* $F(\vec{v})_s = 0$ *if* $v_s = 0$. *Then* F *is surjective.*

5.4.7 Explicit solution for pure models

In this subsection we prove Theorem 5.1.18 and Corollary 5.1.19, obtaining an explicit description of $\widehat{\mathsf{ALG}}$ in the important special case of *pure* models for which

$$\xi(x_1, \dots, x_r) = \prod_{s \in \mathscr{S}} x_s^{a_s}. \tag{5.90}$$

Due to the homogeneity and lack of external field, it is natural to expect that the optimal (p, Φ) is given by $p \equiv 1$ and $\Phi(q) = (q^{b_1}, \dots, q^{b_r})$ for positive constants b_s. (Here we do not require Φ to be admissible, which by Lemma 5.4.9 does not make a difference.) Most of our previous results do not apply directly because ξ violates the non-degeneracy condition, however as mentioned previously we can apply them after adding a small perturbation.

Lemma 5.4.43. *For a pure model described by ξ, there exists Φ^* such that with $p \equiv 1$,*

$$\mathbb{A}(p, \Phi^*; 0) = \widehat{\mathsf{ALG}}.$$

Proof. Let

$$\xi^{(\varepsilon)}(\vec{x}) = \xi(\vec{x}) + \varepsilon \sum_{s,s' \in \mathscr{S}} x_s x_{s'} + \varepsilon \sum_{s,s',s'' \in \mathscr{S}} x_s x_{s'} x_{s''}.$$

Then the preceding results show that optimal solutions $(\Phi^{(\varepsilon)}, p^{(\varepsilon)}, q_0^{(\varepsilon)})$ for $\xi^{(\varepsilon)}$ satisfy $p^{(\varepsilon)} \equiv 1$ and $q_0^{\varepsilon(\varepsilon)} = 0$. Taking a convergent subsequence $\Phi^{(\varepsilon)} \to \Phi^*$ as $\varepsilon \to 0$ in the space $\mathcal{M}$ (shown to be compact in Appendix 5.C.1) implies the result since $\widehat{\mathsf{ALG}}$ is continuous in ξ. $\square$

We first non-rigorously guess the solution by assuming it is of the form (5.90) and also solves the type II equation. By homogeneity, we may assume

$$\sum_{s \in \mathscr{S}} a_s b_s = 1. \tag{5.91}$$

Then

$$\begin{aligned}
\Phi_s'(q) &= b_s q^{b_s - 1}, \\
(\xi^s \circ \Phi)(q) &= \frac{a_s}{\lambda_s} q^{1-b_s}.
\end{aligned}$$

We thus expect that for some constant L independent of s,

$$\begin{aligned}
\Psi_s(q) &= b_s^{-1} q^{1-b_s} \frac{\mathrm{d}}{\mathrm{d}q} \sqrt{\frac{b_s q^{b_s-1}}{q^{1-b_s-1} a_s (1-b_s)/\lambda_s}} \\
&= \sqrt{\frac{\lambda_s}{a_s(1-b_s)b_s}} q^{1-b_s} \frac{\mathrm{d}}{\mathrm{d}q} q^{-\frac{1}{2}+b_s} \\
&= \left(-\frac{1}{2} + b_s\right) \sqrt{\frac{\lambda_s}{a_s(1-b_s)b_s}} q^{-1/2} \\
&= -L^{-1/2} q^{-1/2}.
\end{aligned}$$

(Recall that Ψ_s should be negative.) The resulting quadratic equation in b_s has solution

$$b_s = \frac{1 - \sqrt{\frac{a_s}{a_s + L\lambda_s}}}{2}. \tag{5.92}$$

Finally L is chosen to satisfy (5.91); it is easy to see there is a unique such choice.

Our next step is to verify the computation above and prove uniqueness.

Proof of Theorem 5.1.18.

Part 1: value of ALG Here we assume $p \equiv 1$, relying on Lemma 5.4.43, and determine the value ALG. Using the purity of ξ, a simple scaling argument shows the $\widehat{\mathsf{ALG}}$ value with endpoint $\vec{x} = (x_1, \dots, x_r)$ (cf. Remark 5.1.13) is given by

$$\widehat{\mathsf{ALG}}(\vec{x}) = \widehat{\mathsf{ALG}}(\vec{1}) \cdot \prod_{s \in \mathscr{S}} x_s^{a_s/2}. \tag{5.93}$$

(Recall that ξ is a covariance, hence the factor $1/2$ in the exponent on the right-hand side.) Set $\phi^s_{D-1} = 1 - b_s\delta$ for small δ and $\vec{b} \succeq 0$ satisfying (5.91). This is a fully general choice for ϕ_{D-1} as in Section 5.3. In light of Proposition 5.3.5, we obtain that for small $\delta > 0$,

$$\widehat{\mathsf{ALG}}(\vec{1}) = \max_{\vec{b}\,:\,(5.91)} \Big(\widehat{\mathsf{ALG}}(\phi_{D-1}) + \delta \sum_{s \in \mathscr{S}} \lambda_s \sqrt{\lambda_s^{-1} a_s b_s (1 - b_s)} \Big) + o(\delta). \tag{5.94}$$

Denoting $\widehat{\mathsf{ALG}} = \widehat{\mathsf{ALG}}(\vec{1})$ and using (5.93), we find

$$\begin{aligned}
\widehat{\mathsf{ALG}} &= \max_{\vec{b}\,:\,(5.91)} \Big(\widehat{\mathsf{ALG}} \cdot \prod_{s \in \mathscr{S}} (1 - b_s\delta)^{a_i/2} + \delta \sum_{s \in \mathscr{S}} \lambda_s \sqrt{\lambda_s^{-1} a_s b_s (1 - b_s)} \Big) + o(\delta) \\
&= \max_{\vec{b}\,:\,(5.91)} \Big(\Big(1 - \frac{\delta}{2}\Big) \widehat{\mathsf{ALG}} + \delta \sum_{s \in \mathscr{S}} \sqrt{\lambda_s a_s b_s (1 - b_s)} \Big) + o(\delta).
\end{aligned}$$

Rearranging and sending $\delta \to 0$ yields

$$\widehat{\mathsf{ALG}} = 2 \max_{\vec{b}\,:\,(5.91)} \sum_{s \in \mathscr{S}} \sqrt{\lambda_s a_s b_s (1 - b_s)}. \tag{5.95}$$

First, it is easy to see that any maximizing $\vec{b}^*$ has $b_s^* > 0$ for all s, since otherwise the derivative of the right-hand side in b_s would be infinite. By Lagrange multipliers, for some $C > 0$ any solution will have

$$\begin{aligned}
\sqrt{\frac{a_s}{L\lambda_s}} &= \frac{\mathrm{d}}{\mathrm{d}b_s} \Big(\sqrt{b_s(1 - b_s)} \Big) \\
&= \frac{\frac{1}{2} - b_s}{\sqrt{b_s(1 - b_s)}}
\end{aligned} \tag{5.96}$$

for some $L \in [0, \infty]$ (where division by ∞ gives 0).

Let us first assume $\sum_{s \in \mathscr{S}} a_s \geq 3$. Then (5.91) implies that $b_s < 1/2$ for some s, hence for all s since the signs have to match in (5.96). In particular we have $L < \infty$, and (5.92) above easily follows from (5.96). The resulting formula is as desired:

$$\begin{aligned}
\widehat{\mathsf{ALG}} &= 2 \sum_{s \in \mathscr{S}} \sqrt{\lambda_s a_s} \cdot \Big(\frac{1}{2} - b_s \Big) \sqrt{\frac{L\lambda_s}{a_s}} \\
&= \sum_{s \in \mathscr{S}} \lambda_s \sqrt{\frac{L a_s}{L\lambda_s + a_s}}.
\end{aligned}$$

The only remaining case is $\xi(x_1, x_2) = x_1 x_2$. Then it is clear from (5.95) that $b_1 = b_2 = 1/2$ and

$$\widehat{\mathsf{ALG}} = \sqrt{\lambda_1} + \sqrt{\lambda_2}.$$

(This case of Theorem 5.1.18 is stated with $b_1 = b_2 = 1$ which is an equivalent parametrization.)

Part 2: uniqueness assuming $p \equiv 1$ Next we show the optimal trajectory $\Phi^*(q) = (q^{b_1}, \dots, q^{b_r})$ is unique up to reparametrization when $p \equiv 1$. The maximization problem in (5.95) is strictly convex on the affine subspace defined by (5.91), and hence has a unique minimizer. It follows that if ϕ_d in the preceding equation is defined by any choice $\vec{b}$ bounded away from the optimal one, the obtained value would be strictly worse than ALG. In other words, any optimal trajectory where $p \equiv 1$ must satisfy $\Phi'(1) = \vec{b}$. By scale-invariance, we conclude that $\Phi(q) = (q^{b_1}, \dots, q^{b_r})$ is the unique optimal such trajectory.

Part 3: uniqueness of optimal p Finally we prove that all optimal solutions actually satisfy $p \equiv 1$. Suppose another maximizer (p, Φ) exists. Let

$$q_* = \inf_{q>0} \{q \ : \ \min_{s \in \mathscr{S}} \Phi_s(q) > 0\}.$$

The definition of p on $[0, q_*)$ is irrelevant so we assume without loss of generality that p is constant on $[0, q_*]$ and continuous at q_*. It is easy to see that such a maximizing p must be continuous on all of $[0, 1]$ and satisfy $p(1) = 1$; otherwise p could be strictly increased while keeping p' constant for the purposes of $\mathbb{A}$. The proof of Lemma 5.C.4 implies that p is uniformly Lipschitz on $[q_* + \varepsilon, 1]$ for any $\varepsilon > 0$, so that p' makes sense as a measurable function.

We have seen that if $p \equiv 1$ then ALG is achieved by a unique Φ, so we remains to show that no optimal (p, Φ) satisfies $p \not\equiv 1$ Assuming that $p \not\equiv 1$ we may choose $q > q_*$ a Lebesgue point for both p' and Φ' such that

$$p'(q) > 0.$$

We now derive a contradiction by expanding $\widehat{\mathsf{ALG}}$ around q as in (5.94). In particular, consider $\phi_d = \Phi(q-\delta)$ and $p_d = p(q - \delta)$. Let $\Delta_s = \Phi_s(q) - \phi_{d,s}$ and $\Delta_p = p(q) - p_d$. Since q is a Lebesgue point, we have $\Delta_s = \Phi_s'(q)\delta + o(\delta)$ and $\Delta_p = p'(q) + o(\delta)$.

The computation above for the value $\widehat{\mathsf{ALG}}$ implies

$$\widehat{\mathsf{ALG}}(p_d, \phi_d) = \widehat{\mathsf{ALG}}(\phi_d)\sqrt{p_d}.$$

Here $\widehat{\mathsf{ALG}}(p_d, \phi_d)$ denotes the analog of (5.7) with endpoint value $\Phi(q) = \phi_d$ rather than $q = 1^{\mathscr{S}}$, and $p(q) = p_d$. Therefore

$$\begin{aligned}
\widehat{\mathsf{ALG}}(p(q), \Phi(q)) &= \widehat{\mathsf{ALG}}(\phi_d)\sqrt{p_d} + \sum_{s \in \mathscr{S}} \lambda_s \sqrt{\Delta_s \left(\Delta_p \xi^s(\phi_d) + p_d \sum_{s' \in \mathscr{S}} \partial_{x_{s'}} \xi^s(\phi_d) \Delta_{s'} \right)} + o(\delta) \\
&= \widehat{\mathsf{ALG}}(p(q), \Phi(q)) \cdot \left(1 - \frac{\delta}{2} \times \left(\frac{p'(q)}{p(q)} + \sum_{s \in \mathscr{S}} \frac{a_s \Phi_s'(q)}{\Phi_s(q)} \right) \right) \\
&\quad + \delta \sum_{s \in \mathscr{S}} \lambda_s \sqrt{\Phi_s'(q) \left(p'(q)(\xi^s \circ \Phi)(q) + p(q) \sum_{s' \in \mathscr{S}} \partial_{x_{s'}} (\xi^s \circ \Phi)(q) \Phi_{s'}'(q) \right)} + o(\delta).
\end{aligned}$$

Rearranging and sending $\delta \to 0$ implies

$$\widehat{\mathsf{ALG}}(p(q), \Phi(q))/2 = \frac{\sum_{s \in \mathscr{S}} \lambda_s \sqrt{\Phi_s'(q) \left(p'(q)(\xi^s \circ \Phi)(q) + p(q) \sum_{s' \in \mathscr{S}} \partial_{x_{s'}} (\xi^s \circ \Phi)(q) \Phi_{s'}'(q) \right)}}{\frac{p'(q)}{p(q)} + \sum_{s \in \mathscr{S}} \frac{a_s \Phi_s'(q)}{\Phi_s(q)}}. \tag{5.97}$$

We claim that (5.97) forces $p'(q) = 0$, which completes the proof of uniqueness since q was an arbitrary choice of Lebesgue point. Note that from any solution to (5.97) we immediately get a maximizing (Φ, p) for $\mathbb{A}$ where $p(q)$ and each $\Phi_s(q)$ is a monomial of the form aq^b.

The right-hand side above has maximum value $\widehat{\mathsf{ALG}}(p(q), \Phi(q))/2$, and we already know from Lemma 5.4.43 there exists $(p'(q), \Phi'(q))$ achieving this value with $p'(q) = 0$. Supposing another maximizing $(\widetilde{p}'(q), \widetilde{\Phi}'(q))$ with $\widetilde{p}'(q) > 0$ exists, we suppress the input q and consider a general solution

$$(p_a', \Phi_a') = \left(ap_1' - (a-1)p_0', a\Phi_1' - (a-1)\Phi_0' \right).$$

We always restrict to a such that all derivatives are non-negative. The denominator of the right-hand side of (5.97) is affine in a, while Lemma 5.C.9 implies the numerator is concave. Since (p_0', Φ_0') and (p_1', Φ_1') both maximize the right-hand side we deduce that it takes the constant value $\widehat{\mathsf{ALG}}(p(q), \Phi(q))/2$ on (p_a', Φ_a') for all $a \in [0, 1]$. In particular using again Lemma 5.C.9 we find that each of the r terms in the numerator is actually a linear function of a on the interval such that

$$p_a'(q) \geq 0, \quad \text{and} \quad \min_s \Phi_{a,s}'(q) \geq 0. \tag{5.98}$$

This means equality is achieved for p_a for a satisfying (5.98) (even if $a > 1$) and implies that $\Phi'(q) \neq \widetilde{\Phi}'(q)$. Let $a_* > 0$ be the maximal value satisfying (5.98), so that $\min_s \Phi'_{a_*,s}(q) = \Phi'_{a_*,s_*}(q) = 0$. Then clearly the s_* term of the numerator is not affine on $a \in [a_* - \varepsilon, a_*]$; since Φ'_{a_*} satisfies admissibility it does not equal $\vec{0}$. This gives a contradiction, so we conclude that $p \equiv 1$ holds for all optimal (p, Φ). $\square$

Proof of Corollary 5.1.19 . Here we have $\lambda_s = \frac{a_s}{\sum_{s\in\mathscr{S}} a_s}$ in the preceding formulas. It is easy to see from (5.92) that the values b_s are all equal. From (5.91) we find $b_s = \frac{1}{\sum_{s\in\mathscr{S}} a_s}$ and so

$$
\begin{aligned}
\widehat{\mathsf{ALG}} &= 2\sum_{s\in\mathscr{S}} \sqrt{\lambda_s a_s b_s (1-b_s)} \\
&= 2\sum_{s\in\mathscr{S}} \frac{a_s}{\sqrt{\sum_{s\in\mathscr{S}} a_s}} \cdot \frac{\sqrt{\left(\sum_{s\in\mathscr{S}} a_s\right)-1}}{\sum_{s\in\mathscr{S}} a_s} \\
&= 2\sqrt{\frac{\left(\sum_{s\in\mathscr{S}} a_s\right)-1}{\sum_{s\in\mathscr{S}} a_s}}. \qquad \square
\end{aligned}
$$

We finally show Corollary 5.1.21, recalling the formula for E_∞ from [McK24] and verifying it equals ALG for pure models. It is given as follows, where $\mathbb{H} = \{z \in \mathbb{C} \ : \ \mathsf{Im}(z) > 0\}$ denotes the complex open upper half plane. We recall (a slight generalization of) [McK24, Lemma 2.2]; as written only the bipartite case was considered therein but the general multi-species case is no different. Additionally we point out that the constants α_s appearing in [McK24] continue to vanish in pure models for general r, which we take advantage of in the statement below.

Informally, the point below is simply that $\sum_s \lambda_s M_s$ is the Stieltjes transform of the bulk spectral distribution of an $N \times N$ random matrix with variance profile $\partial_{x_s,x_{s'}}\xi$ with diagonal species-dependent shift $E\xi^s(\vec{1})$. This essentially corresponds to the behavior of the Riemannian Hessian $\nabla^2_{\mathsf{sp}} H_N(\boldsymbol{\sigma})$ at a point $\boldsymbol{\sigma}$ with $H_N(\boldsymbol{\sigma}) = E$, where the diagonal shift corresponds to the induced radial derivative of H_N.

Proposition 5.4.44 (Adaptation of [McK24, Lemma 2.2] with r species and pure ξ)**.** *For $z \in \mathbb{H}$ (resp. $-\mathbb{H}$), there is a unique solution $\vec{M} \in \mathbb{H}^{\mathscr{S}}$ (resp. $-\mathbb{H}^{\mathscr{S}}$) to the matrix Dyson equation*

$$1 + M_s\left(\left(z - E\xi^s(\vec{1})\right) + \partial_{x_s}\xi^s(\vec{1})M_s + \sum_{s'\neq s}(\partial_{x_{s'}}\xi^s(\vec{1}))M_{s'}\right) = 0, \quad \forall s \in \mathscr{S}.$$

The threshold $E_\infty \geq 0$ is the smallest value such that with $z = 0$, $\vec{M}(E)$ extends analytically and continuously at the boundary to $E \in [E_\infty, \infty)$ (and is real-valued on this interval).

When $\xi(\vec{x}) = \prod_{s\in\mathscr{S}} x_s^{a_s}$ is pure and $z = 0$, the vector Dyson equation simplifies to

$$1 + a_s M_s\left(E - \lambda_s M_s + \sum_{s'\in\mathscr{S}} \lambda_{s'} a_{s'} M_{s'}\right) = 0, \quad \forall\, s \in \mathscr{S}. \tag{5.99}$$

Proof of Corollary 5.1.21. For convenience we omit the case $\xi(x_1, x_2) = x_1x_2$ and assume $\sum_s a_s \geq 3$. Setting

$$
\begin{aligned}
K_s &= a_s M_s, \\
K &= \sum_{s\in\mathscr{S}} \lambda_s K_s
\end{aligned}
$$

the system (5.99) can be rearranged to

$$A \equiv K + E = \frac{\lambda_s K_s}{a_s} - \frac{1}{K_s}, \quad \forall\, s \in \mathscr{S}.$$

With $B_s = \frac{a_s}{\lambda_s}$ we find that at $E = E_\infty$,

$$K_s = \frac{AB_s - \sqrt{A^2B_s^2 + 4B_s}}{2}.$$

Here the choice of sign is forced by $M_s < 0$; this easily holds for sufficiently large E (where one can give a power series expansion), and follows by continuity since $K_s \neq 0$ in general.

Note that A above determines each K_s, hence K and hence $E = A - K$. Viewing E as a function the A, its derivative must vanish and so:

$$\begin{aligned}
0 &= \frac{\mathsf{d}E}{\mathsf{d}A} \\
&= 1 - \frac{1}{2}\sum_{s\in\mathscr{S}} a_s \left(1 - \frac{AB_s}{\sqrt{A^2B_s^2 + 4B_s}}\right) && (5.100)\\
&= 1 - \frac{1}{2}\sum_{s\in\mathscr{S}} a_s \left(1 - \frac{1}{\sqrt{1 + 4/(A^2B_s)}}\right) && (5.101)\\
&\overset{(5.91)}{=} 1 - \frac{1}{2}\sum_{s\in\mathscr{S}} a_s \left(1 - \sqrt{\frac{a_s}{a_s + L\lambda_s}}\right) \\
&= 1 - \frac{1}{2}\sum_{s\in\mathscr{S}} a_s \left(1 - \sqrt{\frac{1}{1 + L/B_s}}\right). && (5.102)
\end{aligned}$$

Here we used $\sum_s a_s \geq 3$ to deduce from (5.100) that $A > 0$, thus implying the next line. By monotonicity, equality of (5.101) and (5.102) now implies $A = 2/\sqrt{L}$. Turning to the desired equality, we first write

$$\begin{aligned}
E_\infty &= A - K \\
&= \frac{2}{\sqrt{L}} - \frac{1}{2}\sum_s \lambda_s \left(\frac{2a_s}{\lambda_s\sqrt{L}} - \sqrt{\frac{4a_s^2}{L\lambda_s^2} + \frac{4a_s}{\lambda_s}}\right) \\
&= \frac{2}{\sqrt{L}} - \sum_s \frac{a_s}{\sqrt{L}}\left(1 - \sqrt{\frac{a_s + L\lambda_s}{a_s}}\right).
\end{aligned}$$

With $V_s \equiv \sqrt{a_s + L\lambda_s}$, adding and subtracting $\sum_s \frac{a_s^{3/2}}{V_s\sqrt{L}}$ to get the second equality, we compute

$$\begin{aligned}
E_\infty - \mathsf{ALG} &= \frac{2}{\sqrt{L}} + \sum_s \left(-\frac{a_s}{\sqrt{L}} + \frac{V_s\sqrt{a_s}}{\sqrt{L}} - \frac{\lambda_s\sqrt{La_s}}{V_s}\right) \\
&= \frac{1}{\sqrt{L}}\left(2 + \sum_s\left(-a_s + \frac{a_s^{3/2}}{V_s}\right)\right) + \sum_s \frac{\sqrt{a_s}}{V_s\sqrt{L}}\left(V_s^2 - a_s - L\lambda_s\right) \\
&= 0.
\end{aligned}$$

Here in the last step, we used (5.91) to handle the first contribution (summed over $s \in \mathscr{S}$) and the definition of V_s for the second (for each $s \in \mathscr{S}$). $\square$

Appendix

5.A Equivalence of BOGP and $\mathsf{BOGP}_{\mathrm{loc},0}$

In this section, we prove Proposition 5.2.9 that $\mathsf{BOGP} = \mathsf{BOGP}_{\mathrm{loc},0}$. We introduce two other limits $\mathsf{BOGP}_{\mathrm{den}}$ and $\mathsf{BOGP}_{\mathrm{loc}}$, as follows (restating BOGP and $\mathsf{BOGP}_{\mathrm{loc},0}$ for convenience).

$$\mathsf{BOGP} = \lim_{D\to\infty} \lim_{\eta\to 0} \lim_{k\to\infty} \sup_{\vec{\chi}\in\mathbb{I}(0,1)^{\mathscr{S}}} \inf_{\underline{\vec{\phi}}=\vec{\chi}(\underline{p})} \limsup_{N\to\infty} \frac{1}{N}\mathbb{E} \sup_{\underline{\boldsymbol{\sigma}}\in\mathcal{Q}(\eta)} \mathcal{H}_N(\underline{\boldsymbol{\sigma}}),$$

$$\mathsf{BOGP}_{\mathrm{den}} = \lim_{D\to\infty} \lim_{\eta\to 0} \lim_{k\to\infty} \sup_{\substack{\vec{\chi}\in\mathbb{I}(0,1)^{\mathscr{S}} \\ 1/D^2\text{-separated}}} \inf_{\substack{\underline{\vec{\phi}}=\vec{\chi}(\underline{p}) \\ 6r/D\text{-dense}}} \limsup_{N\to\infty} \frac{1}{N}\mathbb{E} \sup_{\underline{\boldsymbol{\sigma}}\in\mathcal{Q}(\eta)} \mathcal{H}_N(\underline{\boldsymbol{\sigma}}),$$

$$\mathsf{BOGP}_{\mathrm{loc}} = \lim_{D\to\infty} \lim_{\eta\to 0} \lim_{k\to\infty} \sup_{\substack{\vec{\chi}\in\mathbb{I}(0,1)^{\mathscr{S}} \\ 1/D^2\text{-separated}}} \inf_{\substack{\underline{\vec{\phi}}=\vec{\chi}(\underline{p}) \\ 6r/D\text{-dense}}} \limsup_{N\to\infty} \frac{1}{N}\mathbb{E} \sup_{\underline{\boldsymbol{\sigma}}\in\mathcal{Q}_{\mathrm{loc}}(\eta)} \mathcal{H}_N(\underline{\boldsymbol{\sigma}}),$$

$$\mathsf{BOGP}_{\mathrm{loc},0} = \lim_{D\to\infty} \lim_{k\to\infty} \sup_{\substack{\vec{\chi}\in\mathbb{I}(0,1)^{\mathscr{S}} \\ 1/D^2\text{-separated}}} \inf_{\substack{\underline{\vec{\phi}}=\vec{\chi}(\underline{p}) \\ 6r/D\text{-dense}}} \limsup_{N\to\infty} \frac{1}{N}\mathbb{E} \sup_{\underline{\boldsymbol{\sigma}}\in\mathcal{Q}_{\mathrm{loc}}(0)} \mathcal{H}_N(\underline{\boldsymbol{\sigma}}).$$

In the last three lines, the limits in k, η are clearly decreasing, but the limits in D are not, so the existence of these limits needs to be proven. Proposition 5.2.9 follows from the following propositions.

Proposition 5.A.1. *The limit* $\mathsf{BOGP}_{\mathrm{den}}$ *exists and* $\mathsf{BOGP} = \mathsf{BOGP}_{\mathrm{den}}$.

Proposition 5.A.2. *The limit* $\mathsf{BOGP}_{\mathrm{loc}}$ *exists and* $\mathsf{BOGP}_{\mathrm{den}} = \mathsf{BOGP}_{\mathrm{loc}}$.

Proposition 5.A.3. *The limit* $\mathsf{BOGP}_{\mathrm{loc},0}$ *exists and* $\mathsf{BOGP}_{\mathrm{loc}} = \mathsf{BOGP}_{\mathrm{loc},0}$.

5.A.1 Equivalence of BOGP and $\mathsf{BOGP}_{\mathrm{den}}$

Let $\underline{p}' = (p_0, \ldots, p_{D'})$ and $\underline{\vec{\phi}}' = (\vec{\phi}'_0, \ldots, \vec{\phi}'_{D'})$. Say $(\underline{p}', \underline{\vec{\phi}}')$ **refines** $(\underline{p}, \underline{\vec{\phi}})$ if there exists $0 \le a_0 < \cdots < a_D \le D'$ such that $(p_d, \vec{\phi}_d) = (p'_{a_d}, \vec{\phi}'_{a_d})$ for all $0 \le d \le D$.

Lemma 5.A.4. *The value* $\mathbb{E}\max_{\underline{\boldsymbol{\sigma}}\in\mathcal{Q}(\eta)} \mathcal{H}_N(\underline{\boldsymbol{\sigma}})$ *is decreasing under refinement. That is, if* $(\underline{p}', \underline{\vec{\phi}}')$ *refines* $(\underline{p}, \underline{\vec{\phi}})$*, then for any* k, η*,*

$$\mathbb{E} \max_{\underline{\boldsymbol{\sigma}}\in\mathcal{Q}^{k,D,\underline{\vec{\phi}}}(\eta)} \mathcal{H}_N^{k,D,\underline{p}}(\underline{\boldsymbol{\sigma}}) \ge \mathbb{E} \max_{\underline{\boldsymbol{\sigma}}\in\mathcal{Q}^{k,D',\underline{\vec{\phi}}'}(\eta)} \mathcal{H}_N^{k,D',\underline{p}'}(\underline{\boldsymbol{\sigma}}).$$

Proof. Let $I = \{a_0, \ldots, a_D\}$ and $J = [D'] \setminus I$. Define an equivalence relation $\bowtie$ on $\mathbb{L}(k, D')$ by $u \bowtie v$ if $u_d = v_d$ for all $d \in J$. Let

$$\mathcal{Q}' = \left\{ \underline{\boldsymbol{\sigma}} \in \mathcal{B}_N^{\mathbb{L}(k,D')} : \left\| \vec{R}(\boldsymbol{\sigma}(u^1), \boldsymbol{\sigma}(u^2)) - \vec{\phi}'_{u^1\wedge u^2} \right\|_\infty \le \eta,\ \forall u^1 \bowtie u^2 \right\}$$

be the superset of $\mathcal{Q}^{k,D',\underline{\vec{\phi}}'}$ where we only enforce the overlap constraint for $u^1 \bowtie u^2$. Then

$$\begin{aligned}
\mathbb{E} \max_{\underline{\boldsymbol{\sigma}} \in \mathcal{Q}^{k,D',\underline{\vec{\phi}}'}(\eta)} \mathcal{H}_N^{k,D',\underline{p}'}(\underline{\boldsymbol{\sigma}}) &\le \mathbb{E} \max_{\underline{\boldsymbol{\sigma}} \in \mathcal{Q}'} \mathcal{H}_N^{k,D',\underline{p}'}(\underline{\boldsymbol{\sigma}}) \\
&= \mathbb{E} \max_{\underline{\boldsymbol{\sigma}} \in \mathcal{Q}'} \frac{1}{k^{D'-D}} \sum_{u_J \in [k]^{D'-D}} \frac{1}{k^D} \sum_{u_I \in [k]^D} \mathcal{H}_N^{k,D',\underline{p}'}(\underline{\boldsymbol{\sigma}}) \\
&= \mathbb{E} \max_{\underline{\boldsymbol{\sigma}} \in \mathcal{Q}^{k,D,\underline{\vec{\phi}}}(\eta)} \mathcal{H}_N^{k,D,\underline{p}}(\underline{\boldsymbol{\sigma}}). \qquad \square
\end{aligned}$$

Proof of Proposition 5.A.1. Let $\mathsf{BOGP}^+_{\mathrm{den}}$ and $\mathsf{BOGP}^-_{\mathrm{den}}$ be $\mathsf{BOGP}_{\mathrm{den}}$ where the outer limit in D is replaced by lim sup and lim inf, respectively. We will separately prove $\mathsf{BOGP} \ge \mathsf{BOGP}^+_{\mathrm{den}}$ and $\mathsf{BOGP} \le \mathsf{BOGP}^-_{\mathrm{den}}$.

Fix any D, k, η, $1/D^2$-separated $\vec{\chi}$, and (not necessarily $6r/D$-dense) $\underline{p}, \underline{\vec{\phi}}$ satisfying $\underline{\vec{\phi}} = \vec{\chi}(\underline{p})$. Let $\delta = (r+1)/D$. Let $\widetilde{p}_0 = p_0$, and define a sequence $\widetilde{p}_1, \dots, \widetilde{p}_{\tilde{D}}$, where $\widetilde{p}_{d+1}$ is the smallest $p \in [\widetilde{p}_d, 1]$ such that

$$\max\left(p - \widetilde{p}_d, \|\vec{\chi}(p) - \vec{\chi}(\widetilde{p}_d)\|_\infty\right) \ge \delta$$

if such p exists. Let $\tilde{D}$ be the first index d such that no such p exists. Note that if $\Sigma_d = \widetilde{p}_d + \|\vec{\chi}(\widetilde{p}_d)\|_1$, then $0 \le \Sigma_d \le r+1$ and $\Sigma_{d+1} - \Sigma_d \ge \delta$ for all d. Thus $\tilde{D} \le (r+1)/\delta = D$.

Consider either $D' = 2D$ or $D' = 2D+1$. Let $\underline{p}'$ be the sorted union of $\{p_0, \dots, p_D\}$, $\{\widetilde{p}_1, \dots, \widetilde{p}_{\tilde{D}}\}$, and (if necessary) additional arbitrary $p \in [0,1]$, so that $\underline{p}'$ has length D'. Define $\underline{\vec{\phi}}' = \vec{\chi}(\underline{p}')$. Since $(\underline{p}', \underline{\vec{\phi}}')$ refines $(\underline{p}, \underline{\vec{\phi}})$, Lemma 5.A.4 implies

$$\mathbb{E} \max_{\underline{\boldsymbol{\sigma}} \in \mathcal{Q}^{k,D,\underline{\vec{\phi}}}(\eta)} \mathcal{H}_N(\underline{\boldsymbol{\sigma}}) \ge \mathbb{E} \max_{\underline{\boldsymbol{\sigma}} \in \mathcal{Q}^{k,D',\underline{\vec{\phi}}'}(\eta)} \mathcal{H}_N^{k,D',\underline{p}'}(\underline{\boldsymbol{\sigma}}).$$

Moreover, one can check that $\delta \le 6r/D'$, so $(\underline{p}', \underline{\vec{\phi}}')$ is $6r/D'$-dense. Thus, if $f(D)$ and $g(D)$ are the quantities inside the outer limits of BOGP and $\mathsf{BOGP}_{\mathrm{den}}$, we have shown $f(D) \ge g(2D), g(2D+1)$ (as taking the supremum over not necessarily $1/D^2$-separated $\vec{\chi}$ can only increase $f(D)$). This implies $\mathsf{BOGP} \ge \mathsf{BOGP}^+_{\mathrm{den}}$.

For the other direction, fix D, k, η and (not necessarily $1/D^2$-separated) $\vec{\chi}$. Define

$$\vec{\chi}'(p) = (1 - D^{-2})\vec{\chi}(p) + D^{-2}\vec{1},$$

so $\vec{\chi}'$ is $1/D^2$-separated. Consider any $6r/D$-dense $(\underline{p}, \underline{\vec{\phi}}')$ with $\underline{\vec{\phi}}' = \vec{\chi}'(\underline{p})$, and let $\underline{\vec{\phi}} = \vec{\chi}(\underline{p})$.

Let $\underline{\boldsymbol{\sigma}} \in \mathcal{Q}^{k,D,\underline{\vec{\phi}}}(\eta)$. Let $\boldsymbol{x}$ satisfy $\vec{R}(\boldsymbol{x}, \boldsymbol{x}) = \vec{1}$ and $\vec{R}(\boldsymbol{x}, \boldsymbol{\sigma}(u)) = \vec{0}$ for all $u \in \mathbb{L}$. Define

$$\boldsymbol{\rho}(u) = \sqrt{1 - D^{-2}}\boldsymbol{\sigma}(u) + D^{-1}\boldsymbol{x},$$

so that for all $u, v \in \mathbb{L}$,

$$\left\|\vec{R}(\boldsymbol{\rho}(u), \boldsymbol{\rho}(v)) - \vec{\phi}'_{u \wedge v}\right\|_\infty = (1 - D^{-2})\left\|\vec{R}(\boldsymbol{\sigma}(u), \boldsymbol{\sigma}(v)) - \vec{\phi}_{u \wedge v}\right\|_\infty \le \eta.$$

Thus $\underline{\boldsymbol{\rho}} \in \mathcal{Q}^{k,D,\underline{\vec{\phi}}'}(\eta)$, and we can easily check that

$$\frac{1}{\sqrt{N}}\|\boldsymbol{\rho}(u) - \boldsymbol{\sigma}(u)\|_2 = O(D^{-2})$$

for all $u \in \mathbb{L}$. By Proposition 5.1.23, with probability $1 - e^{-\Omega(N)}$ we have $H_N^{(u)} \in K_N$ for all $u \in \mathbb{L}$. On this event,

$$\left|\frac{1}{N}\mathcal{H}_N(\underline{\boldsymbol{\rho}}) - \frac{1}{N}\mathcal{H}_N(\underline{\boldsymbol{\sigma}})\right| \le CD^{-2}$$

for some $C > 0$, and so

$$\frac{1}{N} \sup_{\underline{\boldsymbol{\rho}} \in \mathcal{Q}^{k,D,\underline{\vec{\phi}}'}(\eta)} \mathcal{H}_N(\underline{\boldsymbol{\rho}}) + CD^{-2} \ge \frac{1}{N} \sup_{\underline{\boldsymbol{\sigma}} \in \mathcal{Q}^{k,D,\underline{\vec{\phi}}}(\eta)} \mathcal{H}_N(\underline{\boldsymbol{\sigma}}).$$

By Lemma 5.2.11, both sides of this inequality are subgaussian with fluctuations $O(N^{-1/2})$, so the contribution from the complement of this event is $o_N(1)$, and

$$\limsup_{N\to\infty} \frac{1}{N}\mathbb{E} \sup_{\underline{\boldsymbol{\rho}}\in\mathcal{Q}^{k,D,\underline{\vec{\phi}}'}(\eta)} \mathcal{H}_N(\underline{\boldsymbol{\rho}}) + CD^{-2} \ge \limsup_{N\to\infty} \frac{1}{N}\mathbb{E} \sup_{\underline{\boldsymbol{\sigma}}\in\mathcal{Q}^{k,D,\underline{\vec{\phi}}}(\eta)} \mathcal{H}_N(\underline{\boldsymbol{\sigma}}).$$

Thus $f(D) \le g(D) + CD^{-2}$ (as taking the infimum over $(\underline{p},\underline{\vec{\phi}})$ that are not necessarily the image of a $6r/D$-dense $(\underline{p},\underline{\vec{\phi}}')$ under the above transformation can only decrease $f(D)$). This implies $\mathsf{BOGP} \le \mathsf{BOGP}^-_{\text{den}}$. □

5.A.2 Equivalence of $\mathsf{BOGP}_{\text{den}}$ and $\mathsf{BOGP}_{\text{loc}}$

Lemma 5.A.5. *We have that $\mathcal{Q}(\eta) \subseteq \mathcal{Q}_{\text{loc}}(\eta + \frac{2}{k})$.*

Proof. Consider any $\underline{\boldsymbol{\sigma}} \in \mathcal{Q}(\eta)$. We define $\underline{\boldsymbol{\rho}} \in \mathcal{B}_N^{\mathbb{T}}$ by $\boldsymbol{\rho}(u) = \boldsymbol{\sigma}(u)$ if $u \in \mathbb{L}$, and

$$\boldsymbol{\rho}(u) = \frac{1}{k}\sum_{i=1}^k \boldsymbol{\rho}(ui)$$

for $u \in \mathbb{T}\setminus\mathbb{L}$. We will show that $\underline{\boldsymbol{\rho}} \in \mathcal{Q}_{\text{loc}+}(\eta + \frac{2}{k})$, and so $\underline{\boldsymbol{\sigma}} \in \mathcal{Q}_{\text{loc}}(\eta + \frac{2}{k})$.

Let $v \succeq u$ denote that v is a descendant of u in $\mathbb{T}$. Consider any non-leaf $u \in \mathbb{T}$ and two of its children ui, uj, for $i \neq j$. For any $s \in \mathscr{S}$,

$$|R_s(\boldsymbol{\rho}(ui),\boldsymbol{\rho}(uj)) - \phi^s_{|u|}| \le \frac{1}{k^{2(D-|u|)}} \sum_{\substack{v,v'\in\mathbb{L}\\ v\succeq ui, v'\succeq uj}} |R_s(\boldsymbol{\sigma}(v),\boldsymbol{\sigma}(v')) - \phi^s_{|u|}| \le \eta. \tag{5.103}$$

Moreover,

$$|R_s(\boldsymbol{\rho}(ui),\boldsymbol{\rho}(u)) - \phi^s_{|u|}| \le \frac{1}{k}\sum_{j=1}^k |R_s(\boldsymbol{\sigma}(ui),\boldsymbol{\sigma}(uj)) - \phi^s_{|u|}| \le \eta + \frac{2}{k},$$

where we bounded the terms $j \neq i$ by (5.103) and the term $j = i$ crudely by 2. Thus,

$$|R_s(\boldsymbol{\rho}(u),\boldsymbol{\rho}(u)) - \phi^s_{|u|}| \le \frac{1}{k}\sum_{i=1}^k |R_s(\boldsymbol{\sigma}(ui),\boldsymbol{\sigma}(u)) - \phi^s_{|u|}| \le \eta + \frac{2}{k}.$$ □

For $k' \le k$, define a k'-*ary subtree* of $\mathbb{T}$ to be a subset $T \subseteq \mathbb{T}$ isometric to $\mathbb{T}(k', D)$. The following fact is clear from the definition of $\mathcal{Q}_{\text{loc}}(\eta)$.

Fact 5.A.6. *Let $T \subseteq \mathbb{T}$ be a k'-ary subtree with leaf set L. If $\underline{\boldsymbol{\sigma}} \in \mathcal{Q}_{\text{loc}}(\eta)$, then $(\boldsymbol{\sigma}(u))_{u\in L} \in \mathcal{Q}^{k',D,\underline{\vec{\phi}}}_{\text{loc}}(\eta)$.*

Proof. There exists $\underline{\boldsymbol{\rho}} \in \mathcal{Q}_{\text{loc}+}(\eta)$ such that $\boldsymbol{\rho}(u) = \boldsymbol{\sigma}(u)$ for all $u \in \mathbb{L}$. Then $(\boldsymbol{\rho}(u))_{u\in T} \in \mathcal{Q}^{k',D,\underline{\vec{\phi}}}_{\text{loc}+}(\eta)$, which implies the result. □

Lemma 5.A.7. *Let k' be the largest integer solution to $D(k')^D \le \min(\sqrt{k}, \eta^{-1})$. If $\underline{\boldsymbol{\sigma}} \in \mathcal{Q}_{\text{loc}}(\eta)$, there exists a k'-ary subtree T of $\mathbb{T}$ with leaf set L such that $(\boldsymbol{\sigma}(u))_{u\in L} \in \mathcal{Q}^{k',D,\underline{\vec{\phi}}}(CD^2(k^{-1/4} + \eta^{1/4}))$, for some $C > 0$.*

Proof. Let $\underline{\boldsymbol{\rho}} \in \mathcal{Q}_{\text{loc}+}(\eta)$ such that $\boldsymbol{\rho}(u) = \boldsymbol{\sigma}(u)$ for all $u \in \mathbb{L}$. We will construct T by a breadth-first search: we start from $T = \{\emptyset\}$ and each step *process* a leaf u of T by adding k' children of u to T, until all leaves of T are of depth D.

Suppose we are currently processing vertex u. Let $V = \{\boldsymbol{\rho}(v) : v \in T\}$ and $S = \text{span}(V)$; note that $|V| \le D(k')^D \le \min(\sqrt{k}, \eta^{-1})$. Let P_S denote the projection operator onto S. For $i \in [k]$, write $\boldsymbol{x}^i = \frac{1}{\sqrt{N}}(\boldsymbol{\rho}(ui) - \boldsymbol{\rho}(u))$, and note $\|\boldsymbol{x}^i\|_2 \le 2$. Then

$$\frac{1}{N}\sum_{i=1}^k \|P_S(\boldsymbol{\rho}(ui) - \boldsymbol{\rho}(u))\|_2^2 = \sum_{i=1}^k \|P_S\boldsymbol{x}^i\|_2^2$$

is upper bounded by the sum of the top $|V|$ eigenvalues of the Gram matrix $\boldsymbol{M} = (\langle \boldsymbol{x}^i, \boldsymbol{x}^j \rangle)_{i,j=1}^k$. However, for $i \neq j$,

$$|\langle \boldsymbol{x}^i, \boldsymbol{x}^j \rangle| = \frac{1}{N} |\langle \boldsymbol{\rho}(ui) - \boldsymbol{\rho}(u), \boldsymbol{\rho}(uj) - \boldsymbol{\rho}(u) \rangle| \le \sum_{s \in \mathscr{S}} \lambda_s |R_s(\boldsymbol{\rho}(ui) - \boldsymbol{\rho}(u), \boldsymbol{\rho}(uj) - \boldsymbol{\rho}(u))| \le 4\eta,$$

while $|\langle \boldsymbol{x}^i, \boldsymbol{x}^i \rangle| \le 4$. So, if we let $\boldsymbol{M} = \boldsymbol{D} + \boldsymbol{A}$ where $\boldsymbol{D} = \mathrm{diag}(\boldsymbol{M})$, and let $a_1 \ge \cdots \ge a_{|V|}$ be the top $|V|$ eigenvalues of $\boldsymbol{A}$, then the sum of the top $|V|$ eigenvalues of $\boldsymbol{M}$ is upper bounded by $4|V| + \sum_{i=1}^{|V|} a_i$. However,

$$\sum_{i=1}^{|V|} a_i \le \sqrt{|V| \sum_{i=1}^{|V|} a_i^2} \le \sqrt{|V| \|\boldsymbol{A}\|_F^2} \le 4k\eta\sqrt{|V|}.$$

It follows that

$$\frac{1}{kN} \sum_{i=1}^{k} \|P_S(\boldsymbol{\rho}(ui) - \boldsymbol{\rho}(u))\|_2^2 \le \frac{4|V|}{k} + 4\eta\sqrt{|V|} \le 4(k^{-1/2} + \eta^{1/2}),$$

where the last step follows from $|V| \le \min(\sqrt{k}, \eta^{-1})$. Thus there are k' children ui of u such that

$$\frac{1}{\sqrt{N}} \|P_S(\boldsymbol{\rho}(ui) - \boldsymbol{\rho}(u))\|_2 \le 3(k^{-1/4} + \eta^{1/4}).$$

We choose these as the children of u in T. By constructing T in this manner, we get that for all distinct edges (u, ui), (v, vj) in T,

$$\frac{1}{N} |\langle \boldsymbol{\sigma}(ui) - \boldsymbol{\sigma}(u), \boldsymbol{\sigma}(vj) - \boldsymbol{\sigma}(v) \rangle|, \frac{1}{N} |\langle \boldsymbol{\sigma}(ui) - \boldsymbol{\sigma}(u), \boldsymbol{\sigma}(\emptyset) \rangle| \le 6(k^{-1/4} + \eta^{1/4}).$$

whence

$$\left\| \vec{R}(\boldsymbol{\sigma}(ui), \boldsymbol{\sigma}(u), \boldsymbol{\sigma}(vj) - \boldsymbol{\sigma}(v)) \right\|_\infty, \left\| \vec{R}(\boldsymbol{\sigma}(ui), \boldsymbol{\sigma}(u), \boldsymbol{\sigma}(\emptyset)) \right\|_\infty \le \frac{6(k^{-1/4} + \eta^{1/4})}{\min_s \lambda_s}.$$

We now verify that $(\boldsymbol{\sigma}(u))_{u \in L} \in \mathcal{Q}^{k', D, \underline{\vec{\phi}}}(CD^2(k^{-1/4} + \eta^{1/4}))$. Consider any $u, v \in L$ with least common ancestor w, and let $|w| = d$. Let $(u_0, \ldots, u_{D-d})$ and $(v_0, \ldots, v_{D-d})$ be the paths from w to u, v, with $u_0 = v_0 = w$ and $u_{D-d} = u$, $v_{D-d} = v$, and let $(w_0, \ldots, w_d)$ be the path from $\emptyset$ to w, with $w_0 = \emptyset$, $w_d = w$. Also define as convention $\boldsymbol{\sigma}(w_{-1}) = \boldsymbol{0}$. Then,

$$\begin{aligned}
\left\| \vec{R}(\boldsymbol{\sigma}(u), \boldsymbol{\sigma}(v)) - \vec{\phi}_d \right\|_\infty &\le \left\| \vec{R}(\boldsymbol{\sigma}(w), \boldsymbol{\sigma}(w)) - \vec{\phi}_d \right\|_\infty + \sum_{i=1}^{D-d} \sum_{\ell=0}^{d} \left\| \vec{R}(\boldsymbol{\sigma}(w_\ell) - \boldsymbol{\sigma}(w_{\ell-1}), \boldsymbol{\sigma}(u_i) - \boldsymbol{\sigma}(u_{i-1})) \right\|_\infty \\
&\quad + \sum_{j=1}^{D-d} \sum_{\ell=0}^{d} \left\| \vec{R}(\boldsymbol{\sigma}(w_\ell) - \boldsymbol{\sigma}(w_{\ell-1}), \boldsymbol{\sigma}(v_j) - \boldsymbol{\sigma}(v_{j-1})) \right\|_\infty \\
&\quad + \sum_{i,j=1}^{D-d} \left\| \vec{R}(\boldsymbol{\sigma}(u_i) - \boldsymbol{\sigma}(u_{i-1}), \boldsymbol{\sigma}(v_j) - \boldsymbol{\sigma}(v_{j-1})) \right\|_\infty \\
&\le CD^2(k^{-1/4} + \eta^{1/4}).
\end{aligned}$$

□

Lemma 5.A.8. *There exists $C > 0$ such that with probability $1 - e^{-\Omega(N)}$ over the Hamiltonians $H_N^{(u)}$ the following holds. If $\varepsilon > 0$, $\underline{\boldsymbol{\sigma}} \in \mathcal{Q}_{\mathrm{loc}}(\eta)$, and*

$$\frac{1}{N} \mathcal{H}(\underline{\boldsymbol{\sigma}}) \ge E,$$

then for $k' = \lfloor k\varepsilon / 3CD \rfloor$, there exists a k'-ary subtree T of $\mathbb{T}$ with leaf set L such that

$$\frac{1}{N} H_N^{(u)}(\boldsymbol{\sigma}(u)) \ge E - \varepsilon$$

for all $u \in L$.

Proof. We consider the event that $H_N^{(u)} \in K_N$ for all $u \in \mathbb{L}$, for K_N defined in Proposition 5.1.23. This holds with probability $1 - e^{-\Omega(N)}$, and on this event, $|H_N^{(u)}(\boldsymbol{\sigma}(u))| \le C$ for all $u \in \mathbb{L}$. For $u \in \mathbb{T}$ define

$$F(u) = \frac{1}{Nk^{D-|u|}} \sum_{\substack{v \in \mathbb{L} \\ v \succeq u}} H^{(v)}(\boldsymbol{\sigma}(v)).$$

We will show that for any $u \in \mathbb{T} \setminus \mathbb{L}$, we may find k' distinct children $ui_1, \ldots, ui_{k'}$ such that $F(ui_j) \ge F(u) - \varepsilon/D$ for all j. Indeed, we have

$$F(u) = \frac{1}{k} \sum_{i=1}^{k} F(ui),$$

and $|F(ui)| \le C$ for all i, so the claim follows from Markov's inequality.

We construct the subtree T recursively starting from $\emptyset$, using the above claim to select the k' children of each node. Thus, for all $u, ui \in T$ with ui a child of u, we have $F(ui) \ge F(u) - \varepsilon/D$. Since $F(\emptyset) = \frac{1}{N}\mathcal{H}_N(\underline{\boldsymbol{\sigma}}) \ge E$, the result follows. $\square$

Proof of Proposition 5.A.2. Let $\mathsf{BOGP}^+_{\mathrm{loc}}$ and $\mathsf{BOGP}^-_{\mathrm{loc}}$ be $\mathsf{BOGP}_{\mathrm{loc}}$ where the outer limit in D is replaced by $\limsup$ and $\liminf$, respectively. Lemma 5.A.5 gives $\mathsf{BOGP}_{\mathrm{den}} \le \mathsf{BOGP}^-_{\mathrm{loc}}$, so it suffices to prove $\mathsf{BOGP}_{\mathrm{den}} \ge \mathsf{BOGP}^+_{\mathrm{loc}}$.

Fix arbitrary $\varepsilon > 0$, D, k, η, $1/D^2$-dense $\vec{\chi}$, and $6r/D$-dense $(\underline{p}, \vec{\underline{\phi}})$ satisfying $\vec{\underline{\phi}} = \vec{\chi}(\underline{p})$. If $\underline{\boldsymbol{\sigma}} \in \mathcal{Q}_{\mathrm{loc}}(\eta)$ and $\frac{1}{N}\mathcal{H}_N(\underline{\boldsymbol{\sigma}}) \ge E$, then on an event with probability $1 - e^{-\Omega(N)}$, Lemma 5.A.8 gives a k'-ary subtree $T \subseteq \mathbb{T}$ with leaf set L such that $\frac{1}{N}H_N^{(u)}(\boldsymbol{\sigma}(u)) \ge E - \varepsilon$ for all $u \in L$. However, $(\boldsymbol{\sigma}(u))_{u \in L}$ is itself an element of $\mathcal{Q}_{\mathrm{loc}}^{k',D,\vec{\underline{\phi}}}(\eta)$ by Fact 5.A.6, so Lemma 5.A.7 gives a k''-ary subtree $T' \subseteq T$ with leaf set L' such that $(\boldsymbol{\sigma}(u))_{u \in L'} \in \mathcal{Q}^{k'',D,\vec{\underline{\phi}}}(\eta')$. Here $k' = \lfloor \varepsilon/3CD \rfloor$, k'' is the largest solution to $D(k'')^D \le \min(\sqrt{k'}, \eta^{-1})$, and $\eta' = CD^2((k')^{-1/4} + \eta^{1/4})$.

It follows that for all E,

$$\mathbb{P}\left[\frac{1}{N} \sup_{\underline{\boldsymbol{\sigma}} \in \mathcal{Q}^{k'',D,\vec{\underline{\phi}}}(\eta')} \mathcal{H}_N^{k'',D,\underline{p}}(\underline{\boldsymbol{\sigma}}) \ge E - \varepsilon\right] \ge \binom{k}{k''}^{-D} \mathbb{P}\left[\frac{1}{N} \sup_{\underline{\boldsymbol{\sigma}} \in \mathcal{Q}_{\mathrm{loc}}^{k,D,\vec{\underline{\phi}}}(\eta)} \mathcal{H}_N^{k,D,\underline{p}}(\underline{\boldsymbol{\sigma}}) \ge E\right] - e^{-\Omega(N)}.$$

By Lemma 5.2.11, the random variables in these two probabilities are both subgaussian with fluctuations $O(N^{-1/2})$. So

$$\limsup_{N\to\infty} \frac{1}{N}\mathbb{E} \sup_{\underline{\boldsymbol{\sigma}} \in \mathcal{Q}^{k'',D,\vec{\underline{\phi}}}(\eta')} \mathcal{H}_N^{k'',D,\underline{p}}(\underline{\boldsymbol{\sigma}}) + \varepsilon \ge \limsup_{N\to\infty} \frac{1}{N}\mathbb{E} \sup_{\underline{\boldsymbol{\sigma}} \in \mathcal{Q}_{\mathrm{loc}}^{k,D,\vec{\underline{\phi}}}(\eta)} \mathcal{H}_N^{k,D,\underline{p}}(\underline{\boldsymbol{\sigma}}).$$

For fixed D, as $k \to \infty$ and $\eta \to 0$, we have $k'' \to \infty$ and $\eta' \to 0$. Then taking $D \to \infty$ shows $\mathsf{BOGP}_{\mathrm{den}} + \varepsilon \ge \mathsf{BOGP}^+_{\mathrm{loc}}$. Since ε was arbitrary, the result follows. $\square$

Remark 5.A.9. A byproduct of Lemma 5.A.8 is that defining $\mathcal{H}_N$ as the **minimum** over $u \in \mathbb{L}$ of the energies $H_N^{(u)}(\boldsymbol{\sigma}(u))$, and BOGP in terms of this $\mathcal{H}_N$, gives the same threshold as our definition (5.30) of $\mathcal{H}_N$ as the average of these energies. The minimal energy is actually more directly connected to our proof of Theorem 5.2.3, as seen in the definition (5.34) of S_{solve}. However the average energy is more convenient for our analysis in Section 5.3.

5.A.3 Equivalence of $\mathsf{BOGP}_{\mathrm{loc}}$ and $\mathsf{BOGP}_{\mathrm{loc},0}$

Lemma 5.A.10. *Let $k \in \mathbb{N}$, $0 < q_0 \le q \le 1$ and $q', \varepsilon \in [0,1]$. There exists $\varepsilon' = \varepsilon'(\varepsilon, k, q_0)$, where $\varepsilon' \to 0$ as $\varepsilon \to 0$ for fixed k, q_0, such that the following holds for all q, q'. Suppose that $\boldsymbol{x}, \boldsymbol{y}^1, \ldots, \boldsymbol{y}^k \in \mathbb{R}^N$ and*

$$\boldsymbol{Y} = \begin{bmatrix} \boldsymbol{x} & \boldsymbol{y}^1 & \cdots & \boldsymbol{y}^k \end{bmatrix}$$

satisfies $\boldsymbol{Y}^\top \boldsymbol{Y} = D + E$*, where* $D = \mathrm{diag}(q, q', \ldots, q')$*, all entries of* E *have magnitude at most* ε*, and* $E_{1,1} = 0$*. There exist* $\boldsymbol{z}^1, \ldots, \boldsymbol{z}^k$ *such that for*

$$\boldsymbol{Z} = \begin{bmatrix} \boldsymbol{x} & \boldsymbol{z}^1 & \cdots & \boldsymbol{z}^k \end{bmatrix},$$

we have $\boldsymbol{Z}^\top \boldsymbol{Z} = D$ *and* $\left\|\boldsymbol{z}^i - \boldsymbol{y}^i\right\|_2 \le \varepsilon'$ *for all* $i \in [k]$*.*

Proof. We will take

$$\varepsilon' = \begin{cases} 2 & k^2\varepsilon \ge q_0, \\ 3k^{3/2}\varepsilon^{1/2} & \text{otherwise.} \end{cases}$$

If $k^3\varepsilon \ge q_0$, we let $\boldsymbol{z}^1, \ldots, \boldsymbol{z}^k$ be any orthogonal vectors of norm $\sqrt{q'}$ orthogonal to $\boldsymbol{x}$ and each other. As $\left\|\boldsymbol{y}^i\right\|_2, \left\|\boldsymbol{z}^i\right\|_2 \le 1$, the result follows. Similarly, if $k^3\varepsilon \ge q'$, then

$$\left\|\boldsymbol{y}^i - \boldsymbol{z}^i\right\|_2 \le \left\|\boldsymbol{y}^i\right\|_2 + \left\|\boldsymbol{z}^i\right\|_2 = \sqrt{q' + \varepsilon} + \sqrt{q'} \le 3k^{3/2}\varepsilon^{1/2}.$$

It remains to address the case $k^3\varepsilon \le \min(q_0, q')$. We define $\boldsymbol{z}^1, \ldots, \boldsymbol{z}^k$ by the Gram-Schmidt algorithm, i.e.

$$\widehat{\boldsymbol{z}}^i = \boldsymbol{y}^i - \frac{\langle \boldsymbol{y}^i, \boldsymbol{x}\rangle}{\|\boldsymbol{x}\|_2^2}\boldsymbol{x} - \sum_{j=1}^{i-1} \frac{\langle \boldsymbol{y}^i, \boldsymbol{z}^j\rangle}{\|\boldsymbol{z}^j\|_2}\boldsymbol{z}^j, \qquad \boldsymbol{z}^i = \frac{\sqrt{q'}}{\left\|\widehat{\boldsymbol{z}}^i\right\|_2}\widehat{\boldsymbol{z}}^i.$$

Let $\varepsilon_i = \varepsilon(1 + 3k^{-2})^i$, and note that $\varepsilon \le \varepsilon_i \le 2\varepsilon$ for all $0 \le i \le k$. We will show by induction over i that for all $j \le i < \ell$,

$$|\langle \boldsymbol{y}^\ell, \boldsymbol{z}^j\rangle| \le \varepsilon_i, \tag{5.104}$$

where as the base case this vacuously holds for $i = 0$. Suppose the inductive hypothesis holds for $i - 1$. It suffices to prove (5.104) for $j = i$ because the assertion for the remaining j is implied by the inductive hypothesis, as $\varepsilon_{i-1} \le \varepsilon_i$. We have

$$\left\|\widehat{\boldsymbol{z}}^i\right\|_2^2 = \left\|\boldsymbol{y}^i\right\|_2^2 - \frac{\langle \boldsymbol{y}^i, \boldsymbol{x}\rangle^2}{\|\boldsymbol{x}\|_2^2} - \sum_{j=1}^{i-1} \frac{\langle \boldsymbol{y}^i, \boldsymbol{z}^j\rangle^2}{\|\boldsymbol{z}^j\|_2^2}$$

so

$$\left|\frac{\left\|\widehat{\boldsymbol{z}}^i\right\|_2^2}{q'} - 1\right| \le \left|\frac{\left\|\boldsymbol{y}^i\right\|_2^2}{q'} - 1\right| + \frac{\langle \boldsymbol{y}^i, \boldsymbol{x}\rangle^2}{q'\|\boldsymbol{x}\|_2^2} + \sum_{j=1}^{i-1} \frac{\langle \boldsymbol{y}^i, \boldsymbol{z}^j\rangle^2}{q'\|\boldsymbol{z}^j\|_2^2} \le \frac{\varepsilon_{i-1}}{q'} + \frac{\varepsilon_{i-1}^2}{q'q_0} + \frac{k\varepsilon_{i-1}^2}{(q')^2} \le \frac{2}{k^3}.$$

Thus $\left\|\widehat{\boldsymbol{z}}^i\right\|_2 \ge \sqrt{q'}(1 - 2k^{-3})$. For any $\ell > i$,

$$\begin{aligned} |\langle \widehat{\boldsymbol{z}}^i, \boldsymbol{y}^\ell\rangle| &\le |\langle \boldsymbol{y}^i, \boldsymbol{y}^\ell\rangle| + \frac{|\langle \boldsymbol{y}^i, \boldsymbol{x}\rangle||\langle \boldsymbol{x}, \boldsymbol{y}^\ell\rangle|}{\|\boldsymbol{x}\|_2^2} + \sum_{j=1}^{i-1} \frac{|\langle \boldsymbol{y}^i, \boldsymbol{z}^j\rangle||\langle \boldsymbol{z}^j, \boldsymbol{y}^\ell\rangle|}{\|\boldsymbol{z}^j\|_2^2} \\ &\le \varepsilon_{i-1} + \frac{\varepsilon_{i-1}^2}{q_0} + \frac{k\varepsilon_{i-1}^2}{q_0'} \le \varepsilon_{i-1}\left(1 + \frac{2}{k^2}\right) \end{aligned}$$

Thus

$$|\langle \boldsymbol{z}^i, \boldsymbol{y}^\ell\rangle| \le \varepsilon_{i-1} \cdot \frac{1 + 2k^{-2}}{1 - 2k^{-3}} \le \varepsilon_i,$$

completing the induction. Finally, note that

$$\left\|\widehat{\boldsymbol{z}}^i - \boldsymbol{y}^j\right\|_2^2 = \frac{\langle \boldsymbol{y}^i, \boldsymbol{x}\rangle^2}{\|\boldsymbol{x}\|_2^2} + \sum_{j=1}^{i-1} \frac{\langle \boldsymbol{y}^i, \boldsymbol{z}^j\rangle^2}{\|\boldsymbol{z}^j\|_2^2} \le \frac{\varepsilon_{i-1}^2}{q_0} + \frac{k\varepsilon_{i-1}^2}{q'} \le \frac{5\varepsilon}{k^2}$$

and

$$\left| \left\| \widehat{\boldsymbol{z}}^i \right\|_2 - \sqrt{q'} \right| \le \frac{\left| \left\| \widehat{\boldsymbol{z}}^i \right\|_2^2 - q' \right|}{\sqrt{q'}} \le \varepsilon_{i-1} + \frac{\varepsilon_{i-1}^2}{q' q_0} + \frac{k\varepsilon_{i-1}^2}{(q')^2} \le 3\varepsilon.$$

Thus

$$\left\| \boldsymbol{z}^i - \boldsymbol{y}^i \right\|_2 \le \left\| \widehat{\boldsymbol{z}}^i - \boldsymbol{y}^i \right\|_2 + \left| \left\| \widehat{\boldsymbol{z}}^i \right\|_2 - \sqrt{q'} \right| \le \frac{\sqrt{5\varepsilon}}{k} + 3\varepsilon \le \varepsilon'. \qquad \square$$

Proof of Proposition 5.A.3. Let $\mathsf{BOGP}^+_{\mathrm{loc},0}$ and $\mathsf{BOGP}^-_{\mathrm{loc},0}$ be $\mathsf{BOGP}_{\mathrm{loc},0}$ where the outer limit in D is replaced by $\limsup$ and $\liminf$, respectively. It is clear that $\mathsf{BOGP}_{\mathrm{loc}} \ge \mathsf{BOGP}^+_{\mathrm{loc},0}$, so it suffices to prove $\mathsf{BOGP}_{\mathrm{loc}} \le \mathsf{BOGP}^-_{\mathrm{loc},0}$.

Fix D, k, η, $1/D^2$-separated $\vec{\chi}$, and $6r/D$-dense $(\underline{p}, \underline{\vec{\phi}})$ with $\underline{\vec{\phi}} = \vec{\chi}(\underline{p})$. Consider $\underline{\boldsymbol{\sigma}} \in \mathcal{Q}_{\mathrm{loc}}(\eta)$ and let $\underline{\boldsymbol{\rho}} \in \mathcal{Q}_{\mathrm{loc}+}(\eta)$ such that $(\boldsymbol{\rho}(u))_{u\in\mathbb{L}} = \underline{\boldsymbol{\sigma}}$. Define $\varepsilon_0 = \eta D$ and $\varepsilon_d = \varepsilon'(6\varepsilon_{d-1} + 4\eta, k, D^{-2})$ for $1 \le d \le D$, where ε' is given by Lemma 5.A.10. We will now construct $\underline{\boldsymbol{\tau}} \in \mathcal{Q}_{\mathrm{loc}+}(0)$ approximating $\underline{\boldsymbol{\rho}}$ in the sense that for all $u \in \mathbb{T}, s \in \mathscr{S}$,

$$\sqrt{R_s(\boldsymbol{\tau}(u) - \boldsymbol{\rho}(u), \boldsymbol{\tau}(u) - \boldsymbol{\rho}(u))} \le \varepsilon_{|u|}. \tag{5.105}$$

We define $\boldsymbol{\tau}(\emptyset)$ by

$$\boldsymbol{\tau}(\emptyset)_s = \boldsymbol{\rho}(\emptyset)_s \sqrt{\frac{\phi_0^s}{R_s(\boldsymbol{\rho}(\emptyset), \boldsymbol{\rho}(\emptyset))}}$$

for all $s \in \mathscr{S}$. Thus $R_s(\boldsymbol{\tau}(\emptyset), \boldsymbol{\tau}(\emptyset)) = \phi_0^s$ and

$$\sqrt{R_s(\boldsymbol{\tau}(\emptyset) - \boldsymbol{\rho}(\emptyset), \boldsymbol{\tau}(\emptyset) - \boldsymbol{\rho}(\emptyset))} = \left| \sqrt{R_s(\boldsymbol{\rho}(\emptyset), \boldsymbol{\rho}(\emptyset))} - \sqrt{\phi_0^s} \right| \le \frac{\eta}{\sqrt{\phi_0^s}} \le \varepsilon_0,$$

where the second-last inequality holds for all sufficiently small $\eta > 0$. This proves (5.105) for $u = \emptyset$. We construct $\boldsymbol{\tau}(u)$ for the remaining $u \in \mathbb{T}$ recursively. Suppose we have constructed $\boldsymbol{\tau}(u)$ satisfying (5.105). Then, for each $s \in \mathscr{S}$, $i, j \in [k]$,

$$\begin{aligned} R_s(\boldsymbol{\rho}(ui) - \boldsymbol{\tau}(u), \boldsymbol{\rho}(uj) - \boldsymbol{\tau}(u)) = &\, R_s(\boldsymbol{\rho}(ui) - \boldsymbol{\rho}(u), \boldsymbol{\rho}(uj) - \boldsymbol{\rho}(u)) \\ &+ R_s(\boldsymbol{\rho}(ui) - \boldsymbol{\rho}(u), \boldsymbol{\rho}(u) - \boldsymbol{\tau}(u)) \\ &+ R_s(\boldsymbol{\rho}(u) - \boldsymbol{\tau}(u), \boldsymbol{\rho}(uj) - \boldsymbol{\rho}(u)) \\ &+ R_s(\boldsymbol{\rho}(u) - \boldsymbol{\tau}(u), \boldsymbol{\rho}(u) - \boldsymbol{\tau}(u)), \end{aligned}$$

so

$$|R_s(\boldsymbol{\rho}(ui) - \boldsymbol{\tau}(u), \boldsymbol{\rho}(uj) - \boldsymbol{\tau}(u)) - (\phi^s_{|u|+1} - \phi^s_{|u|})\mathbf{1}\{i = j\}| \le 6\varepsilon_{|u|} + 4\eta.$$

Similarly,

$$\begin{aligned} |R_s(\boldsymbol{\rho}(ui) - \boldsymbol{\tau}(u), \boldsymbol{\tau}(u))| = &\, |R_s(\boldsymbol{\rho}(ui) - \boldsymbol{\rho}(u), \boldsymbol{\rho}(u))| \\ &+ |R_s(\boldsymbol{\rho}(u) - \boldsymbol{\tau}(u), \boldsymbol{\rho}(u))| \\ &+ |R_s(\boldsymbol{\rho}(ui) - \boldsymbol{\rho}(u), \boldsymbol{\tau}(u) - \boldsymbol{\rho}(u))| \\ &+ |R_s(\boldsymbol{\rho}(u) - \boldsymbol{\tau}(u), \boldsymbol{\tau}(u) - \boldsymbol{\rho}(u))| \\ \le &\, 6\varepsilon_{|u|} + 4\eta. \end{aligned}$$

We apply Lemma 5.A.10 on the vectors

$$\frac{\boldsymbol{\tau}(u)_s}{\sqrt{\lambda_s N}}, \frac{\boldsymbol{\rho}(u1)_s - \boldsymbol{\tau}(u)_s}{\sqrt{\lambda_s N}}, \ldots, \frac{\boldsymbol{\rho}(uk)_s - \boldsymbol{\tau}(u)_s}{\sqrt{\lambda_s N}}$$

with $q = \phi^s_{|u|} \ge D^{-2}$, $q' = \phi^s_{|u|+1} - \phi^s_{|u|}$, and $\varepsilon = 6\varepsilon_{|u|} + 4\eta$. This gives us $\boldsymbol{\tau}(u1)_s, \ldots, \boldsymbol{\tau}(uk)_s$ satisfying (5.105), such that

$$R_s(\boldsymbol{\tau}(ui) - \boldsymbol{\tau}(u), \boldsymbol{\tau}(ui) - \boldsymbol{\tau}(u)) = \phi^s_{|u|+1} - \phi^s_{|u|}$$

and the vectors $\boldsymbol{\tau}(u)_s$, $\boldsymbol{\tau}(u1)_s - \boldsymbol{\tau}(u)_s$, $\boldsymbol{\tau}(uk)_s - \boldsymbol{\tau}(u)_s$ are pairwise orthogonal. From this we can see that

$$\begin{aligned}
\vec{R}(\boldsymbol{\tau}(ui), \boldsymbol{\tau}(u)) &= \vec{\phi}_{|u|}, \\
\vec{R}(\boldsymbol{\tau}(ui), \boldsymbol{\tau}(ui)) &= \vec{\phi}_{|u|+1}, \\
\vec{R}(\boldsymbol{\tau}(ui), \boldsymbol{\tau}(uj)) &= \vec{\phi}_{|u|} \quad \text{if } i \neq j.
\end{aligned}$$

Thus the $\underline{\boldsymbol{\tau}}$ constructed this way is an element of $\mathcal{Q}_{\mathrm{loc}+}(0)$. Finally, let $\underline{\boldsymbol{\sigma}}' = (\boldsymbol{\tau}(u))_{u\in\mathbb{L}}$, so $\underline{\boldsymbol{\sigma}}' \in \mathcal{Q}_{\mathrm{loc}}(0)$. Equation (5.105) implies that for all $u \in \mathbb{L}$,

$$\frac{1}{\sqrt{N}} \|\boldsymbol{\sigma}'(u) - \boldsymbol{\sigma}(u)\|_2 \le \varepsilon_D.$$

By Proposition 5.1.23, with probability $1 - e^{-\Omega(N)}$ we have $H_N^{(u)} \in K_N$ for all $u \in \mathbb{L}$. On this event,

$$\left| \frac{1}{N}\mathcal{H}_N(\underline{\boldsymbol{\sigma}}') - \frac{1}{N}\mathcal{H}_N(\underline{\boldsymbol{\sigma}}) \right| \le C\varepsilon_D$$

for some $C > 0$, and so

$$\frac{1}{N} \sup_{\underline{\boldsymbol{\sigma}}' \in \mathcal{Q}_{\mathrm{loc}}^{k,D,\underline{\vec{\phi}}'}(\eta)} \mathcal{H}_N(\underline{\boldsymbol{\sigma}}') + C\varepsilon_D \ge \frac{1}{N} \sup_{\underline{\boldsymbol{\sigma}} \in \mathcal{Q}_{\mathrm{loc}}^{k,D,\underline{\vec{\phi}}}(0)} \mathcal{H}_N(\underline{\boldsymbol{\sigma}}).$$

By Lemma 5.2.11, both sides of this inequality are subgaussian with fluctuations $O(N^{-1/2})$, so the contribution from the complement of this event is $o_N(1)$, and

$$\limsup_{N\to\infty} \frac{1}{N}\mathbb{E} \sup_{\underline{\boldsymbol{\sigma}}' \in \mathcal{Q}_{\mathrm{loc}}^{k,D,\underline{\vec{\phi}}'}(0)} \mathcal{H}_N(\underline{\boldsymbol{\sigma}}') + C\varepsilon_D \ge \limsup_{N\to\infty} \frac{1}{N}\mathbb{E} \sup_{\underline{\boldsymbol{\sigma}} \in \mathcal{Q}_{\mathrm{loc}}^{k,D,\underline{\vec{\phi}}}(\eta)} \mathcal{H}_N(\underline{\boldsymbol{\sigma}}).$$

Taking $\eta \to 0$ (which forces $\varepsilon_D \to 0$) followed by $D, k \to \infty$ implies $\mathsf{BOGP}_{\mathrm{loc}} \le \mathsf{BOGP}^-_{\mathrm{loc},0}$, as desired. $\square$

5.B Ground state energy of multi-species spherical SK with external field

We adopt the notations of Lemma 5.3.4. In this section, we will prove this lemma by showing that

$$\limsup_{N\to\infty} \mathbb{E}\mathrm{GS}_N(W, \vec{v}, k) \le \sum_{s\in\mathscr{S}} \lambda_s \sqrt{v_s^2 + 2\sum_{s'\in\mathscr{S}} \lambda_{s'} w_{s,s'}^2} \le \liminf_{N\to\infty} \mathbb{E}\mathrm{GS}_N(W, \vec{v}, k).$$

5.B.1 Upper bound for $\vec{v} = \vec{0}$, $k = 1$

The following (exact) upper bound for the case $\vec{v} = \vec{0}$, $k = 1$ follows from the results of [BBvH23]. We will prove Lemma 5.3.4 using only this result and elementary techniques.

Proposition 5.B.1. *For W as in Lemma 5.3.4,*

$$\limsup_{N\to\infty} \mathbb{E}\mathrm{GS}_N(W, \vec{0}, 1) \le \sum_{s\in\mathscr{S}} \lambda_s \sqrt{2\sum_{s'\in\mathscr{S}} \lambda_{s'} w_{s,s'}^2}.$$

Proof. In this proof, abbreviate $\mathrm{GS}_N = \mathrm{GS}_N(W, \vec{v}, k)$. Let $\boldsymbol{G} \in \boldsymbol{R}^{N\times N}$ have i.i.d. standard Gaussian entries. Thus $G = \frac{1}{2}\left(\boldsymbol{G} + \boldsymbol{G}^\top\right)$ is symmetric with $\mathcal{N}(0,1)$ diagonal entries, $\mathcal{N}(0,1/2)$ off-diagonal entries, and independent entries on and above the diagonal. Define $M \in \mathbb{R}^{N\times N}$ by $M_{i,j} = N^{-1/2} w_{s(i),s(j)} G_{i,j}$. It is clear by homogeneity that

$$\mathrm{GS}_N = \frac{1}{N} \max_{\boldsymbol{\sigma}\in\mathcal{B}_N} \boldsymbol{\sigma}^\top M \boldsymbol{\sigma}.$$

Let $\vec{C} \in \mathbb{R}_{>0}^{\mathscr{S}}$ be a vector of constants we will set later. We consider the rescaled matrix $\widetilde{M} = \sqrt{\vec{C}}^{\otimes 2} \diamond M$. This can be generated by $\widetilde{M} = \widehat{M} + \overline{M}$, where $\widehat{M}$ is a random symmetric matrix with independent entries on and above the diagonal

$$\widehat{M}_{i,j} \sim \mathcal{N}\left(0, \frac{C_{s(i)}C_{s(j)}w_{s(i),s(j)}^2}{2N}\right)$$

and $\overline{M}$ is a random diagonal matrix with independent entries

$$\overline{M}_{i,i} \sim \mathcal{N}\left(0, \frac{C_{s(i)}^2 w_{s(i),s(i)}^2}{2N}\right).$$

Clearly $\mathbb{E}\|\overline{M}\|_{\mathsf{op}} = O(\sqrt{N^{-1}\log N})$. [BBvH23, Theorem 1.2] states that

$$\mathbb{E}\|\widehat{M}\|_{\mathsf{op}} \le \|X_{\text{free}}\|_{\mathsf{op}} + O\left(v^{1/2}\sigma^{1/2}(\log N)^{3/4}\right),$$

where $\|X_{\text{free}}\|_{\mathsf{op}}, \sigma, v$ are defined as follows. We have

$$\sigma = \sqrt{\mathbb{E}\|\widehat{M}^2\|_{\mathsf{op}}} = O(1), \qquad v = \sqrt{\|\mathsf{Cov}(\widehat{M})\|_{\mathsf{op}}},$$

where $\mathsf{Cov}(\widehat{M}) \in \mathbb{R}^{N^2 \times N^2}$ is the covariance matrix of the entries of $\widehat{M}$ and has operator norm $O(1/N)$. It follows that the error term $v^{1/2}\sigma^{1/2}(\log N)^{3/4}$ contributes $o_N(1)$. Finally [BBvH23, Lemma 3.2] states that in our setting,

$$\|X_{\text{free}}\|_{\mathsf{op}} = 2 \sup_{\substack{a \in [0,1]^N \\ \sum_i a_i = 1}} \sum_{i \in [N]} \sqrt{a_i \sum_{i' \in [N]} \frac{C_{s(i)}C_{s(i')}w_{s(i),s(i')}^2 a_{i'}}{2N}}$$

It is not difficult to see by concavity of the square-root that, for $\lambda_{s,N} = |\mathcal{I}_s|/N$ (so $\lambda_{s,N} \to \lambda_s$) replacing all a_i such that $i \in \mathcal{I}_s$ with

$$A_s = \lambda_{s,N}^{-1} \sum_{i : s(i) = s} a_i$$

only improves the right-hand side. Substituting $B_s = C_s A_s$, we conclude that

$$\begin{aligned}
\|X_{\text{free}}\|_{\mathsf{op}} &= \sup_{\substack{\vec{A} \in \mathbb{R}_{\ge 0}^{\mathscr{S}} \\ \sum_s \lambda_{s,N} A_s = 1}} \sum_{s \in \mathscr{S}} \lambda_{s,N} \sqrt{2A_s \sum_{s' \in \mathscr{S}} \lambda_{s'} C_s C_{s'} w_{s,s'}^2 A_{s'}} \\
&= \sup_{\substack{\vec{B} \in \mathbb{R}_{\ge 0}^{\mathscr{S}} \\ \sum_s C_s^{-1} \lambda_{s,N} B_s = 1}} \sum_{s \in \mathscr{S}} \lambda_{s,N} \sqrt{2B_s \sum_{s' \in \mathscr{S}} \lambda_{s'} w_{s,s'}^2 B_{s'}}.
\end{aligned}$$

From the above discussion, $\|\widetilde{M}\|_{\mathsf{op}} \le \|X_{\text{free}}\|_{\mathsf{op}} + o_N(1)$. Moreover we observe that

$$\begin{aligned}
\mathrm{GS}_N &= \frac{1}{N} \max_{\|\boldsymbol{\sigma}_s\|_2^2 \le \lambda_s N} \boldsymbol{\sigma}^\top M \boldsymbol{\sigma} = \frac{1}{N} \max_{\|\boldsymbol{\sigma}_s\|_2^2 \le C_s^{-1}\lambda_s N} \boldsymbol{\sigma}^\top \widetilde{M} \boldsymbol{\sigma} \\
&\le \frac{1}{N} \max_{\|\boldsymbol{\sigma}\|_2^2 \le \sum_{s \in \mathscr{S}} C_s^{-1}\lambda_s N} \boldsymbol{\sigma}^\top \widetilde{M} \boldsymbol{\sigma} = \left(\sum_{s \in \mathscr{S}} C_s^{-1}\lambda_s\right) \|\widetilde{M}\|_{\mathsf{op}}.
\end{aligned}$$

Combining and using homogeneity, we find

$$
\begin{aligned}
\mathbb{E}\mathrm{GS}_N &\le \left(\sum_{s\in\mathscr{S}} C_s^{-1}\lambda_s\right)\mathbb{E}\|\widetilde{M}\|_{\mathsf{op}} \\
&= \left(\sum_{s\in\mathscr{S}} C_s^{-1}\lambda_s\right) \sup_{\substack{\vec{B}\in\mathbb{R}_{\ge 0}^{\mathscr{S}} \\ \sum_s C_s^{-1}\lambda_{s,N}B_s=1}} \sum_{s\in\mathscr{S}} \lambda_{s,N}\sqrt{2B_s\sum_{s'\in\mathscr{S}}\lambda_{s',N}w_{s,s'}^2 B_{s'}} + o_N(1) \\
&= \frac{\sum_{s\in\mathscr{S}} C_s^{-1}\lambda_s}{\sum_{s\in\mathscr{S}} C_s^{-1}\lambda_{s,N}} \cdot \sup_{\substack{\vec{D}\in\mathbb{R}_{\ge 0}^{\mathscr{S}} \\ \sum_s C_s^{-1}\lambda_{s,N}D_s=\sum_{s\in\mathscr{S}} C_s^{-1}\lambda_{s,N}}} \sum_{s\in\mathscr{S}} \lambda_{s,N}\sqrt{2D_s\sum_{s'\in\mathscr{S}}\lambda_{s',N}w_{s,s'}^2 D_{s'}} + o_N(1). \qquad (5.106)
\end{aligned}
$$

If the supremum in (5.106) is attained at $\vec{D}=\vec{1}$, then (because $\lambda_{s,N}\to\lambda_s$) we get the desired bound

$$
\mathbb{E}\mathrm{GS}_N \le \sum_{s\in\mathscr{S}}\lambda_s\sqrt{2\sum_{s'\in\mathscr{S}}\lambda_{s'}w_{s,s'}^2} + o_N(1).
$$

Crucially, we observe that the expression

$$
F(\vec{D}) = \sum_{s\in\mathscr{S}} \lambda_{s,N}\sqrt{2D_s\sum_{s'\in\mathscr{S}}\lambda_{s',N}w_{s,s'}^2 D_{s'}} \qquad (5.107)
$$

is concave in $\vec{D}$. Therefore if $\vec{D}=\vec{1}$ is a critical point of F within the set satisfying $\sum_s C_s^{-1}\lambda_{s,N}D_s = \sum_{s\in\mathscr{S}} C_s^{-1}\lambda_{s,N}$, then it also attains the supremum in (5.106). For the choice $C_s=\frac{\lambda_{s,N}}{\partial_{D_s}F}$, $\vec{D}=\vec{1}$ is a critical point of F. This concludes the proof. □

5.B.2 General upper bound

In this subsection, we will prove the following upper bound for the case $k=1$.

Proposition 5.B.2. *For $W,\vec{v}$ as in Lemma 5.3.4,*

$$
\limsup_{N\to\infty}\mathbb{E}\mathrm{GS}_N(W,\vec{v},1) \le \sum_{s\in\mathscr{S}}\lambda_s\sqrt{v_s^2+2\sum_{s'\in\mathscr{S}}\lambda_{s'}w_{s,s'}^2}.
$$

By slight abuse of notation, let $H_N=H_{N,1}^1$ and $\mathrm{GS}_N(W,\vec{v})=\mathrm{GS}_N(W,\vec{v},1)$. Recall that

$$
H_N(\boldsymbol{\sigma}) = \langle \vec{v}\diamond\boldsymbol{g},\boldsymbol{\sigma}\rangle + \widetilde{H}_N(\boldsymbol{\sigma}), \qquad \widetilde{H}_N(\boldsymbol{\sigma}) = \frac{1}{\sqrt{N}}\langle W\diamond\boldsymbol{G},\boldsymbol{\sigma}^{\otimes 2}\rangle
$$

where $\boldsymbol{g}\in\mathbb{R}^N$, $\boldsymbol{G}\in\mathbb{R}^{N\times N}$ have i.i.d. standard Gaussian entries. Define

$$
A(W,\vec{v}) = \limsup_{N\to\infty}\mathbb{E}\mathrm{GS}_N(W,\vec{v}).
$$

We first establish some basic properties of this limit.

Lemma 5.B.3. *A satisfies the following properties.*

(a) For any $c>0$, $A(cW,c\vec{v})=cA(W,\vec{v})$.

(b) $A(0,\vec{v})=\sum_{s\in\mathscr{S}}\lambda_s v_s$.

(c) $A(W,\vec{0})\le\sum_{s\in\mathscr{S}}\lambda_s\sqrt{2\sum_{s'\in\mathscr{S}}\lambda_{s'}w_{s,s'}^2}$.

(d) $A(W,\vec{v})\le A(W,\vec{0})+A(0,\vec{v})$.

Proof. Part (a) is obvious. Part (b) follows from

$$\mathbb{E}\mathrm{GS}_N(0,\vec{v}) = \frac{1}{N}\mathbb{E}\max_{\boldsymbol{\sigma}\in\mathcal{S}_N}\langle \vec{v}\diamond \boldsymbol{g},\boldsymbol{\sigma}\rangle = \frac{1}{N}\sum_{s\in\mathscr{S}}\sqrt{\lambda_s N}v_s\mathbb{E}\|\boldsymbol{g}_s\|_2 = \sum_{s\in\mathscr{S}}\lambda_s v_s + o_N(1).$$

Part (c) follows from Proposition 5.B.1. Part (d) follows from

$$\begin{aligned}
\mathrm{GS}_N(W,\vec{v}) &= \frac{1}{N}\max_{\boldsymbol{\sigma}\in\mathcal{S}_N}\left(\langle \vec{v}\diamond \boldsymbol{g},\boldsymbol{\sigma}\rangle + \frac{1}{\sqrt{N}}\langle W\diamond \boldsymbol{G},\boldsymbol{\sigma}^{\otimes 2}\rangle\right) \\
&\geq \frac{1}{N}\max_{\boldsymbol{\sigma}\in\mathcal{S}_N}\langle \vec{v}\diamond \boldsymbol{g},\boldsymbol{\sigma}\rangle + \frac{1}{N}\max_{\boldsymbol{\sigma}\in\mathcal{S}_N}\frac{1}{\sqrt{N}}\langle W\diamond \boldsymbol{G},\boldsymbol{\sigma}^{\otimes 2}\rangle \\
&= \mathrm{GS}_N(W,\vec{0}) + \mathrm{GS}_N(0,\vec{v}). \qquad (5.108)
\end{aligned}$$

□

Next we show some a priori regularity conditions on A.

Proposition 5.B.4. *Let*

$$C(W,\vec{v}) = 4\left(\sum_{s\in\mathscr{S}}\lambda_s v_s^2 + \sum_{s,s'\in\mathscr{S}}\lambda_s\lambda_{s'}w_{s,s'}^2\right).$$

Then, for sufficiently large N and all $t>0$,

$$\mathbb{P}\left[|\mathrm{GS}_N(W,\vec{v}) - \mathbb{E}\mathrm{GS}_N(W,\vec{v})| > t\right] \leq 2\exp\left(-\frac{Nt^2}{C(W,\vec{v})}\right).$$

Proof. Let $C = C(W,\vec{v})$. For any $\boldsymbol{\sigma}\in\mathcal{S}_N$,

$$\begin{aligned}
\mathbb{E}H_N(\boldsymbol{\sigma})^2 &= \|\vec{v}\diamond\boldsymbol{\sigma}\|_2^2 + \frac{1}{N}\left\|W\diamond\boldsymbol{\sigma}^{\otimes 2}\right\|_F^2 \\
&= N\left(\sum_{s\in\mathscr{S}}\lambda_{s,N}v_s^2 + \sum_{s,s'\in\mathscr{S}}\lambda_{s,N}\lambda_{s',N}w_{s,s'}^2\right) \leq \frac{CN}{2}.
\end{aligned}$$

for large enough N. By the Borell-TIS inequality, $\max_{\boldsymbol{\sigma}\in\mathcal{S}_N} H_N(\boldsymbol{\sigma})$ is $CN/2$-subgaussian, so $\mathrm{GS}_N(W,\vec{v})$ is $C/2N$-subgaussian, which implies the result. □

For $\vec{a} = (a_s)_{s\in\mathscr{S}'}\in[0,1]^{\mathscr{S}}$, define $W(W,\vec{v},\vec{a}) = (w'_{s,s'})_{s,s'\in\mathscr{S}}$ and $\vec{v}(W,\vec{v},\vec{a}) = (v'_s)_{s\in\mathscr{S}}$ where

$$w'_{s,s'} = \sqrt{(1-a_s)(1-a_{s'})}w_{s,s'}, \qquad v'_s = \sqrt{2(1-a_s)\left(\sum_{s'\in\mathscr{S}}\lambda_{s'}a_{s'}w_{s,s'}^2\right)}.$$

We will prove Proposition 5.B.2 using the following recursive upper bound in A.

Lemma 5.B.5. *For $W,\vec{v}$ as in Lemma 5.3.4,*

$$A(W,\vec{v}) \leq \max_{\vec{a}\in[0,1]^{\mathscr{S}}}\sum_{s\in\mathscr{S}}\lambda_s v_s\sqrt{a_s} + A\left(W(W,\vec{v},\vec{a}),\vec{v}(W,\vec{v},\vec{a})\right). \qquad (5.109)$$

Proof. Define $\widehat{\boldsymbol{g}}\in\mathcal{S}_N$ by $\widehat{\boldsymbol{g}}_s = \frac{\sqrt{\lambda_s N}\boldsymbol{g}_s}{\|\boldsymbol{g}_s\|_2}$ for each $s\in\mathscr{S}$. For $\vec{a}\in[0,1]^{\mathscr{S}}$, define

$$\mathrm{GS}_N(W,\vec{v};\vec{a}) = \frac{1}{N}\max_{\boldsymbol{\sigma}\in\mathcal{R}_N(\vec{a})}H_N(\boldsymbol{\sigma}), \qquad \mathcal{R}_N(\vec{a}) = \left\{\boldsymbol{\sigma}\in\mathcal{S}_N : R(\boldsymbol{\sigma},\widehat{\boldsymbol{g}}) = \sqrt{\vec{a}}\right\}.$$

For a non-random $\vec{a}$ and any $\boldsymbol{\sigma}\in\mathcal{R}_N(\vec{a})$,

$$\langle \vec{v}\diamond\boldsymbol{g},\boldsymbol{\sigma}\rangle = N\sum_{s\in\mathscr{S}}\lambda_s v_s\sqrt{a_s}\frac{\|\boldsymbol{g}_s\|_2}{\sqrt{\lambda_s N}}.$$

For $\boldsymbol{\sigma} \in \mathcal{R}_N(\vec{a})$, we may write $\boldsymbol{\sigma} = \sqrt{\vec{a}} \diamond \widehat{\boldsymbol{g}} + \sqrt{\vec{1} - \vec{a}} \diamond \boldsymbol{\rho}$ for $\boldsymbol{\rho} \in \mathcal{R}_N(\vec{0})$. Define the Gaussian process $\widehat{H}_N^{\vec{a}}(\boldsymbol{\rho}) = \widetilde{H}_N\big(\sqrt{\vec{a}} \diamond \widehat{\boldsymbol{g}} + \sqrt{\vec{1} - \vec{a}} \diamond \boldsymbol{\rho}\big)$, which is supported on $\mathcal{R}_N(\vec{0})$. We next calculate the covariance of this process. Recall that the covariance of $\widetilde{H}_N$ is

$$\mathbb{E}\widetilde{H}_N(\boldsymbol{\sigma})\widetilde{H}_N(\boldsymbol{\sigma'}) = N\xi(R(\boldsymbol{\sigma}, \boldsymbol{\sigma'})), \qquad \xi(\vec{x}) = \left\langle W \odot W, (\vec{\lambda} \odot \vec{x})^{\otimes 2} \right\rangle.$$

Because $\boldsymbol{g}, \boldsymbol{G}$ are independent, the covariance of $\widehat{H}_N^{\vec{a}}$ is

$$\mathbb{E}\widehat{H}_N^{\vec{a}}(\boldsymbol{\rho})\widehat{H}_N^{\vec{a}}(\boldsymbol{\rho'}) = N\xi_{\vec{a}}(R(\boldsymbol{\rho}, \boldsymbol{\rho'})), \tag{5.110}$$

where, for $W' = W(W, \vec{v}, \vec{a})$ and $\vec{v}' = \vec{v}(W, \vec{v}, \vec{a})$,

$$\begin{aligned}\xi_{\vec{a}}(\vec{x}) &= \xi\left(\vec{a} + (\vec{1} - \vec{a}) \odot \vec{x}\right) = \left\langle W \odot W, (\vec{\lambda} \odot \vec{a} + \vec{\lambda} \odot (1 - \vec{a}) \odot \vec{x})^{\otimes 2} \right\rangle \\ &= \left\langle W' \odot W', (\vec{\lambda} \odot \vec{x})^{\otimes 2} \right\rangle + \left\langle \vec{v}' \odot \vec{v}', \vec{\lambda} \odot \vec{x} \right\rangle + \left\langle W \odot W, (\vec{\lambda} \odot \vec{a})^{\otimes 2} \right\rangle. \end{aligned} \tag{5.111}$$

We may construct a Gaussian process $\overline{H}_N^{\vec{a}}$ (conditional on $\boldsymbol{g}$) on $\mathcal{S}_N$ with covariance (5.110) whose restriction to $\mathcal{R}_N(\vec{0})$ agrees with $\widehat{H}_N^{\vec{a}}$. Thus

$$\begin{aligned}\mathrm{GS}_N(W, \vec{v}; \vec{a}) &= \sum_{s \in \mathscr{S}} \lambda_s v_s \sqrt{a_s} \frac{\|\boldsymbol{g}_s\|_2}{\sqrt{\lambda_s N}} + \frac{1}{N} \max_{\boldsymbol{\rho} \in \mathcal{R}_N(\vec{0})} \widehat{H}_N^{\vec{a}}(\boldsymbol{\rho}) \\ &\le \sum_{s \in \mathscr{S}} \lambda_s v_s \sqrt{a_s} \frac{\|\boldsymbol{g}_s\|_2}{\sqrt{\lambda_s N}} + \frac{1}{N} \max_{\boldsymbol{\rho} \in \mathcal{S}_N} \overline{H}_N^{\vec{a}}(\boldsymbol{\rho}).\end{aligned}$$

Moreover,

$$\frac{1}{N} \max_{\boldsymbol{\rho} \in \mathcal{S}_N} \overline{H}_N^{\vec{a}}(\boldsymbol{\rho}) =_d \mathrm{GS}(W(W, \vec{v}, \vec{a}), \vec{v}(W, \vec{v}, \vec{a})) + \frac{1}{\sqrt{N}} \left\langle W \odot W, (\vec{\lambda} \odot \vec{a})^{\otimes 2} \right\rangle^{1/2} Z$$

for an independent $Z \sim \mathcal{N}(0, 1)$. Let $\mathcal{D} = \{0, \frac{1}{N}, \dots, \frac{N-1}{N}, 1\}^{\mathscr{S}}$. Let $\mathcal{E}$ be the event that

(a) For a constant L, $H_N(\boldsymbol{\sigma})$ is $L\sqrt{N}$-Lipschitz on $\boldsymbol{\sigma} \in \mathcal{S}_N$. By Proposition 5.1.23, this occurs with probability $1 - \exp(-CN)$.

(b) For all $s \in \mathscr{S}$, $|\|\boldsymbol{g}_s\|_2 - \sqrt{\lambda_{s,N} N}| \le N^{1/4}$; by standard concentration inequalities this holds with probability $1 - r\exp(-CN^{1/2})$.

(c) For all $\vec{a} \in \mathcal{D}$, $|\frac{1}{N} \max_{\boldsymbol{\rho} \in \mathcal{S}_N} \overline{H}_N^{\vec{a}}(\boldsymbol{\rho}) - \mathbb{E}\mathrm{GS}_N(W(W, \vec{v}, \vec{a}), \vec{v}(W, \vec{v}, \vec{a}))| \le N^{-1/4}$; by Proposition 5.B.4 and standard tail bounds on Z this holds with probability $1 - 2(N+1)^r \exp(-CN^{1/2})$. Here we use that for $\vec{a} \in \mathcal{D}$, the constants $C(W(W, \vec{v}, \vec{a}), \vec{v}(W, \vec{v}, \vec{a}))$ in Proposition 5.B.4 are uniformly upper bounded.

By adjusting C, $\mathbb{P}(\mathcal{E}) \ge 1 - \exp(-CN^{1/2})$. On $\mathcal{E}$, if $\boldsymbol{\sigma} \in \mathcal{S}_N$ maximizes H_N, we can find $\boldsymbol{\sigma'} \in \bigcup_{\vec{a} \in \mathcal{D}} \mathcal{R}_N(\vec{a})$ with $\|\boldsymbol{\sigma'} - \boldsymbol{\sigma}\|_2 \le O(1/\sqrt{N})$. By the Lipschitz condition (a), $|H(\boldsymbol{\sigma}) - H(\boldsymbol{\sigma'})| \le O(1)$. So,

$$\begin{aligned}\mathrm{GS}_N(W, \vec{v}) &= \frac{1}{N} H_N(\boldsymbol{\sigma}) \le \frac{1}{N} H_N(\boldsymbol{\sigma'}) + O(1/N) \\ &\le \max_{\vec{a} \in \mathcal{D}} \mathrm{GS}_N(W, \vec{v}; \vec{a}) + O(1/N) \\ &\le \max_{\vec{a} \in \mathcal{D}} \left(\sum_{s \in \mathscr{S}} \lambda_s v_s \sqrt{a_s} + \mathbb{E}\mathrm{GS}_N(W(W, \vec{v}, \vec{a}), \vec{v}(W, \vec{v}, \vec{a})) \right) + o_N(1).\end{aligned}$$

The subgaussianity from Proposition 5.B.4 implies that the contribtion to $\mathbb{E}\mathrm{GS}_N(W, \vec{v})$ from $\mathcal{E}^c$ is $o_N(1)$, so

$$\begin{aligned}\mathbb{E}\mathrm{GS}_N(W, \vec{v}) &\le \max_{\vec{a} \in \mathcal{D}} \left(\sum_{s \in \mathscr{S}} \lambda_s v_s \sqrt{a_s} + \mathbb{E}\mathrm{GS}_N(W(W, \vec{v}, \vec{a}), \vec{v}(W, \vec{v}, \vec{a})) \right) + o_N(1) \\ &\le \max_{\vec{a} \in [0,1]^{\mathscr{S}}} \left(\sum_{s \in \mathscr{S}} \lambda_s v_s \sqrt{a_s} + \mathbb{E}\mathrm{GS}_N(W(W, \vec{v}, \vec{a}), \vec{v}(W, \vec{v}, \vec{a})) \right) + o_N(1).\end{aligned}$$

Taking $\limsup_{N \to \infty}$ on both sides yields the result. $\square$

Proof of Proposition 5.B.2. We will show that any A satisfying the properties in Lemma 5.B.3 and the bound (5.109) must satisfy

$$A(W,\vec{v}) \le A_*(W,\vec{v}) \equiv \sum_{s\in\mathscr{S}} \lambda_s \sqrt{v_s^2 + 2\sum_{s'\in\mathscr{S}} \lambda_{s'} w_{s,s'}^2}.$$

Clearly A_* satisfies the conclusions of Lemma 5.B.3, with equality in assertion (c). For any $\vec{a}\in[0,1]^{\mathscr{S}}$,

$$\begin{aligned}
A_*\left(W(W,\vec{v},\vec{a}),\vec{v}(W,\vec{v},\vec{a})\right) &= \sum_{s\in\mathscr{S}} \lambda_s \sqrt{2(1-a_s)\sum_{s'\in\mathscr{S}} a_{s'}\lambda_{s'} w_{s,s'}^2 + 2\sum_{s'\in\mathscr{S}} \lambda_{s'}(1-a_s)(1-a_{s'}) w_{s,s'}^2} \\
&= \sum_{s\in\mathscr{S}} \lambda_s \sqrt{2(1-a_s)\sum_{s'\in\mathscr{S}} \lambda_{s'} w_{s,s'}^2}, \\
\implies \sum_{s\in\mathscr{S}} \lambda_s v_s\sqrt{a_s} + A_*\left(W(W,\vec{v},\vec{a}),\vec{v}(W,\vec{v},\vec{a})\right) &= \sum_{s\in\mathscr{S}} \lambda_s \left(\sqrt{a_s} v_s + \sqrt{1-a_s}\sqrt{2\sum_{s'\in\mathscr{S}} \lambda_{s'} w_{s,s'}^2}\right) \\
&\le \sum_{s\in\mathscr{S}} \lambda_s \sqrt{v_s^2 + 2\sum_{s'\in\mathscr{S}} \lambda_{s'} w_{s,s'}^2} = A_*(W,\vec{v})
\end{aligned}$$

by Cauchy-Schwarz. Equality holds when

$$a_s = \frac{v_s^2}{v_s^2 + 2\sum_{s'\in\mathscr{S}} \lambda_{s'} w_{s,s'}^2} \tag{5.112}$$

for all $s\in\mathscr{S}$, and so A_* satisfies (5.109) with equality.

Suppose A satisfies the conclusions of Lemma 5.B.3 and the inequality (5.109), and there exists $(W,\vec{v})$ with $A(W,\vec{v}) > A_*(W,\vec{v})$. By homogeneity (Lemma 5.B.3(a)), we can assume $1 = \|W\|_1 \equiv \sum_{s,s'\in\mathscr{S}} w_{s,s'}$. For any small $\delta > 0$, we may choose $(W^*,\vec{v}^*)$ such that $\|W^*\|_1 = 1$ and

$$A(W^*,\vec{v}^*) - A_*(W^*,\vec{v}^*) \ge (1-\delta) \sup_{(W,\vec{v}):\|W\|_1=1} \left(A(W,\vec{v}) - A_*(W,\vec{v})\right) > 0.$$

Set

$$\vec{a}^* = \underset{\vec{a}\in[0,1]^{\mathscr{S}}}{\arg\max} \sum_{s\in\mathscr{S}} \lambda_s v_s^* \sqrt{a_s} + A\left(W(W^*,\vec{v}^*,\vec{a}),\vec{v}(W^*,\vec{v}^*,\vec{a})\right),$$

and $W' = W(W^*,\vec{v}^*,\vec{a}^*), \vec{v}' = \vec{v}(W^*,\vec{v}^*,\vec{a}^*)$, so

$$\begin{aligned}
A(W^*,\vec{v}^*) &\le \sum_{s\in\mathscr{S}} \lambda_s v_s^* \sqrt{a_s^*} + A(W',\vec{v}'), \\
A_*(W^*,\vec{v}^*) &\ge \sum_{s\in\mathscr{S}} \lambda_s v_s^* \sqrt{a_s^*} + A_*(W',\vec{v}').
\end{aligned}$$

Here, the second inequality uses that A_* satisfies (5.109) with equality. Therefore

$$A(W',\vec{v}') - A_*(W',\vec{v}') \ge A(W^*,\vec{v}^*) - A_*(W^*,\vec{v}^*) \ge (1-\delta) \sup_{(W,\vec{v}):\|W\|_1=1} \left(A(W,\vec{v}) - A_*(W,\vec{v})\right). \tag{5.113}$$

By homogeneity, this implies $\|W'\|_1 \ge 1-\delta$. Let $\mathscr{S}_0 \subseteq \mathscr{S}$ be the set of s for which there exists s' with $w_{s,s'} \ge \delta_1 \equiv \sqrt{2\delta}$. For such s, s',

$$\delta \ge \|W^*\|_1 - \|W'\|_1 \ge w_{s,s'}^* - w_{s,s'}' \ge \left(1-\sqrt{1-a_s}\right) w_{s,s'} \ge \frac{1}{2} a_s w_{s,s'} \ge \frac{1}{2} a_s \delta_1.$$

Thus, for $s\in\mathscr{S}_0$, $a_s \le \delta_1$. Of course, for $s\in\mathscr{S}\setminus\mathscr{S}_0$, $w_{s,s'} \le \delta_1$ for all $s'\in\mathscr{S}$. Thus for all $s\in\mathscr{S}$,

$$v_s' \le \sqrt{2\sum_{s'\in\mathscr{S}} \lambda_{s'}\delta_1} = \sqrt{2\delta_1} \equiv \delta_2.$$

By parts (d), (b), and (c) of Lemma 5.B.3,

$$A(W', \vec{v}') \le A(W', \vec{0}) + A(0, \vec{v}') \le A(W', \vec{0}) + \delta_2 \le A_*(W', \vec{0}) + \delta_2.$$

By inspection, $A_*(W', \vec{v}') \ge A_*(W', \vec{0})$. Thus

$$A(W', \vec{v}') - A_*(W', \vec{v}') \le \delta_2.$$

For small enough $\delta > 0$, this contradicts (5.113). $\square$

Finally, the upper bound for $k = 1$ directly implies the upper bound for general k.

Corollary 5.B.6. *For $W, \vec{v}$ as in Lemma 5.3.4,*

$$\limsup_{N\to\infty} \mathbb{E}\mathrm{GS}_N(W, \vec{v}, k) \le \sum_{s\in\mathscr{S}} \lambda_s \sqrt{v_s^2 + 2\sum_{s'\in\mathscr{S}} \lambda_{s'} w_{s,s'}^2}.$$

Proof. Note that (recall (5.49))

$$\mathrm{GS}_N(W, \vec{v}, k) = \frac{1}{kN} \max_{\vec{\boldsymbol{\sigma}}\in\mathcal{S}_N^{k,\perp}} H_{N,k} \le \frac{1}{k}\sum_{i=1}^{k} \frac{1}{N} \max_{\boldsymbol{\sigma}^i\in\mathcal{S}_N} H_{N,k}^i(\boldsymbol{\sigma}^i).$$

Taking expectations yields $\mathbb{E}\mathrm{GS}_N(W, \vec{v}, k) \le \mathbb{E}\mathrm{GS}_N(W, \vec{v}, 1)$. This and Proposition 5.B.2 imply the result. $\square$

Remark 5.B.7. The proof of Proposition 5.B.2 via the recursive inequality (5.109) extends to the ground state energies in multi-species spherical spin glasses with general (non-quadratic) interactions. It thus gives an elementary way to upper bound the ground state energy for spin glasses with external field given the ground state energy of spin glasses without external field, when the latter is known. As we will see in the next subsection, it is possible to construct points where this recursive inequality holds with (approximate) equality, so the upper bound is sharp.

5.B.3 Lower bound

In this subsection, we will constructively prove the matching lower bound to Corollary 5.B.6.

Proposition 5.B.8. *For $W, \vec{v}$ as in Lemma 5.3.4,*

$$\liminf_{N\to\infty} \mathbb{E}\mathrm{GS}_N(W, \vec{v}, k) \ge \sum_{s\in\mathscr{S}} \lambda_s \sqrt{v_s^2 + 2\sum_{s'\in\mathscr{S}} \lambda_{s'} w_{s,s'}^2}.$$

Lemma 5.B.9. *Let $S_N = \{\boldsymbol{x} \in \mathbb{R}^N : \|\boldsymbol{x}\|_2 = \sqrt{N}\}$. Suppose $\boldsymbol{y}^1, \ldots, \boldsymbol{y}^k \in S_N$ satisfy $|\langle \boldsymbol{y}^i, \boldsymbol{y}^j\rangle| \le N^{2/3}$ for all $i \neq j$. Then there exist pairwise orthogonal $\boldsymbol{z}^1, \ldots, \boldsymbol{z}^k \in S_N$ such that $\mathsf{span}(\boldsymbol{z}^1, \ldots, \boldsymbol{z}^k) = \mathsf{span}(\boldsymbol{y}^1, \ldots, \boldsymbol{y}^k)$ and $\langle \boldsymbol{y}^i, \boldsymbol{z}^i\rangle \ge N - 4kN^{1/3}$.*

Proof. We define $\boldsymbol{z}^1, \ldots, \boldsymbol{z}^k$ by applying the Gram-Schmidt algorithm to $\boldsymbol{y}^1, \ldots, \boldsymbol{y}^k$: let $\boldsymbol{z}^1 = \boldsymbol{y}^1$, and for $2 \le i \le k$, let

$$\widetilde{\boldsymbol{y}}^i = \boldsymbol{y}^i - \sum_{j=1}^{i-1} \frac{\langle \boldsymbol{z}^j, \boldsymbol{y}^i\rangle}{N} \boldsymbol{z}^j, \qquad \boldsymbol{z}^i = \frac{\sqrt{N}}{\left\|\widetilde{\boldsymbol{y}}^i\right\|_2} \widetilde{\boldsymbol{y}}^i.$$

Clearly $\mathsf{span}(\boldsymbol{z}^1, \ldots, \boldsymbol{z}^k) = \mathsf{span}(\boldsymbol{y}^1, \ldots, \boldsymbol{y}^k)$. We will prove by induction on i that for all $\ell > i$, $|\langle \boldsymbol{z}^i, \boldsymbol{y}^\ell\rangle| \le 2N^{2/3}$. The base case $i = 1$ is true by hypothesis. For $i > 1$, we have

$$\left\|\widetilde{\boldsymbol{y}}^i\right\|_2^2 = N\left(1 - \sum_{j=1}^{i-1} \frac{\langle \boldsymbol{z}^j, \boldsymbol{y}^i\rangle^2}{N}\right) \in \left[N(1 - 4kN^{-2/3}), N\right],$$

using the inductive hypothesis. Moreover, for $\ell > i$,

$$|\langle \widetilde{\boldsymbol{y}}^i, \boldsymbol{y}^\ell \rangle| \le |\langle \boldsymbol{y}^i, \boldsymbol{y}^\ell \rangle| + \sum_{j=1}^{i-1} \frac{|\langle \boldsymbol{z}^j, \boldsymbol{y}^i \rangle||\langle \boldsymbol{z}^j, \boldsymbol{y}^\ell \rangle|}{N} \le N^{2/3} \left(1 + 4kN^{-1/3}\right).$$

Therefore

$$|\langle \boldsymbol{z}^i, \boldsymbol{y}^\ell \rangle| \le N^{2/3} \left(1 + 4kN^{-1/3}\right) \left(1 - kN^{-2/3}\right)^{-1/2} \le 2N^{2/3},$$

completing the induction. Now $\langle \widetilde{\boldsymbol{y}}^i, \boldsymbol{y}^i \rangle = \left\|\widetilde{\boldsymbol{y}}^i\right\|_2^2$, so

$$\langle \boldsymbol{z}^i, \boldsymbol{y}^i \rangle = \sqrt{N}\left\|\widetilde{\boldsymbol{y}}^i\right\|_2 \ge N \left(1 - 4kN^{-2/3}\right)^{1/2} \ge N - 4kN^{1/3}.$$ □

Recall that $\lambda_{s,N} = |\mathcal{I}_s|/N$. Let $\delta_N = \max_{s\in\mathscr{S}} |\frac{\lambda_{s,N}}{\lambda_s} - 1|$.

Lemma 5.B.10. *There exists an event $\mathcal{E} \in \sigma(\boldsymbol{g}^1, \dots, \boldsymbol{g}^k)$ with $\mathbb{P}(\mathcal{E}) \ge 1 - \exp(-CN^{1/3})$ such that on this event, there exists $\vec{\boldsymbol{x}} = (\boldsymbol{x}^1, \dots, \boldsymbol{x}^k) \in \mathcal{S}_N^{k,\perp}$ such that the following properties hold.*

(a) For all i, $\left\|R(\boldsymbol{g}^i, \boldsymbol{g}^i) - \vec{1}\right\|_\infty \le \delta_N + N^{-1/4}$.

(b) For all i, $\left\|R(\boldsymbol{g}^i, \boldsymbol{x}^i) - \vec{1}\right\|_\infty \le \delta_N + N^{-1/4}$.

(c) For all i, $R(\boldsymbol{x}^i, \boldsymbol{x}^i) = \vec{1}$.

(d) For all $s \in \mathscr{S}$, $\mathsf{span}(\boldsymbol{x}_s^1, \dots, \boldsymbol{x}_s^k) = \mathsf{span}(\boldsymbol{g}_s^1, \dots, \boldsymbol{g}_s^k)$.

Proof. By standard concentration inequalities, for each $i \in [k]$ and $s \in \mathscr{S}$, $|\langle \boldsymbol{g}_s^i, \boldsymbol{g}_s^i \rangle - \lambda_{s,N}N| \le N^{2/3}$ with probability $1 - \exp(-CN^{1/3})$, which implies

$$\left|\frac{\langle \boldsymbol{g}_s^i, \boldsymbol{g}_s^i \rangle}{\lambda_s N} - 1\right| \le \left|\frac{\lambda_{s,N}}{\lambda_s} - 1\right| + \frac{N^{2/3}}{\lambda_s N} \le \delta_N + N^{-1/4}.$$

When this holds for all $i \in [k]$, $s \in \mathscr{S}$, part (a) follows.

For each $i \in [k]$, define $\widehat{\boldsymbol{g}}^i \in \mathbb{R}^N$ by $\widehat{\boldsymbol{g}}_s^i = \frac{\sqrt{\lambda_{s,N}N}}{\|\boldsymbol{g}_s^i\|_2}\boldsymbol{g}_s^i$ for all $s \in \mathscr{S}$. Note that each $\widehat{\boldsymbol{g}}_s^i$ is a uniformly random point on the sphere of radius $\sqrt{\lambda_{s,N}N}$ supported on the coordinates $\mathcal{I}_s$.

Fix $s \in \mathscr{S}$. By standard concentration inequalities, for each pair of distinct $i, j \in [k]$, $|\langle \widehat{\boldsymbol{g}}_s^i, \widehat{\boldsymbol{g}}_s^j \rangle| \le (\lambda_{s,N}N)^{2/3}$ with probability $1 - \exp(-CN^{1/3})$. If this holds for all s, i, j, Lemma 5.B.9 implies the existence of orthogonal $\boldsymbol{z}_s^1, \dots, \boldsymbol{z}_s^k$ on the sphere of radius $\sqrt{\lambda_{s,N}N}$ supported on coordinates $\mathcal{I}_s$ with

$$\mathsf{span}(\boldsymbol{z}_s^1, \dots, \boldsymbol{z}_s^k) = \mathsf{span}(\widehat{\boldsymbol{g}}_s^1, \dots, \widehat{\boldsymbol{g}}_s^k) \tag{5.114}$$

and

$$\lambda_{s,N}N - 4k(\lambda_{s,N}N)^{1/3} \le \langle \boldsymbol{z}_s^k, \widehat{\boldsymbol{g}}_s^i \rangle \le \lambda_{s,N}N.$$

Let $\boldsymbol{x}_s^i = \boldsymbol{z}_s^i \cdot \sqrt{\lambda_s/\lambda_{s,N}}$, so

$$\frac{\langle \boldsymbol{x}_s^i, \boldsymbol{g}_s^i \rangle}{\lambda_s N} = \frac{\langle \boldsymbol{z}_s^i, \widehat{\boldsymbol{g}}_s^i \rangle}{\lambda_{s,N}N} \cdot \sqrt{\frac{\lambda_{s,N}}{\lambda_s}} \cdot \frac{\|\boldsymbol{g}_s^i\|}{\sqrt{\lambda_{s,N}N}} = (1 + O(N^{-1/3})\sqrt{\frac{\lambda_{s,N}}{\lambda_s}}.$$

Thus

$$\left|\frac{\langle \boldsymbol{x}_s^i, \boldsymbol{g}_s^i \rangle}{\lambda_s N} - 1\right| \le \left|\sqrt{\frac{\lambda_{s,N}}{\lambda_s}} - 1\right| + O(N^{-1/3}) \le \delta_N + N^{-1/4}.$$

If this holds for all s, part (b) follows. By a union bound, adjusting C as necessary, the above events simultaneously hold with probability $1 - \exp(-CN^{1/3})$. By construction, $R(\boldsymbol{x}^i, \boldsymbol{x}^i) = \vec{1}$ and $R(\boldsymbol{x}^i, \boldsymbol{x}^j) = \vec{0}$ for all $i \neq j$, which implies part (c) and $\vec{\boldsymbol{x}} \in \mathcal{S}_N^{k,\perp}$. The relation (5.114) implies part (d). □

The following recursive lower bound for $\mathbb{E}\mathrm{GS}_N(W,\vec{v},k)$ is a converse to Lemma 5.B.5 and is the main step in the proof of Proposition 5.B.8.

Lemma 5.B.11. *Let $W,\vec{v}$ be as in Lemma 5.3.4 and $\vec{a}\in[0,1]^{\mathscr{S}}$, and set $W'=W(W,\vec{v},\vec{a})$, $\vec{v}'=\vec{v}(W,\vec{v},\vec{a})$. Then,*

$$\mathbb{E}\mathrm{GS}_N(W,\vec{v},k)\geq\sum_{s\in\mathscr{S}}\lambda_s v_s\sqrt{a_s}+\mathbb{E}\mathrm{GS}_{N-kr}(W',\vec{v}',k)-o_N(1),$$

where GS_{N-kr} denotes the ground state energy (see (5.50)) of a dimension $N-kr$ multi-species quadratic spin glass with species sizes $\widetilde{\mathcal{I}}_s=\mathcal{I}_s-k$.

Proof. Suppose for now the event $\mathcal{E}$ in Lemma 5.B.10 holds and let $\vec{\boldsymbol{x}}=(\boldsymbol{x}^1,\ldots,\boldsymbol{x}^k)$ be as in this lemma. Let

$$\mathcal{S}_{N,\perp}\equiv\left\{\boldsymbol{\rho}\in\mathcal{S}_N:R(\boldsymbol{\rho},\boldsymbol{x}^i)=\vec{0}\ \forall i\in[k]\right\}=\left\{\boldsymbol{\rho}\in\mathcal{S}_N:R(\boldsymbol{\rho},\boldsymbol{g}^i)=\vec{0}\ \forall i\in[k]\right\}\tag{5.115}$$

where the second equality follows from Lemma 5.B.10(d) and

$$\mathcal{S}_{N,\perp}^{k,\perp}\equiv\left\{\vec{\boldsymbol{\rho}}=(\boldsymbol{\rho}^1,\ldots,\boldsymbol{\rho}^k)\in\mathcal{S}_{N,\perp}^k:R(\boldsymbol{\rho}^i,\boldsymbol{\rho}^j)=\vec{0}\ \forall i\neq j\right\}.\tag{5.116}$$

For each $i\in[k]$ let

$$\boldsymbol{\sigma}^i=\sqrt{\vec{a}}\diamond\boldsymbol{x}^i+\sqrt{\vec{1}-\vec{a}}\diamond\boldsymbol{\rho}^i\tag{5.117}$$

where $\vec{\boldsymbol{\rho}}=(\boldsymbol{\rho}^1,\ldots,\boldsymbol{\rho}^k)\in\mathcal{S}_{N,\perp}^{k,\perp}$. The orthogonality relations in (5.115) and (5.116) imply $\vec{\boldsymbol{\sigma}}=(\boldsymbol{\sigma}^1,\ldots,\boldsymbol{\sigma}^k)\in\mathcal{S}_N^{k,\perp}$. Then,

$$\frac{1}{N}H_{N,k}(\vec{\boldsymbol{\sigma}})=\frac{1}{kN}\sum_{i=1}^k\langle\vec{v}\diamond\boldsymbol{g}^i,\sqrt{\vec{a}}\diamond\boldsymbol{x}^i\rangle+\frac{1}{kN^{3/2}}\sum_{i=1}^k\left\langle W\diamond\boldsymbol{G},(\sqrt{\vec{a}}\diamond\boldsymbol{x}^i+\sqrt{\vec{1}-\vec{a}}\diamond\boldsymbol{\rho}^i)^{\otimes 2}\right\rangle.\tag{5.118}$$

By Lemma 5.B.10(a, b),

$$\frac{1}{N}\langle W\diamond\boldsymbol{G},\sqrt{\vec{a}}\diamond\boldsymbol{x}^i\rangle=\sum_{s\in\mathscr{S}}\lambda_s v_s\sqrt{a_s}+o_N(1).$$

Note that the state space $\mathcal{S}_{N,\perp}$ is $\mathcal{S}_N$ with k fewer dimensions in each species, and these dimensions (and $\vec{\boldsymbol{x}}$) are independent of $\boldsymbol{G}$. So, optimizing the second term of (5.118) over $\vec{\boldsymbol{\rho}}\in\mathcal{S}_{N,\perp}^{k,\perp}$ is equivalent to optimizing a dimension $N-kr$ multi-species quadratic spin glass. The same covariance calculation as (5.111) shows that

$$\sup_{\vec{\boldsymbol{\rho}}\in\mathcal{S}_{N,\perp}^{k,\perp}}\frac{1}{N^{3/2}}\left\langle W\diamond\boldsymbol{G},(\sqrt{\vec{a}}\diamond\boldsymbol{x}^i+\sqrt{\vec{1}-\vec{a}}\diamond\boldsymbol{\rho}^i)^{\otimes 2}\right\rangle=_d\sqrt{\frac{N-kr}{N}}\mathrm{GS}_{N-kr}(W',\vec{v}')+O(N^{-1/2})Z,$$

where $Z\sim\mathcal{N}(0,1)$ is independent of GS_{N-kr}. Thus

$$\mathbb{E}\mathrm{GS}_N(W,\vec{v},k)\geq\mathbb{E}\mathbf{1}\{\mathcal{E}\}\frac{1}{N}H_{N,k}(\vec{\boldsymbol{\sigma}})\geq\sum_{s\in\mathscr{S}}\lambda_s v_s\sqrt{a_s}+\mathbb{E}\mathrm{GS}_{N-kr}(W',\vec{v}',k)-o_N(1).\qquad\square$$

Lemma 5.B.11 suggests a natural way to construct an approximate ground state of $H_{N,k}$. First, use Gram-Schmidt orthogonalization to produce $\vec{\boldsymbol{x}}=(\boldsymbol{x}^1,\ldots,\boldsymbol{x}^k)$ from the external fields $\boldsymbol{g}^1,\ldots,\boldsymbol{g}^k$, as in Lemma 5.B.10. Choose $\vec{a}\in[0,1]^{\mathscr{S}}$ and set $\vec{\boldsymbol{\sigma}}$ as in (5.117), for $\vec{\boldsymbol{\rho}}\in\mathcal{S}_{N,\perp}^{k,\perp}$ to be determined. The correlations of the $\boldsymbol{\sigma}^i$ with the external fields $\boldsymbol{g}^i$ contribute energy $\sum_{s\in\mathscr{S}}\lambda_s v_s\sqrt{a_s}$, while the optimization over $\vec{\boldsymbol{\rho}}$ is equivalent to optimizing another quadratic multi-species spin glass, whose parameters depend on $\vec{a}$. Finally, recursively optimize $\vec{\boldsymbol{\rho}}$. The following proof demonstrates that when $\vec{v}>\vec{0}$, there exists a sequence of choices of $\vec{a}$ such that running this algorithm to a large constant recursion depth finds a near ground state $\vec{\boldsymbol{\sigma}}\in\mathcal{S}_N^{k,\perp}$ of $H_{N,k}$. (If some entries of $\vec{v}$ are zero, the algorithm succeeds after first introducing a small artificial external field.)

Proof of Proposition 5.B.8. Assume for now that $\vec{v} \succ \vec{0}$ where the inequality is strict in each coordinate. Define $W^{(0)} = W$, $\vec{v}^{(0)} = \vec{v}$. Denote the relation (5.112) by $\vec{a} = \vec{a}(W, \vec{v})$. Let T be a large constant to be determined, and for $0 \le t \le T-1$ define

$$\vec{a}^{(t)} = \vec{a}(W^{(t)}, \vec{v}^{(t)}), \qquad W^{(t+1)} = W(W^{(t)}, \vec{v}^{(t)}, \vec{a}^{(t)}), \qquad \vec{v}^{(t+1)} = \vec{v}(W^{(t)}, \vec{v}^{(t)}, \vec{a}^{(t)}).$$

Further define

$$E^{(t)} = \sum_{s \in \mathscr{S}} \lambda_s \sqrt{(v_s^{(t)})^2 + 2 \sum_{s' \in \mathscr{S}} \lambda_{s'} (w_{s,s'}^{(t)})^2}, \qquad F^{(t)} = \sum_{s \in \mathscr{S}} \lambda_s v_s^{(t)} \sqrt{a_s^{(t)}}.$$

Let $\delta > 0$ be arbitrary; we will show that $\mathbb{E}\mathrm{GS}_N(W, \vec{v}) \ge E^{(0)} - \delta$ for all sufficiently large N. Lemma 5.B.11 with the choice $\vec{a} = \vec{a}^{(t)}$ implies that

$$\mathbb{E}\mathrm{GS}_{N-tkr}(W^{(t)}, \vec{v}^{(t)}) \ge F^{(t)} + \mathbb{E}\mathrm{GS}_{N-(t+1)kr}(W^{(t+1)}, \vec{v}^{(t+1)}) - o_N(1),$$

and summing yields

$$\mathbb{E}\mathrm{GS}_N(W, \vec{v}) \ge \sum_{t=0}^{T-1} F^{(t)} - o_N(1).$$

Note that

$$F^{(t)} = \sum_{s \in \mathscr{S}} \lambda_s \sqrt{(v_s^{(t)})^2 + 2 \sum_{s' \in \mathscr{S}} \lambda_{s'} (w_{s,s'}^{(t)})^2 \cdot a_s^{(t)}},$$
$$E^{(t+1)} = \sum_{s \in \mathscr{S}} \lambda_s \sqrt{(v_s^{(t)})^2 + 2 \sum_{s' \in \mathscr{S}} \lambda_{s'} (w_{s,s'}^{(t)})^2 \cdot (1 - a_s^{(t)})},$$

so $F^{(t)} = E^{(t)} - E^{(t+1)}$. Thus

$$\mathbb{E}\mathrm{GS}_N(W, \vec{v}) \ge E^{(0)} - E^{(T)} - o_N(1).$$

Since

$$(v_s^{(t+1)})^2 + 2 \sum_{s' \in \mathscr{S}} \lambda_{s'} (w_{s,s'}^{(t+1)})^2 = 2(1 - a_s^{(t)}) \sum_{s' \in \mathscr{S}} \sum_{s'} \lambda_{s'} (w_{s,s'}^{(t)})^2$$

we have

$$a_s^{(t+1)} = \frac{(v_s^{(t+1)})^2}{(v_s^{(t+1)})^2 + 2 \sum_{s' \in \mathscr{S}} \lambda_{s'} (w_{s,s'}^{(t+1)})^2} = \frac{\sum_{s' \in \mathscr{S}} a_{s'}^{(t)} \lambda_{s'} (w_{s,s'}^{(t)})^2}{\sum_{s' \in \mathscr{S}} \lambda_{s'} (w_{s,s'}^{(t)})^2}.$$

It follows that, for $\alpha^{(t)} = \min_{s \in \mathscr{S}} a_s^{(t)}$, we have $\alpha^{(t+1)} \ge \alpha^{(t)}$. The assumption $\vec{v} > \vec{0}$ ensures $\alpha^{(0)} > 0$. Because $E^{(t+1)}/E^{(t)} \le 1 - \alpha^{(t)}$, we have

$$E^{(T)} \le E^{(0)} (1 - \alpha^{(0)})^T < \delta/2$$

for sufficiently large constant T. This implies $\mathbb{E}\mathrm{GS}_N(W, \vec{v}) \ge E^{(0)} - \delta/2 - o_N(1) \ge E^{(0)} - \delta$. This proves the result when $\vec{v} \succ \vec{0}$.

If some coordinates of $\vec{v}$ are zero, we apply this result to $\vec{v} + \eta \vec{1}$ for small $\eta > 0$. By (5.108),

$$\mathrm{GS}_N(W, \vec{v}) \ge \mathrm{GS}_N(W, \vec{v} + \eta \vec{1}) - \mathrm{GS}_N(0, \eta \vec{1}),$$

so for sufficiently large N,

$$\mathbb{E}\mathrm{GS}_N(W, \vec{v}) \ge \sum_{s \in \mathscr{S}} \lambda_s \sqrt{v_s^2 + 2 \sum_{s' \in \mathscr{S}} \lambda_{s'} w_{s,s'}^2} - \delta - \eta.$$

As this holds for any $\delta, \eta > 0$ the result follows. $\square$

Proof of Lemma 5.3.4. Follows from Corollary 5.B.6 and Proposition 5.B.8. $\square$

5.C Deferred proofs from Section 5.4

5.C.1 Existence of a maximizer: proof of Proposition 5.4.10

Proposition 5.4.10. *There exists a maximizer $(p,\Phi,q_0)\in\mathcal{M}$ for $\mathbb{A}$ and $\mathbb{A}(p,\Phi;q_0)<\infty$.*

Given $(p,\Phi,q_0)\in\mathcal{M}$ we extend p,Φ to domain $[0,1]$ by setting $p(q)=0$ for $q\in[0,q_0)$ and making Φ linear on $[0,q_0]$ with $\Phi(0)=\vec{0}$. Using this canonical extension we equip $\mathcal{M}$ with the metric

$$d\left((p^1,\Phi^1,q_0^1),(p^2,\Phi^2,q_0^2)\right)=\left\|p^1-p^2\right\|_{L^1([0,1])}+\left\|\Phi^1-\Phi^2\right\|_{L^1([0,1])}+|q_0^1-q_0^2|. \tag{5.119}$$

We will prove that $\mathcal{M}$ is a compact space on which $\mathbb{A}$ is upper semi-continuous. Existence of a triple $(p,\Phi;q_0)\in\mathcal{M}$ maximizing (5.7) within this space then follows.

Proposition 5.C.1. *The space $\mathcal{M}$ with metric (5.119) is compact.*

Proof. Given an infinite sequence $(p^n,\Phi^n,q_0^n)_{n\geq 0}$ of points in $\mathcal{M}$, we show there is a limit point. First find a subsequence (a_n) along which the convergence $q_0^{a_n}\to q_0$ holds. Then the subsequence $(p^{a_n})_{n\geq 0}$ has a subsubsequential limit in the space $L^1([q_0,1])$; similarly for $(\Phi_s^{a_n})_{n\geq 0}$, for each $s\in\mathscr{S}$. Thus we may choose a subsequence b_n of a_n on which $p^{b_n}\to p$ and $\Phi_s^{b_n}\to\Phi_s$ (for all $s\in\mathscr{S}$) in $L^1([q_0,1])$. It is easy to see that p and each Φ_s vanishes on $[0,q_0)$, and that Φ satisfies admissibility. It is easy to see that

$$\|p^{b_n}-p\|_{L^1([0,1])}\leq\|p^{b_n}-p\|_{L^1([q_0,1])}+|q_0-q_0^{b_n}|$$

and

$$\|\Phi_s^{b_n}-\Phi_s\|_{L^1([0,1])}\leq\|\Phi_s^{b_n}-\Phi_s\|_{L^1([q_0,1])}+|q_0-q_0^{b_n}|.$$

It follows that $(p^{b_n},\Phi^{b_n},q_0^{b_n})\to(p,\Phi,q_0)$ in $\mathcal{M}$. This completes the proof. $\square$

Proposition 5.C.2. *The function $\mathbb{A}$ is uniformly bounded on $\mathcal{M}$.*

Proof. For any admissible Φ we have by Cauchy-Schwarz

$$\begin{aligned}\sum_{s\in\mathscr{S}}\lambda_s\int_0^1\sqrt{\Phi_s'(q)(p\times\xi^s\circ\Phi)'(q)}\mathsf{d}q&\leq\sum_{s\in\mathscr{S}}\lambda_s\int_0^1\left(\Phi_s'(q)+(p\times\xi^s\circ\Phi)'(q)\right)\mathsf{d}q\\&\leq\sum_{s\in\mathscr{S}}\lambda_s\left(1+\xi^s(\vec{1})-\xi^s(\vec{0})\right).\end{aligned}\tag{5.120}$$

The first term of $\mathbb{A}$ is clearly uniformly bounded, so the result follows. $\square$

Proposition 5.C.3. *$\mathbb{A}$ is upper semi-continuous on $\mathcal{M}$.*

Proof. Suppose $(p^{b_n},\Phi^{b_n},q_0^{b_n})\to(p,\Phi,q_0)$ in $\mathcal{M}$. We write

$$\begin{aligned}|\mathbb{A}(p^{b_n},\Phi^{b_n};q_0^{b_n})-\mathbb{A}(p,\Phi;q_0)|\leq&\int_{q_0}^1\left|\sqrt{(\Phi_s^{b_n})'(q)(p^{b_n}\times\xi^s\circ\Phi^{b_n})'(q)}-\sqrt{\Phi_s'(q)(p\times\xi^s\circ\Phi)'(q)}\right|\mathsf{d}q\\&+C_\lambda\left(\left|\int_{q_0^{b_n}}^{q_0}\sqrt{p'(q)+1}\,\mathsf{d}q\right|+\sum_{s\in\mathscr{S}}\left|\sqrt{\Phi_s^n(q_0^n)}-\sqrt{\Phi_s(q_0)}\right|\right).\end{aligned}$$

The sum over $s\in\mathscr{S}$ obviously tends to 0. Moreover by Cauchy–Schwarz,

$$\begin{aligned}\left|\int_{q_0^{b_n}}^{q_0}\sqrt{p'(q)+1}\,\mathsf{d}q\right|&\leq|q_0-q_0^{b_n}|^{1/2}\cdot\sqrt{p(q_0)-p(q_0^{b_n})+1}\\&\leq C_\lambda'(q_0-q_0^{b_n})^{1/2}.\end{aligned}$$

Therefore it suffices to show the first term above tends to 0. Since the map

$$(p,\Phi)\mapsto(p\times\xi^s\circ\Phi)$$

from $L^1([0,1])^{|\mathscr{S}|+1} \to L^1([0,1])$ is continuous and returns a non-decreasing function, it suffices to show that

$$G(f,g) = \int_0^1 \sqrt{f'(q)g'(q)}\,\mathsf{d}q$$

is upper semi-continuous on $L^1([0,1]) \times L^1([0,1])$ when restricted to non-decreasing functions. This is essentially equivalent to upper semi-continuity of Hellinger distance which is well-known. $\square$

Combining the results above implies Proposition 5.4.10.

5.C.2 A priori regularity of maximizers

Let $(p, \Phi, q_0) \in \mathcal{M}$ be a maximizer of $\mathbb{A}$, which exists by Proposition 5.4.10. In this subsection we will prove the following two propositions.

Proposition 5.4.11. *The functions p, Φ are continuously differentiable on $[q_0+\varepsilon, 1]$ for any $\varepsilon > 0$. Moreover, there exists $L > 0$ (possibly depending on $(p, \Phi; q_0)$ as well as ξ) such that $L^{-1}\vec{1} \preceq \Phi'(q) \preceq L\vec{1}$ for almost all $q \in (q_0, 1]$.*

Proposition 5.4.12. *The function p satisfies $p(q) > 0$ for all $q > q_0$, $p(1) = 1$, and $p(q_0) = 0$ if $q_0 > 0$.*

Lemma 5.C.4. *The function p is absolutely continuous and $p(1) = 1$. Moreover p' is uniformly bounded on compact subsets of $(q_0, 1)$.*

Proof. Given any non-decreasing $p : [q_0, 1] \to [q_0, 1]$, we may treat p' as a positive measure via

$$p'(x)dx = f(x)dx + \mu(dx) \tag{5.121}$$

for μ a singular-plus-atomic measure and $f \in L^1([q_0, 1]; \mathbb{R}_{\geq 0})$. We may then replace p by $\bar{p}$ such that

$$\bar{p}'(x)dx = f(x)dx, \quad \text{and} \quad \bar{p}(1) = 1.$$

Then $\bar{p}(x) \geq p(x)$ for all $x \in [q_0, 1]$, and $\bar{p}'(x)$ agrees with $p'(x)$ except for a singular-plus-atomic part. It follows that

$$\mathbb{A}(p, \Phi; q_0) \leq \mathbb{A}(\bar{p}, \Phi; q_0).$$

Moreover it is easy to see that strict inequality $\mathbb{A}(p, \Phi; q_0) < \mathbb{A}(\bar{p}, \Phi; q_0)$ holds whenever $p \neq \bar{p}$. We conclude that p is absolutely continuous and $p(1) = 1$.

To show the latter statement, we use a similar argument with more care. Let $q \in (q_0, 1)$ and choose a large constant $C = C(q_0, q)$. Recalling (5.121), suppose $\|f(x)\|_{L^\infty([q,1])} > C$ for a large constant C and let

$$c \equiv \frac{\int_q^1 (f(x) - C)_+ \,\mathsf{d}x}{q - q_0}.$$

We may replace f by

$$f_C(x) = \begin{cases} f(x), & x \in [0, q_0) \\ f(x) + c, & x \in [q_0, q) \\ \min(C, f(x)), & x \in [q, 1] \end{cases}$$

and similarly replace p by p_C with

$$\bar{p}_C'(x)dx = f_C(x)\,\mathsf{d}x, \quad \text{and} \quad \bar{p}(1) = 1.$$

It is easy to see that $p_C(x) \geq p(x)$ for each $x \in [0,1]$. Keeping Φ the same, we consider the change in $\mathbb{A}$. The decrease in $\mathbb{A}$ on $[q,1]$ is at most

$$
\begin{aligned}
&\sum_{s\in\mathscr{S}} \lambda_s \int_q^1 \sqrt{\Phi_s'(x)(p \times \xi^s \circ \Phi)'(x)} - \sqrt{\Phi_s'(x)(p_C \times \xi^s \circ \Phi)'(x)}\ \mathsf{d}x \\
&= \sum_{s\in\mathscr{S}} \lambda_s \int_q^1 \sqrt{\Phi_s'(x)\left(p'(x)\xi^s\big(\Phi(x)\big) + p(x)\langle \Phi'(x), \nabla\xi^s\big(\Phi(x)\big)\rangle\right)} \\
&\qquad\qquad - \sqrt{\Phi_s'(x)\left(p_C'(x)\xi^s\big(\Phi(x)\big) + p_C(x)\langle \Phi'(x), \nabla\xi^s\big(\Phi(x)\big)\rangle\right)}\ \mathsf{d}x \\
&\leq \sum_{s\in\mathscr{S}} \lambda_s \int_q^1 \sqrt{\Phi_s'(x)\left(p'(x)\xi^s\big(\Phi(x)\big) + p(x)\langle \Phi'(x), \nabla\xi^s\big(\Phi(x)\big)\rangle\right)} \\
&\qquad\qquad - \sqrt{\Phi_s'(x)\left(p_C'(x)\xi^s\big(\Phi(x)\big) + p(x)\langle \Phi'(x), \nabla\xi^s\big(\Phi(x)\big)\rangle\right)}\ \mathsf{d}x \qquad (5.122) \\
&\leq \sum_{s\in\mathscr{S}} \lambda_s \int_q^1 \sqrt{\Phi_s'(x)p'(x)\cdot\xi^s\big(\Phi(x)\big)} - \sqrt{\Phi_s'(x)p_C'(x)\cdot\xi^s\big(\Phi(x)\big)}\ \mathsf{d}x \\
&\leq O(1)\cdot \int_q^1 \sqrt{p'(x)} - \sqrt{p_C'(x)}\mathsf{d}x \\
&\leq O(1)\cdot \int_q^1 C^{-1/2}(f(x)-C)_+\ \mathsf{d}x \\
&\leq O\left(\frac{c(q-q_0)}{\sqrt{C}}\right).
\end{aligned}
$$

(In the second inequality we used $\sqrt{x+z} - \sqrt{y+z} \leq \sqrt{x} - \sqrt{y}$ for $x \geq y \geq 0$, and in the third we used that Φ_s' is uniformly bounded by admissibility.) On $x \in [q_0, q]$, we find that changing from p to p_C increases the value of $\mathbb{A}$:

$$
\begin{aligned}
&\sum_{s\in\mathscr{S}} \lambda_s \int_{q_0}^q \sqrt{\Phi_s'(x)(p_C \times \xi^s \circ \Phi)'(x)} - \sqrt{\Phi_s'(x)(p \times \xi^s \circ \Phi)'(x)}\ \mathsf{d}x \\
&= \sum_{s\in\mathscr{S}} \lambda_s \int_{q_0}^q \sqrt{\Phi_s'(x)\left(p_C'(x)\xi^s\big(\Phi(x)\big) + p_C(x)\langle \Phi'(x), \nabla\xi^s\big(\Phi(x)\big)\rangle\right)} \\
&\qquad - \sqrt{\Phi_s'(x)\left(p'(x)\xi^s\big(\Phi(x)\big) + p(x)\langle \Phi'(x), \nabla\xi^s\big(\Phi(x)\big)\rangle\right)}\ \mathsf{d}x \\
&\geq \sum_{s\in\mathscr{S}} \lambda_s \int_{q_0}^q \sqrt{\Phi_s'(x)\left((p'(x)+c)\xi^s\big(\Phi(x)\big) + p(x)\langle \Phi'(x), \nabla\xi^s\big(\Phi(x)\big)\rangle\right)} \\
&\qquad - \sqrt{\Phi_s'(x)\left(p'(x)\xi^s\big(\Phi(x)\big) + p(x)\langle \Phi'(x), \nabla\xi^s\big(\Phi(x)\big)\rangle\right)}\ \mathsf{d}x \\
&\geq \Omega(c)\cdot \int_{q_0}^q \sum_{s\in\mathscr{S}} \frac{\mathsf{d}x}{\sqrt{\Phi_s'(x)\left((p'(x)+c)\xi^s\big(\Phi(x)\big) + p(x)\langle \Phi'(x), \nabla\xi^s\big(\Phi(x)\big)\rangle\right)}}.
\end{aligned}
$$

By Markov's inequality, $p'(x) \leq \frac{2}{(q-q_0)}$ on a set of $x \in [q_0, q]$ of measure at least $\frac{q-q_0}{2}$. For each such x, we have $\Phi_s'(x) \leq O(1)$ and $\xi^s(\Phi(x)) \leq O(1)$. We thus find

$$
\Omega(c)\cdot \int_{q_0}^q \sum_{s\in\mathscr{S}} \frac{\mathsf{d}x}{\sqrt{\Phi_s'(x)\left((p'(x)+c)\xi^s\big(\Phi(x)\big) + p(x)\langle \Phi'(x), \nabla\xi^s\big(\Phi(x)\big)\rangle\right)}} \geq \Omega\left(\frac{c(q-q_0)}{\sqrt{q-q_0+c}}\right)
$$

Since $c \leq \frac{1}{q-q_0}$, for C sufficiently large, combining with (5.122) above implies that

$$
\sum_{s\in\mathscr{S}} \lambda_s \int_q^1 \sqrt{\Phi_s'(x)(p_C \times \xi^s \circ \Phi)'(x)} - \sqrt{\Phi_s'(x)(p \times \xi^s \circ \Phi)'(x)}\ \mathsf{d}x > 0.
$$

Since $p(x) = p_C(x)$ for $x \le q_0$, we find $\mathbb{A}(p, \Phi; q_0) < \mathbb{A}(p_C, \Phi; q_0)$, contradicting maximality of $\mathbb{A}(p, \Phi; q_0)$. Having reached a contradiction for C sufficiently large, we conclude that p' is uniformly bounded on $[q, 1]$ for each $q \in (q_0, 1)$ as desired. $\square$

Lemma 5.C.5. *$p(q) > 0$ holds for all $q > q_0$.*

Proof. Suppose not. Then $p(q) = 0$ for all $q \in [q_0, q_0 + \varepsilon]$, for some $\varepsilon > 0$. For $\delta > 0$ small, define

$$p_\delta(q) = \delta + (1 - \delta)p(q).$$

Then

$$\begin{aligned}\sum_{s\in\mathscr{S}} \lambda_s \int_{q_0}^{q_0+\varepsilon} \sqrt{\Phi_s'(q)(p_\delta \times \xi^s \circ \Phi)'(q)}\mathsf{d}q &= \delta^{1/2} \sum_{s\in\mathscr{S}} \lambda_s \int_{q_0}^{q_0+\varepsilon} \sqrt{\Phi_s'(q)(\xi^s \circ \Phi)'(q)}\mathsf{d}q \\ &\ge \delta^{1/2} c(\xi) \sum_{s\in\mathscr{S}} \lambda_s \int_{q_0}^{q_0+\varepsilon} \sqrt{\Phi_s'(q)^2}\mathsf{d}q \\ &= \delta^{1/2} c(\xi).\end{aligned}$$

while

$$\sum_{s\in\mathscr{S}} \lambda_s \int_{q_0}^{q_0+\varepsilon} \sqrt{\Phi_s'(q)(p \times \xi^s \circ \Phi)'(q)}\mathsf{d}q = 0.$$

On the other hand since $p_\delta(q) \ge p(q)$ for all $q \in [q_0, 1]$ and $(p_\delta)' = (1 - \delta)(p)'$ as measures, we obtain

$$\sum_{s\in\mathscr{S}} \lambda_s \int_{q_0+\varepsilon}^{1} \sqrt{\Phi_s'(q)(p_\delta \times \xi^s \circ \Phi)'(q)}\mathsf{d}q \ge (1 - \delta) \sum_{s\in\mathscr{S}} \lambda_s \int_{q_0+\varepsilon}^{1} \sqrt{\Phi_s'(q)(p \times \xi^s \circ \Phi)'(q)}\mathsf{d}q.$$

Combining the above implies $\mathbb{A}(p_\delta, \Phi; q_0) > \mathbb{A}(p, \Phi; q_0)$ for small enough δ, a contradiction. $\square$

Lemma 5.C.6. *For all $s \in \mathscr{S}$ and $q \in (0, 1)$, we have $\Phi_s(q) < 1$.*

Proof. Suppose $\Phi_{s_0}(q_*) = 1$; this implies $0 < q_0 \le q_* < 1$. For small $\delta > 0$ we consider the perturbation Φ_δ with $\Phi_{\delta,s} = \Phi_s$ for $s \ne s_0$ and:

$$\Phi_{\delta,s_0}'(q) = \begin{cases} \Phi_{s_0}'(q) \cdot (1 - \delta(1 - q_*)), & q \in [0, q_*], \\ \delta, & q \in [q_*, 1] \end{cases}$$

Note that $\Phi_{\delta,s}'(q) \ge (1 - O(\delta))\Phi_s'(q)$ and so also $\Phi_{\delta,s}(q) \ge (1 - O(\delta))\Phi_s(q)$ for all $s \in \mathscr{S}$ and $q \in [0, 1]$. As $\mathbb{A}$ is uniformly bounded, we can thus bound

$$\begin{aligned}\mathbb{A}(p, \Phi_\delta; q_0) - \mathbb{A}(p, \Phi; q_0) &= \sum_{s\in\mathscr{S}} h_s\lambda_s\sqrt{\Phi_{\delta,s}(q_0)} + \lambda_s \int_{q_0}^{1} \sqrt{\Phi_{\delta,s}'(q)(p \times \xi^s \circ \Phi_\delta)'(q)}\ \mathsf{d}q \\ &\quad - \sum_{s\in\mathscr{S}} h_s\lambda_s\sqrt{\Phi_s(q_0)} - \lambda_s \int_{q_0}^{1} \sqrt{\Phi_s'(q)(p \times \xi^s \circ \Phi)'(q)}\ \mathsf{d}q \\ &\ge -O(\delta) + \lambda_{s_0} \int_{\frac{1+q_*}{2}}^{1} \sqrt{\Phi_{\delta,s}'(q)(p \times \xi^s \circ \Phi)'(q)} - \sqrt{\Phi_s'(q)(p \times \xi^s \circ \Phi)'(q)}\ \mathsf{d}q.\end{aligned}$$

Using Lemma 5.C.5, admissibility and non-degeneracy of ξ, we find that $(p \times \xi^s \circ \Phi)'(q) \ge \Omega(q)$ for all $q \ge \frac{1+q_*}{2}$. Therefore

$$\lambda_{s_0} \int_{\frac{1+q_*}{2}}^{1} \sqrt{\Phi_{\delta,s}'(q)(p \times \xi^s \circ \Phi)'(q)} - \sqrt{\Phi_s'(q)(p \times \xi^s \circ \Phi)'(q)}\ \mathsf{d}q \ge \Omega(\delta^{1/2})$$

for small δ. Since $\delta^{1/2}$ is of larger order than δ we conclude that $\mathbb{A}(p, \Phi_\delta; q_0) > \mathbb{A}(p, \Phi; q_0)$. This is a contradiction (recall Lemma 5.4.9) and completes the proof. $\square$

Next we turn our attention to Φ'. Similarly to Lemma 5.C.5, the idea is that the square root function has infinite derivative at 0.

Lemma 5.C.7. *There exists $\eta > 0$ such that $\Phi'(q) \succeq \eta\vec{1}$ almost everywhere in $q \in [q_0, 1]$.*

Proof. First, given $(p, \Phi; q_0)$ choose for some $s \in \mathscr{S}$ (specified below) a Lebesgue point $q_s \in (q_0, 1)$ of Φ' with

$$\Phi'_s(q_s) \geq a \tag{5.123}$$

for $a > 0$. Lemma 5.C.6 ensures this is possible for some a depending only on q_0 and $\Phi(q_0)$ (as long as $q_0 < 1$, else there is nothing to prove). In fact we can actually find two distinct such points $q_s^{(1)}, q_s^{(2)}$ (which will be helpful below).

Next for small $\varepsilon > 0$ depending only on (p, Φ), define the interval

$$J_{s,\varepsilon} = (q_s - \varepsilon, q_s + \varepsilon).$$

By (5.123) and the fact that q_s is a Lebesgue point of Φ', there is a subset $I_{s,\varepsilon} \subseteq J_{s,\varepsilon}$ of Lebesgue measure at least $|I_{s,\varepsilon}| \geq \frac{|J_{s,\varepsilon}|}{2} = \varepsilon$ such that

$$\Phi'_s(q) \geq \frac{a}{2}, \quad \forall q \in I_{s,\varepsilon}. \tag{5.124}$$

as long as $\varepsilon > 0$ is chosen sufficiently small. A simple consequence is the estimate

$$C_\varepsilon := \Phi_s(q_s + \varepsilon) - \Phi_s(q_s - \varepsilon) = \int_{q_s-\varepsilon}^{q_s+\varepsilon} \Phi'_s(q)\mathsf{d}q \geq \frac{a\varepsilon}{2}. \tag{5.125}$$

With the setup above complete (except that s is not yet specified), suppose the conclusion is false let η be sufficiently small depending on (p, Φ, ε) for ε as above. Then there exist $s, s_0 \in \mathscr{S}$ and $\hat{q}_{s_0} \in (q_0, 1)$ which is a Lebesgue point for $\nabla\Phi$ such that

$$\Phi'_s(\hat{q}_{s_0}) \leq \eta, \tag{5.126}$$
$$\Phi'_{s_0}(\hat{q}_{s_0}) \geq 1. \tag{5.127}$$

Indeed if q is any Lebesgue point of $\nabla\Phi$ satisfying (5.126) for some s, then (5.127) holds for some $s_0 \neq s$ by admissibility and we define $\hat{q}_{s_0} = q$ this way. The bound (5.126) determines the species s chosen initially.

As $\hat{q}_{s_0}$ is also a Lebesgue point of Φ', in light of (5.126) and (5.127), there exists a set $I_{s_0,\eta} \subseteq J_{s_0,\varepsilon} = (\hat{q}_{s_0} - \varepsilon, \hat{q}_{s_0} + \varepsilon)$ of positive Lebesgue measure such that the inequalities

$$\Phi'_s(q) \leq 2\eta, \tag{5.128}$$
$$\Phi'_{s_0}(q) \geq \frac{a}{2}. \tag{5.129}$$

both hold for all $q \in I_{s_0,\eta}$. Moreover we can assume $J_{s,\varepsilon}, J_{s_0,\varepsilon}$ are disjoint, i.e. $|q_s - \hat{q}_{s_0}| > 2\varepsilon$. Indeed as noted earlier we can choose two candidate points $q_s^{(1)}, q_s^{(2)}$. If $\varepsilon < |q_s^{(1)} - q_s^{(2)}|/5$ is taken, at least one of them suffices for any $\hat{q}_{s_0} \in (q_0, 1)$.

Next choose $\delta \in (0, \eta)$ small and consider the perturbation Φ_δ with $\Phi_\delta(q_0) = \Phi(q_0)$ and

$$\Phi'_{\delta,s}(q) = \begin{cases} \Phi'_s(q) + \delta, & q \in I_{s_0,\eta} \\ \Phi'_s(q)\left(1 - \frac{\delta|I_{s_0,\eta}|}{C_\varepsilon}\right), & \forall q \in J_{s,\varepsilon} \\ \Phi'_s(q), & \text{otherwise} \end{cases}$$

and $\Phi_{\delta,s'} = \Phi_{s'}$ for all $s' \in \mathscr{S}\backslash\{s\}$. (Note we used disjointness of $J_{s,\varepsilon}, J_{s_0,\varepsilon}$ for this definition to make sense.) By Lemma 5.4.9, we must have $\mathbb{A}(p, \Phi_\delta; q_0) \geq \mathbb{A}(p, \Phi; q_0)$ although Φ_δ may not be admissible. Then for

$\delta \le \eta$,

$$\int_{I_{s_0,\eta}} \sqrt{\Phi'_{\delta,s}(q)(p\times\xi^s\circ\Phi)'(q)} - \sqrt{\Phi'_s(q)(p\times\xi^s\circ\Phi)'(q)}\mathsf{d}q$$
$$\begin{aligned}
&\overset{(5.128)}{\ge} (\sqrt{2\eta+\delta}-\sqrt{2\eta})\int_{I_{s_0,\eta}} \sqrt{(p\times\xi^s\circ\Phi)'(q)}\mathsf{d}q \\
&\ge \frac{\delta p(q_s-\varepsilon)^{1/2}}{10\eta^{1/2}}\int_{I_{s_0,\eta}} \sqrt{(\xi^s\circ\Phi)'(q)}\mathsf{d}q \\
&\ge \frac{\delta p(q_s/2)^{1/2}c(\xi)}{10\eta^{1/2}}\int_{I_{s_0,\eta}} \sqrt{\Phi'_{s_0}(q)}\mathsf{d}q \\
&\overset{(5.129)}{\ge} \frac{\delta a^{1/2}p(q_s/2)^{1/2}c(\xi)|I_{s_0,\eta}|}{20\eta^{1/2}}.
\end{aligned} \tag{5.130}$$

We used non-degeneracy of ξ in the penultimate step. On the other hand recalling (5.125), it follows that for all $\widetilde{s}\in\mathscr{S}$ and almost all $q\in[q_0,1]$:

$$\Phi'_{\delta,\widetilde{s}}(q) \ge \left(1-O\left(\frac{\delta|I_{s_0,\eta}|}{\varepsilon}\right)\right)\Phi'_{\widetilde{s}}(q). \tag{5.131}$$

Integrating on $[q_0,q]$, we find

$$\Phi_{\delta,\widetilde{s}}(q) \ge \left(1-O\left(\frac{\delta|I_{s_0,\eta}|}{\varepsilon}\right)\right)\Phi_{\widetilde{s}}(q) \tag{5.132}$$

for all $q\in[q_0,1]$. By the chain rule we similarly obtain that for all $\widetilde{s}\in\mathscr{S}$,

$$(p\times\xi^{\widetilde{s}}\circ\Phi_\delta)' \ge \left(1-O\left(\frac{\delta|I_{s_0,\eta}|}{\varepsilon}\right)\right)(p\times\xi^{\widetilde{s}}\circ\Phi)', \tag{5.133}$$
$$(p\times\xi^{\widetilde{s}}\circ\Phi_\delta)(q) \ge \left(1-O\left(\frac{\delta|I_{s_0,\eta}|}{\varepsilon}\right)\right)(p\times\xi^{\widetilde{s}}\circ\Phi)(q). \tag{5.134}$$

It follows from (5.131), (5.132), (5.133), (5.134) that

$$\begin{aligned}
\int_{I_{s_0,\eta}} \sqrt{\Phi'_{\delta,s}(q)(p\times\xi^s\circ\Phi_\delta)'(q)} &\ge \left(1-O\left(\frac{\delta|I_{s_0,\eta}|}{\varepsilon}\right)\right)\int_{I_{s_0,\eta}} \sqrt{\Phi'_{\delta,s}(q)(p\times\xi^s\circ\Phi)'(q)}\mathsf{d}q \\
&\overset{(5.120)}{\ge} \int_{I_{s_0,\eta}} \sqrt{\Phi'_{\delta,s}(q)(p\times\xi^s\circ\Phi)'(q)}\mathsf{d}q - O\left(\frac{\delta|I_{s_0,\eta}|}{\varepsilon}\right).
\end{aligned} \tag{5.135}$$

Since Φ_δ and Φ differ only inside $[q_0,1]$ we use $\mathbb{A}_{[q_0,1]}$ below to denote the second term of $\mathbb{A}$. We have:

$$\begin{aligned}
\mathbb{A}_{[q_0,1]}(p,\Phi) &= \sum_{\widetilde{s}\in\mathscr{S}}\lambda_{\widetilde{s}}\int_{q_0}^1 \sqrt{\Phi'_{\widetilde{s}}(q)(p\times\xi^{\widetilde{s}}\circ\Phi)'(q)}\mathsf{d}q \\
&= \lambda_s\int_{I_{s_0,\eta}} \sqrt{\Phi'_s(q)(p\times\xi^s\circ\Phi)'(q)}\mathsf{d}q + \lambda_s\int_{[q_0,1]\setminus I_{s_0,\eta}} \sqrt{\Phi'_s(q)(p\times\xi^s\circ\Phi)'(q)}\mathsf{d}q \\
&\quad + \sum_{\widetilde{s}\in\mathscr{S}\setminus\{s\}}\lambda_{\widetilde{s}}\int_{q_0}^1 \sqrt{\Phi'_{\widetilde{s}}(q)(p\times\xi^{\widetilde{s}}\circ\Phi)'(q)}\mathsf{d}q \\
&\equiv \mathrm{I}+\mathrm{I}I+\mathrm{I}II.
\end{aligned}$$

Similarly for J instead of I,

$$\begin{aligned}\mathbb{A}_{[q_0,1]}(p,\Phi_\delta) &= \sum_{\widetilde{s}\in\mathscr{S}} \lambda_{\widetilde{s}} \int_{q_0}^1 \sqrt{\Phi'_{\delta,\widetilde{s}}(q)(p\times\xi^{\widetilde{s}}\circ\Phi_\delta)'(q)}\mathsf{d}q \\ &= \lambda_s \int_{I_{s_0,\eta}} \sqrt{(\Phi_{\delta,s})'(q)(p\times\xi^s\circ\Phi_\delta)'(q)}\mathsf{d}q + \lambda_s \int_{[q_0,1]\setminus I_{s_0,\eta}} \sqrt{(\Phi_{\delta,s})'(q)(p\times\xi^s\circ\Phi_\delta)'(q)}\mathsf{d}q \\ &\quad + \sum_{\widetilde{s}\in\mathscr{S}\setminus\{s\}} \lambda_{\widetilde{s}} \int_{q_0}^1 \sqrt{\Phi'_{\delta,\widetilde{s}}(q)(p\times\xi^{\widetilde{s}}\circ\Phi_\delta)'(q)}\mathsf{d}q \\ &\equiv \mathrm{I}_\delta + \mathrm{II}_\delta + \mathrm{I}II_\delta.\end{aligned}$$

Using (5.131), (5.132), (5.133), (5.134) again, we obtain

$$\mathrm{II}_\delta \geq \left(1 - O\left(\frac{\delta|I_{s_0,\eta}|}{\varepsilon}\right)\right)\mathrm{I}I,$$
$$\mathrm{I}II_\delta \geq \left(1 - O\left(\frac{\delta|I_{s_0,\eta}|}{\varepsilon}\right)\right)\mathrm{I}II.$$

Meanwhile (5.130) and (5.135) imply that for δ small compared to η,

$$\mathrm{I}_\delta \geq \left(1 - O\left(\frac{\delta|I_{s_0,\eta}|}{\varepsilon}\right)\right)\mathrm{I} + \frac{\delta a^{1/2}p(q_s/2)^{1/2}c(\xi)|I_{s_0,\eta}|}{20\eta^{1/2}}.$$

Combining, we find

$$\mathbb{A}_{[q_0,1]}(p,\Phi_\delta) \geq \mathbb{A}_{[q_0,1]}(p,\Phi) + \frac{\delta a^{1/2}p(q_s/2)^{1/2}c(\xi)|I_{s_0,\eta}|}{20\eta^{1/2}} - O\left(\frac{\delta|I_{s_0,\eta}|}{\varepsilon}\right).$$

Taking $\eta \ll \varepsilon^2 ap(q_s/2)c(\xi)^2$ and then δ sufficiently small contradicts the maximality of (p,Φ,q_0), thus completing the proof. □

Proposition 5.C.8. *If $q_0 > 0$, then $p(q_0) = 0$.*

Proof. Assume that $p(q_0) > 0$. Consider the perturbation

$$\widetilde{p}(q) = \begin{cases} p(q) + (q - q_0 - \varepsilon)\delta, & q < q_0 + \varepsilon \\ p(q), & q \geq q_0 + \varepsilon. \end{cases}$$

The function $\tilde{p}$ is non-decreasing, and is non-negative for sufficiently small $\varepsilon, \delta > 0$. For $q < q_0 + \varepsilon$ we find

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\delta}(p\times\xi^s\circ\Phi)'(q) &= \frac{\mathrm{d}}{\mathrm{d}\delta}\left(p'(q)\xi^s(\Phi(q)) + p(q)(\xi^s\circ\Phi)'(q)\right) \\ &= \xi^s(\Phi(q)) - (q_0 + \varepsilon - q)(\xi^s\circ\Phi)'(q) \\ &\geq \xi^s(\Phi(q)) - O(\varepsilon).\end{aligned}$$

If $q_0 > 0$, then $\xi^s(\Phi(q)) \geq c(q_0) > 0$ by admissibility and non-degeneracy of Φ. This contradicts optimality of (p,Φ,q_0) and completes the proof. □

Proof of Proposition 5.4.12. Follows from Lemmas 5.C.4 and 5.C.5 and Proposition 5.C.8. □

Continuous differentiability on $(q_0, 1]$

Here we show that p and Φ are continuously differentiable on compact subsets of $(q_0, 1]$ using another local perturbation argument.

Lemma 5.C.9. *The function $f(x,y) = \sqrt{xy}$ is concave on $\mathbb{R}^2_{>0}$, with strict concavity on all lines except for those passing through the origin.*

Proof. Given $x_0, y_0, x_1, y_1 > 0$ with $(x_0, y_0) \neq (x_1, y_1)$ and $c \in (0,1)$, we have

$$\begin{aligned}
(x_0 y_1 - x_1 y_0)^2 &\geq 0 \\
\implies x_0^2 y_1^2 + x_1^2 y_0^2 + 2x_0 x_1 y_0 y_1 &\geq 4x_0 x_1 y_0 y_1 \\
\implies (x_0 y_1 + x_1 y_0) &\geq 2\sqrt{x_0 x_1 y_0 y_1} \\
\implies c(1-c) \cdot (x_0 y_1 + x_1 y_0) &\geq 2c(1-c)\sqrt{x_0 x_1 y_0 y_1} \\
\implies c^2 x_0 y_0 + (1-c)^2 x_0 y_0) + c(1-c) \cdot (x_0 y_1 + x_1 y_0) &\geq c^2 x_0 y_0 + (1-c)^2 x_0 y_0) + 2c(1-c)\sqrt{x_0 x_1 y_0 y_1} \\
\implies \sqrt{(cx_0 + (1-c)x_1)(cy_0 + (1-c)y_1)} &\geq c\sqrt{x_0 y_0} + (1-c)\sqrt{x_1 y_1}.
\end{aligned}$$

Moreover equality holds if and only if it holds in the first step. □

Lemma 5.C.10. *Both p and Φ are continuously differentiable on compact subsets of $(q_0, 1]$.*

Proof. We assume that $q_0 < 1$ (else there is nothing to prove), and recall Lemma 5.C.5 throughout. Admissibility implies that Φ is uniformly Lipschitz, and Lemma 5.C.4 shows that p is uniformly Lipschitz on compact subsets of $(q_0, 1)$. Hence both $p'(x)$ and Φ'_s exist as non-negative, integrable functions which are uniformly bounded away from q_0.

By an elementary result of [Zaa86], if a measurable function $[q_0, 1] \to \mathbb{R}$ does not agree with any continuous function on a full measure set, then it possesses a genuine point of discontinuity $q_* \in (q_0, 1)$ such that F cannot be made continuous at q_* even by modification on a measure zero set. We fix such a point q_* for sake of contradiction. By definition, this means that for some $\eta > 0$ depending only on (p, Φ, q_*) and for arbitrarily small $\varepsilon > 0$, there exist measurable sets $I, J \subseteq (q_* - \varepsilon, q_* + \varepsilon)$ and $a \in \mathbb{R}$ such that:

$$\begin{aligned}
|I| &= \varepsilon_1 > 0, \\
|J| &= \varepsilon_1 > 0, \\
f(q) &\geq a + \eta, \quad \forall q \in I, \\
f(q) &\leq a - \eta, \quad \forall q \in J.
\end{aligned} \tag{5.136}$$

Here $f(q) = p'(q)$ or $f(q) = \Phi'_s(q)$ for some $s \in \mathscr{S}$.

Let $\gamma_I : [0, \varepsilon_1] \to I$ and $\gamma_J : [0, \varepsilon_1] \to J$ be increasing, measure-preserving bijections (and note that their inverse functions are also measurable). For convenience we set $q_{I,x} = \gamma_I(x)$ and $q_{J,x} = \gamma_J(x)$. We construct perturbations $\tilde{p}, \tilde{\Phi}$ of p and Φ by averaging derivatives on $q_{I,x}$ and $q_{J,x}$:

$$\begin{aligned}
\tilde{p}'(q_{I,x}) = \tilde{p}'(q_{J,x}) &= \frac{p'(q_{I,x}) + p'(q_{J,x})}{2}\,; \\
\tilde{p}'(q) &= p'(q), \quad q \notin I \cup J\,; \\
\tilde{\Phi}'_s(q_{I,x}) = \tilde{\Phi}'_s(q_{J,x}) &= \frac{\Phi'_s(q_{I,x}) + \Phi'_s(q_{J,x})}{2}\,; \\
\tilde{\Phi}'_s(q) &= \Phi'_s(q), \quad q \notin I \cup J.
\end{aligned}$$

We claim that for fixed q_*, η and sufficiently small $\varepsilon > 0$, we have

$$\mathbb{A}(\tilde{p}, \tilde{\Phi}; q_0) > \mathbb{A}(p, \Phi; q_0). \tag{5.137}$$

This contradicts maximality of (p, Φ) and thus implies the desired continuity of (p', Φ').

To begin proving (5.137), recall from Lemma 5.C.4 that p' is uniformly bounded away from q_0, hence on $(q_* - \varepsilon, q_* + \varepsilon)$. Moreover Φ' is uniformly bounded by definition. It follows that for all $s \in \mathscr{S}$ and $q \in (q_* - \varepsilon, q_* + \varepsilon)$,

$$\begin{aligned}
|p(q) - \tilde{p}(q)| &\leq O(\varepsilon_1), \\
|p(q) - p(q_*)| &\leq O(\varepsilon), \\
|\Phi_s(q) - \tilde{\Phi}_s(q)| &\leq O(\varepsilon_1), \\
|\xi^s(\Phi(q)) - \xi^s(\tilde{\Phi}(q))| &\leq O(\varepsilon_1), \\
|\Phi_s(q) - \Phi_s(q_*)| &\leq O(\varepsilon), \\
|\xi^s(\Phi(q)) - \xi^s(\Phi(q_*))| &\leq O(\varepsilon).
\end{aligned} \tag{5.138}$$

These estimates will let us treat the above functions as almost constant while proving (5.137), so we can focus on the more important changes in their derivatives. First for $q \notin [q_* - \varepsilon, q_* + \varepsilon]$, we have $p(q) = \tilde{p}(q)$ and $\Phi(q) = \tilde{\Phi}(q)$, so it suffices to analyze the discrepancy within $q \in [q_* - \varepsilon, q_* + \varepsilon]$. Next, the estimates (5.138) together with the fact that Φ_s' is uniformly bounded below (by Lemma 5.C.7) imply that

$$\left|\sqrt{\Phi_s'(q)(p \times \xi^s \circ \Phi)'(q)} - \sqrt{\tilde{\Phi}_s'(q)(\tilde{p} \times \xi^s \circ \tilde{\Phi})'(q)}\right| \le O(\varepsilon_1), \quad \forall\ q \in [q_* - \varepsilon, q_* + \varepsilon] \backslash (I \cup J). \tag{5.139}$$

Integrating, we obtain

$$\int_{q \in [q_* - \varepsilon, q_* + \varepsilon] \backslash (I \cup J)} \left|\sqrt{\Phi_s'(q)(p \times \xi^s \circ \Phi)'(q)} - \sqrt{\tilde{\Phi}_s'(q)(\tilde{p} \times \xi^s \circ \tilde{\Phi})'(q)}\right| \, \mathsf{d}q \le O(\varepsilon_1 \varepsilon). \tag{5.140}$$

Next we fix $x \in [0, \varepsilon_1]$ and analyze the joint effect of the pertubation at the pair of points $q_{I,x}$ and $q_{J,x}$. This is given by

$$\begin{aligned} &\sqrt{\tilde{\Phi}_s'(q_{I,x})(\tilde{p} \times \xi^s \circ \tilde{\Phi})'(q_{I,x})} - \sqrt{\Phi_s'(q_{I,x})(p \times \xi^s \circ \Phi)'(q_{I,x})} \\ &+ \sqrt{\tilde{\Phi}_s'(q_{J,x})(\tilde{p} \times \xi^s \circ \tilde{\Phi})'(q_{J,x})} - \sqrt{\Phi_s'(q_{J,x})(p \times \xi^s \circ \Phi)'(q_{J,x})}. \end{aligned} \tag{5.141}$$

Recalling again (5.138), we have

$$\begin{aligned} (\tilde{p} \times \xi^s \circ \tilde{\Phi})'(q_{I,x}) &= \tilde{p}(q_{I,x}) \sum_{s' \in \mathscr{S}} \partial_{x_{s'}} \xi^s(\tilde{\Phi}(q_{I,x})) \cdot \tilde{\Phi}_{s'}'(q_{I,x}) + \tilde{p}'(q_{I,x}) \cdot \xi^s(\tilde{\Phi}(q_{I,x})) \\ &= p(q_*) \sum_{s' \in \mathscr{S}} \partial_{x_{s'}} \xi^s(\Phi(q_*)) \cdot \tilde{\Phi}_{s'}'(q_{I,x}) + p'(q_{I,x}) \cdot \xi^s(\Phi(q_*)) \pm O(\varepsilon). \end{aligned} \tag{5.142}$$

Similarly to (5.139), we now control the first two terms of (5.141):

$$\begin{aligned} &\sqrt{\tilde{\Phi}_s'(q_{I,x})(\tilde{p} \times \xi^s \circ \tilde{\Phi})'(q_{I,x})} - \sqrt{\Phi_s'(q_{I,x})(p \times \xi^s \circ \Phi)'(q_{I,x})} + O(\varepsilon) \\ &\overset{(5.142)}{\ge} \sqrt{\tilde{\Phi}_s'(q_{I,x}) \cdot \left(p(q_*) \sum_{s' \in \mathscr{S}} \partial_{x_{s'}} \xi^s(\Phi(q_*)) \cdot \tilde{\Phi}_{s'}'(q_{I,x}) + \tilde{p}'(q_{I,x}) \cdot \xi^s(\Phi(q_*))\right)} \\ &\quad - \sqrt{\Phi_s'(q_{I,x}) \cdot \left(p(q_*) \sum_{s' \in \mathscr{S}} \partial_{x_{s'}} \xi^s(\Phi(q_*)) \cdot \Phi_{s'}'(q_{I,x}) + p'(q_{I,x}) \cdot \xi^s(\Phi(q_*))\right)} \end{aligned} \tag{5.143}$$

and analogously for J instead of I.

It remains to lower-bound the right hand side of (5.143). We break into cases depending on whether Φ' is continuous (if so, then p' must be discontinuous). In both cases, the idea is to argue that the concavity of the square root function yields an increase in the value of $\mathbb{A}$.

Case 1: Φ' is continuous at q_* In this case p' is discontinuous, and (5.136) applies with $f = p$. We estimate the right-hand side of (5.143): as $|\Phi_s'(q) - \tilde{\Phi}_s'(q')| \le o_{\varepsilon \to 0}(1)$ uniformly in $q, q' \in (q_* - \varepsilon, q_* + \varepsilon)$ by definition,

$$\begin{aligned}
&\sqrt{\tilde{\Phi}_s'(q_{I,x})\cdot\left(p(q_*)\sum_{s'\in\mathscr{S}}\partial_{x_{s'}}\xi^s(\Phi(q_*))\cdot\tilde{\Phi}_{s'}'(q_{I,x})+\tilde{p}'(q_{I,x})\cdot\xi^s(\Phi(q_*))\right)}\\
&\quad-\sqrt{\Phi_s'(q_{I,x})\cdot\left(p(q_*)\sum_{s'\in\mathscr{S}}\partial_{x_{s'}}\xi^s(\Phi(q_*))\cdot\Phi_{s'}'(q_{I,x})+p'(q_{I,x})\cdot\xi^s(\Phi(q_*))\right)}\\
&=\sqrt{\Phi_s'(q_*)\cdot\left(p(q_*)\sum_{s'\in\mathscr{S}}\partial_{x_{s'}}\xi^s(\Phi(q_*))\cdot\Phi_{s'}'(q_*)+\tilde{p}'(q_{I,x})\cdot\xi^s(\Phi(q_*))\right)}\\
&\quad-\sqrt{\Phi_s'(q_*)\cdot\left(p(q_*)\sum_{s'\in\mathscr{S}}\partial_{x_{s'}}\xi^s(\Phi(q_*))\cdot\Phi_{s'}'(q_*)+p'(q_{I,x})\cdot\xi^s(\Phi(q_*))\right)}\pm o_{\varepsilon\to 0}(1).
\end{aligned}\tag{5.144}$$

We analyze the last term, combined with the analogous expression for J, using the strict concavity in Lemma 5.C.9 of $x\mapsto\sqrt{x}$ together with (5.136) applied to p. We find that

$$\begin{aligned}
&\sqrt{\Phi_s'(q_*)\cdot\left(p(q_*)\sum_{s'\in\mathscr{S}}\partial_{x_{s'}}\xi^s(\Phi(q_*))\cdot\Phi_{s'}'(q_*)+\tilde{p}'(q_{I,x})\cdot\xi^s(\Phi(q_*))\right)}\\
&\quad-\sqrt{\Phi_s'(q_*)\cdot\left(p(q_*)\sum_{s'\in\mathscr{S}}\partial_{x_{s'}}\xi^s(\Phi(q_*))\cdot\Phi_{s'}'(q_*)+p'(q_{I,x})\cdot\xi^s(\Phi(q_*))\right)}\\
&\quad+\sqrt{\Phi_s'(q_*)\cdot\left(p(q_*)\sum_{s'\in\mathscr{S}}\partial_{x_{s'}}\xi^s(\Phi(q_*))\cdot\Phi_{s'}'(q_*)+\tilde{p}'(q_{J,x})\cdot\xi^s(\Phi(q_*))\right)}\\
&\quad-\sqrt{\Phi_s'(q_*)\cdot\left(p(q_*)\sum_{s'\in\mathscr{S}}\partial_{x_{s'}}\xi^s(\Phi(q_*))\cdot\Phi_{s'}'(q_*)+p'(q_{J,x})\cdot\xi^s(\Phi(q_*))\right)}\\
&\geq c(\eta).
\end{aligned}\tag{5.145}$$

Indeed, all quantities except $p'(\cdot)$ and $\tilde{p}'(\cdot)$ are the same in the four expressions and are bounded away from 0 and infinity. Furthermore all other expressions differ by $O(\varepsilon_1)$ thanks to (5.138), which is small compared to the discrepancy η between the values of p' and $\widetilde{p}'$. Hence for η fixed and ε small enough, they are bounded away from the equality cases of Lemma 5.C.9.

Combining (5.143), (5.144), and (5.145) implies that for each $x\in[0,\varepsilon_1]$ and small enough ε,

$$\begin{aligned}
&\sqrt{\tilde{\Phi}_s'(q_{I,x})(\tilde{p}\times\xi^s\circ\tilde{\Phi})'(q_{I,x})}-\sqrt{\Phi_s'(q_{I,x})(p\times\xi^s\circ\Phi)'(q_{I,x})}\\
&+\sqrt{\tilde{\Phi}_s'(q_{J,x})(\tilde{p}\times\xi^s\circ\tilde{\Phi})'(q_{J,x})}-\sqrt{\Phi_s'(q_{J,x})(p\times\xi^s\circ\Phi)'(q_{J,x})}\\
&\geq c(\eta)-o_{\varepsilon\to 0}(1)\\
&\geq c(\eta)/2.
\end{aligned}$$

Integrating over $x\in[0,\varepsilon_1]$ and combining with (5.140), we conclude that (5.137) holds in Case 1.

Case 2: Φ' is discontinuous at q_*. (Note that p' might also be discontinuous.)

Define for each $s\in\mathscr{S}$ the function

$$F_s(A_1,\dots,A_r,B)=\sqrt{A_s\cdot\left(p(q_*)\sum_{s'\in\mathscr{S}}\partial_{x_{s'}}\xi^s(\Phi(q_*))A_{s'}+B\xi^s(\Phi(q_*))\right)}.$$

Lemma 5.C.9 implies that each function F_s is concave on $\mathbb{R}_{\geq 0}^{r+1}$, since both A_s and $p(q_*)\sum_{s'\in\mathscr{S}}\partial_{x_{s'}}\xi^s(\Phi(q_*))A_{s'}+B\xi^s(\Phi(q_*))$ are linear functions of $(A_1,\dots,A_r,B)$. In particular, for each $(s,x)\in\mathscr{S}\times[0,\varepsilon_1]$ the function

$$f_{s,x}(t)\equiv F_s\left(\frac{(1-t)\Phi_1'(q_{I,x})+t\Phi_1'(q_{J,x})}{2},\dots,\frac{(1-t)\Phi_r'(q_{I,x})+t\Phi_r'(q_{J,x})}{2},\frac{(1-t)p'(q_{I,x})+tp'(q_{J,x})}{2}\right)$$

is concave for $t\in[0,1]$. Recalling the definitions of $\tilde{p}$ and $\tilde{\Phi}$, we expand the inequality $2f_{s,x}(1/2)\geq f_{s,x}(0)+f_{s,x}(1)$ to obtain

$$\begin{aligned}&\sqrt{\tilde{\Phi}_s'(q_{I,x})\cdot\left(p(q_*)\sum_{s'\in\mathscr{S}}\partial_{x_{s'}}\xi^s(\Phi(q_*))\cdot\tilde{\Phi}_{s'}'(q_{I,x})+\tilde{p}'(q_{I,x})\cdot\xi^s(\Phi(q_*))\right)}\\&\quad-\sqrt{\Phi_s'(q_{I,x})\cdot\left(p(q_*)\sum_{s'\in\mathscr{S}}\partial_{x_{s'}}\xi^s(\Phi(q_*))\cdot\Phi_{s'}'(q_{I,x})+p'(q_{I,x})\cdot\xi^s(\Phi(q_*))\right)}\\&\quad+\sqrt{\tilde{\Phi}_s'(q_{J,x})\cdot\left(p(q_*)\sum_{s'\in\mathscr{S}}\partial_{x_{s'}}\xi^s(\Phi(q_*))\cdot\tilde{\Phi}_{s'}'(q_{J,x})+\tilde{p}'(q_{J,x})\cdot\xi^s(\Phi(q_*))\right)}\\&\quad-\sqrt{\Phi_s'(q_{J,x})\cdot\left(p(q_*)\sum_{s'\in\mathscr{S}}\partial_{x_{s'}}\xi^s(\Phi(q_*))\cdot\Phi_{s'}'(q_{J,x})+p'(q_{J,x})\cdot\xi^s(\Phi(q_*))\right)}\\&\geq 0.\end{aligned}\tag{5.146}$$

In light of (5.143), this means that perturbing $(p,\Phi)\to(\tilde{p},\tilde{\Phi})$ can only hurt the contribution from a given $s\in\mathscr{S}$ by $O(\varepsilon)$. To complete the proof we will show that the contribution from some $s\in\mathscr{S}$ is positive and of a larger order. Which of these must occur will depend on the ratio $\frac{p'(q_{I,x})}{p'(q_{J,x})}$.

We will get this contribution from either $s_{\max}$ or $s_{\min}$, defined now. For each $x\in[0,\varepsilon_1]$, let

$$\begin{aligned}s_{\max}(x)&=\arg\max_{s\in\mathscr{S}}\frac{\Phi_s'(q_{I,x})}{\Phi_s'(q_{J,x})},\\ s_{\min}(x)&=\arg\min_{s\in\mathscr{S}}\frac{\Phi_s'(q_{I,x})}{\Phi_s'(q_{J,x})}.\end{aligned}$$

(Both are defined up to almost everywhere equivalence if ties are broken lexicographically.) Recall the functions $\Phi_s'(x)$ are uniformly bounded above and below. It follows from (5.136) that

$$\frac{\Phi_{s_{\min}}'(q_{I,x})}{\Phi_{s_{\min}}'(q_{J,x})}\leq 1-\eta'\leq 1+\eta'\leq\frac{\Phi_{s_{\max}}'(q_{I,x})}{\Phi_{s_{\max}}'(q_{J,x})}\tag{5.147}$$

for some η' depending only on (η,ξ,h). (Discontinuity of Φ' gives one side, and admissibility forces another $s\in\mathscr{S}$ to change in the opposite direction.)

Without loss of generality, suppose that

$$\frac{p'(q_{I,x})}{p'(q_{J,x})}\leq 1.\tag{5.148}$$

In this case, the assumption (5.148) implies

$$\frac{p(q_*)\sum_{s'\in\mathscr{S}}\partial_{x_{s'}}\xi^{s_{\max}}(\Phi(q_*))\cdot\Phi_{s'}'(q_{I,x})+p'(q_{I,x})\cdot\xi^{s_{\max}}(\Phi(q_*))}{p(q_*)\sum_{s'\in\mathscr{S}}\partial_{x_{s'}}\xi^{s_{\max}}(\Phi(q_*))\cdot\Phi_{s'}'(q_{J,x})+p'(q_{J,x})\cdot\xi^{s_{\max}}(\Phi(q_*))}\leq\frac{\Phi_{s_{\max}}'(q_{I,x})}{\Phi_{s_{\max}}'(q_{J,x})}-\eta_1$$

for a constant $\eta_1>0$ depending only on (η,q_*,ξ,h). Since all quantities are bounded away from 0 and infinity, applying a simple compactness argument to the equality case in Lemma 5.C.9 implies

$$2f_{s_{\max},x}(1/2)\geq f_{s_{\max},x}(0)+f_{s_{\max},x}(1)+c(\eta_1).\tag{5.149}$$

Similarly if (5.148) does not hold, then we find (5.149) with $s_{\min}$ in place of $s_{\max}$.

Combining the above with $\varepsilon \ll \eta$, we find that for each $x \in [0, \varepsilon_1]$,

$$\sum_{s\in\mathscr{S}} 2f_{s,x}(1/2) \geq \sum_{s\in\mathscr{S}} \Big(f_{s,x}(0) + f_{s,x}(1)\Big) + c(\eta_1)/2.$$

Integrating over x and recalling (5.140) and (5.143), we conclude that (5.137) also holds in Case 2. This completes the proof. $\square$

Proof of Proposition 5.4.11. Follows from Lemmas 5.C.4, 5.C.7, and 5.C.10. The upper bound on Φ' comes from admissibility (5.5), which implies that $\Phi'_s \leq \lambda_s^{-1}$. $\square$

5.C.3 Type $\mathrm{I}I$ solutions

Here we show that the type $\mathrm{I}I$ equation implicitly takes the form of a second order ordinary differential equation in which $\Phi''(q)$ is Lipschitz in $(\Phi(q), \Phi'(q))$. It follows that a unique type $\mathrm{I}I$ solution exists given any first-order initial condition $(\Phi(q_1), \Phi'(q_1))$, and that the type $\mathrm{I}I$ ODE is satisfied at *all* points in $(q_1, 1)$. We will often enforce the admissibility conditions

$$\langle \vec{\lambda}, \vec{\Phi}'(q)\rangle = 1, \tag{5.150}$$

$$\langle \vec{\lambda}, \vec{\Phi}''(q)\rangle = 0. \tag{5.151}$$

In particular, we denote by $A_{\geq 0}$ the set of vectors $v \in \mathbb{R}_{\geq 0}^{\mathscr{S}}$ satisfying $\langle \vec{\lambda}, v\rangle = 1$. The following important but rather lengthy Lemma 5.4.38 ensures that type $\mathrm{I}I$ solutions are described by a Lipschitz ODE. In it, the value q is actually irrelevant and just serves as a placeholder. Importantly there is no issue when $\Phi_s(q)$ or $\Phi'_s(q)$ is near zero, thanks to non-degeneracy.

Lemma 5.4.38. *Fix $\varepsilon > 0$. For $\Phi(q) \in \mathbb{R}_{\geq 0}^{\mathscr{S}}$ and $\Phi'(q) \in A_{\geq 0}(q)$, the* ***type*** $\mathrm{I}I$ ***equation***

$$\Psi_s(q) = \Psi_{s'}(q) \quad \forall s, s' \in \mathscr{S};$$
$$\langle \vec{\lambda}, \Phi''(q)\rangle = 0$$

is equivalent (for each fixed q) to

$$\Phi''(q) = F(\Phi(q), \Phi'(q))$$

for a locally Lipschitz function $F : \mathbb{R}_{\geq 0}^{\mathscr{S}} \times A_{\geq 0}^{\mathscr{S}} \to \mathbb{R}^{\mathscr{S}}$. Moreover,

$$|\Phi''_s(q)| \leq O(|\Phi'_s(q)|), \quad \forall s \in \mathscr{S}.$$

with a uniform constant for bounded $\Phi'(q)$.

Proof. Write $\Psi(q)$ for $\Psi_s(q)$, which is independent of $s \in \mathscr{S}$ by assumption. We assume throughout that $\Phi(q)$ lies in a bounded set in writing $O(\cdot)$ and $\Omega(\cdot)$ expressions. Note that $\vec{\Phi}''(q)$ exists as an L^1 function for $q \in (q_1, 1]$ since $\vec{\Phi}'$ is absolutely continuous. We write

$$\begin{aligned}
2\Psi(q) &= \frac{2}{\Phi'_s(q)}\frac{\mathrm{d}}{\mathrm{d}q}\sqrt{\frac{\Phi'_s(q)}{(\xi^s\circ\Phi)'(q)}} \\
&= \sqrt{\frac{(\xi^s\circ\Phi)'(q)}{\Phi'_s(q)^3}}\frac{\mathrm{d}}{\mathrm{d}q}\frac{\Phi'_s(q)}{(\xi^s\circ\Phi)'(q)} \\
&= \sqrt{\frac{(\xi^s\circ\Phi)'(q)}{\Phi'_s(q)^3}} \cdot \frac{\Phi''_s(q)(\xi^s\circ\Phi)'(q) - \Phi'_s(q)(\xi^s\circ\Phi)''(q)}{(\xi^s\circ\Phi)'(q)^2} \\
&= \frac{1}{\sqrt{\Phi'_s(q)^3(\xi^s\circ\Phi)'(q)^3}} \cdot \left(\Phi''_s(q)(\xi^s\circ\Phi)'(q) - \Phi'_s(q)(\xi^s\circ\Phi)''(q)\right).
\end{aligned} \tag{5.152}$$

Moreover we have

$$\begin{aligned}
&\Phi_s''(q)(\xi^s\circ\Phi)'(q)-\Phi_s'(q)(\xi^s\circ\Phi)''(q)\\
&=\Phi_s''(q)\sum_{s'\in\mathscr{S}}\partial_{x_{s'}}\xi^s(\Phi(q))\cdot\Phi_{s'}'(q)\\
&\quad-\Phi_s'(q)\left(\sum_{s'\in\mathscr{S}}\partial_{x_{s'}}\xi^s(\Phi(q))\cdot\Phi_{s'}''(q)+\sum_{s',s''\in\mathscr{S}}\partial_{x_{s'}}\partial_{x_{s''}}\xi^s(\Phi(q))\cdot\Phi_{s'}'(q)\right)
\end{aligned}$$

Let

$$B_s(q)=\sum_{s'\in\mathscr{S}}\partial_{x_{s'}}\xi^s(\Phi(q))\cdot\Phi_{s'}'(q).$$

Note that by non-degeneracy each $\partial_{x_{s'}}\xi^s(\Phi(q))$ is bounded away from 0 and ∞ for all $\Phi(q)\in[0,1]^{\mathscr{S}}$. Meanwhile $\sum_{s\in\mathscr{S}}\lambda_s\Phi_s'(q)=1$. Thus for $\Phi'(q)$ obeying (5.150), each $B_s(q)$ is uniformly bounded away from 0 and ∞.

Next let $M(q)\in\mathbb{R}^{\mathscr{S}\times\mathscr{S}}$ be a square matrix with entries

$$M(q)_{s,s'}=\frac{\Phi_s'(q)\cdot\partial_{x_{s'}}\xi^s(\Phi(q))}{B_s(q)}$$

and let I denote the identity $\mathscr{S}\times\mathscr{S}$ matrix. Then the above equations for all $s\in\mathscr{S}$ can be expressed more succinctly as

$$(M-I)\Phi''(q)=-w_1(\Phi(q),\Phi'(q))-\Psi(q)\cdot w_2(\Phi(q),\Phi'(q)) \tag{5.153}$$

for Lipschitz functions $w_1,w_2:[0,1]^{2r}\to\mathbb{R}^r_{>0}$ given explicitly by

$$\begin{aligned}
(w_1)_s&=\frac{\Phi_s'(q)\sum_{s',s''\in\mathscr{S}}\partial_{x_{s'}}\partial_{x_{s''}}\xi^s(\Phi(q))\cdot\Phi_{s'}'(q)}{B_s(q)};\\
(w_2)_s&=\frac{2\sqrt{\Phi_s'(q)^3(\xi^s\circ\Phi)'(q)^3}}{B_s(q)}.
\end{aligned} \tag{5.154}$$

Since B_s is bounded below, both w_1 and w_2 have uniformly bounded entries. Moreover B and w_1,w_2 are uniformly Lipschitz in $(\Phi(q),\Phi'(q))$. Note also that w_2 is entry-wise non-negative.

As a first observation, observe that

$$(M-I)\Phi'(q)=0.$$

Because $\Phi'(q)\succeq 0$ and M has positive entries, this means $\Phi'(q)$ is the unique right Perron-Frobenius eigenvector of M, and thus $\mathrm{rank}(M-I)=r-1$. It follows that for given $(\Phi(q),\Phi'(q))$, a unique solution $(\Phi''(q),\Psi(q))$ to (5.153) exists so long as

$$w_2\notin\mathrm{range}(M-I). \tag{5.155}$$

In fact (5.155) is always true. To see this, note that M has a left Perron-Frobenius eigenvector $v\in\mathbb{R}^r_{>0}$ with $v(M-I)=0$. Then if $w_2=(M-I)w$ for $w\in\mathbb{R}^{\mathscr{S}}$, we find $\langle v,w_2\rangle=0$. This is a contradiction: $\langle v,w_2\rangle>0$ since all entries are strictly positive in both vectors. We denote by $\Lambda(q)\in\mathbb{R}^{\mathscr{S}}$ the value of $\Phi''(q)$ in the aforementioned unique solution.

Our primary aim is now to show that $\Lambda(q)$ is a Lipschitz function of $(\Phi(q),\Phi'(q))\in\mathbb{R}^{\mathscr{S}}\times A_{\geq 0}$. We would like to apply Perron-Frobenius arguments to M, but the fact that $M_{s,s'}\asymp\Phi_s'(q)$ may be very small poses an issue. To rectify this, we define $\widetilde{M}(q)$ with entries

$$\widetilde{M}(q)_{s,s'}=\frac{\Phi_{s'}'(q)\partial_{x_{s'}}\xi^s(\Phi(q))}{B_s(q)}. \tag{5.156}$$

Then defining the diagonal $\mathscr{S}\times\mathscr{S}$ matrix $D(\Phi'(q))$ with entries

$$D(\Phi'(q))_{s,s}=\Phi_s'(q)$$

we have

$$\widetilde{M}(q) = D(\Phi'(q))^{-1}\, M\, D(\Phi'(q)).$$

The key property obeyed by $\widetilde{M}$ but not M is that for any $v \in \mathbb{R}^{\mathscr{S}}_{>0}$, the entries of $\widetilde{M}v$ are of the same order. Namely, all ratios $\frac{(\widetilde{M}v)_s}{(\widetilde{M}v)_{s'}}$ are uniformly bounded because the ratios $M_{s,s'}/M_{s'',s'}$ are uniformly bounded. In particular Lemma 5.C.12 and hence Lemma 5.C.11 (see below) apply to $\widetilde{M}$.

Note that $\widetilde{M}$ has Perron-Frobenius eigenvector $\vec{1}$ and $\widetilde{M}$ is Lipschitz in $(\Phi(q), \Phi'(q))$. We set

$$\begin{aligned} \widetilde{V}(q) &= D(\Phi'(q))^{-1}\Lambda(q), \quad \text{i.e. } \widetilde{V}(q)_s = \frac{\Lambda_s(q)}{\Phi'_s(q)}; \\ V(q)_s &= \widetilde{V}(q)_s - \frac{\sum_{s'\in\mathscr{S}} \widetilde{V}(q)_{s'}}{r}. \end{aligned} \tag{5.157}$$

By construction, $\sum_s V(q)_s = 0$. Moreover

$$(\widetilde{M}(q) - I)V(q) = (\widetilde{M}(q) - I)\widetilde{V}(q) \tag{5.158}$$

since $V(q) - \widetilde{V}(q)$ is proportional to $\vec{1}$.

A priori estimate on $\Lambda(q)$ We now prove (5.88), which will also serve as a useful intermediate step. Note first that w_1 satisfies $|(w_1)_s| = O(\Phi'_s(q))$ (recall that B_s is bounded below), while all entries of w_2 are non-negative. Therefore the entries of $w_1(q) + \Psi(q)w_2(q)$ are bounded either above or below by $O(\Phi'_s(q))$. Furthermore by definition,

$$\begin{aligned} -w_1(q) - \Psi(q)w_2(q) &= (M(q) - I)\Lambda(q) \\ &= D(\Phi'(q))(\widetilde{M}(q) - I)\widetilde{V} \\ &\overset{(5.158)}{=} D(\Phi'(q))(\widetilde{M}(q) - I)V. \end{aligned}$$

We conclude that

$$\min\Big(\|((\widetilde{M}(q) - I)V)_+\|_1, \|((\widetilde{M}(q) - I)V)_-\|_1\Big) \le O(1).$$

Lemma 5.C.11 below now implies that

$$\|V(q)\|_1 \le O(1). \tag{5.159}$$

Note that $\langle \widetilde{V}, \lambda \odot \Phi'(q)\rangle = 0$ by (5.151) and (5.157). The second part of the latter also implies $V(q) - \widetilde{V}(q)$ is proportional to $\vec{1}$, and so

$$\begin{aligned} \left|(V(q) - \widetilde{V}(q))_s\right| &= \left|\langle V(q) - \widetilde{V}(q), \lambda \odot \Phi'(q)\rangle\right| \\ &= |\langle V(q), \lambda \odot \Phi'(q)\rangle| \\ &\overset{(5.159)}{\le} O(1) \end{aligned} \tag{5.160}$$

Using again (5.159) and (5.157) we find that $\|\widetilde{V}(q)\|_1 \le O(1)$ as well. Finally since $\Lambda(q) = \widetilde{V}(q) \odot \Phi'(q)$, we get (5.88) as desired.

Controlling Ψ We take a second detour to show that $\Psi(q)$ is bounded and Lipschitz. Using that $\|w_1\|_1 \le O(1)$ and $\|w_2\|_1 \ge \Omega(1)$ in the first step below, we find

$$\begin{aligned} \Omega(|\Psi(q)|) - O(1) &\le \|w_1(q) + \Psi(q)w_2(q)\|_1 \\ &= \|(M(q) - I)\Lambda(q)\|_1 \\ &\overset{(5.153)}{\le} O(1). \end{aligned}$$

The just-proved estimate (5.88) implies the weaker bound $\|\Lambda(q)\|_1 \le O(1)$, which was used in the last step.

We conclude that $\Psi(q)$ is uniformly bounded:

$$|\Psi(q)| \le O(1). \tag{5.161}$$

Next we show that $\Psi(q)$ is Lipschitz in $(\Phi(q), \Phi'(q))$. We begin by writing

$$\begin{aligned}(M(q)-I)\Lambda(q) - (M(q')-I)\Lambda(q') &= w_1(q') - w_1(q) + \Psi(q')w_2(q') - \Psi(q)w_2(q)\\ &= w_1(q') - w_1(q) + \Psi(q')\big(w_2(q') - w_2(q)\big) + \big(\Psi(q') - \Psi(q)\big)w_2(q)\\ &= O\big(\|\Phi(q) - \Phi(q')\| + \|\Phi'(q) - \Phi'(q')\|\big) + \big(\Psi(q') - \Psi(q)\big)w_2(q).\end{aligned}$$

(Note that the latter $O(\cdot)$ notation hides a vector in $\mathbb{R}^r$.) We will rely on the fact that $w_2(q)$ is entrywise positive and $\|w_2(q)\| \ge \Omega(1)$. To analyze the left-hand side above, we write

$$\begin{aligned}(M(q)-I)\Lambda(q) - (M(q')-I)\Lambda(q') &= (M(q) - M(q'))\Lambda(q) + (M(q')-I)\big(\Lambda(q) - \Lambda(q')\big)\\ &\le O\big(\|\Phi(q) - \Phi(q')\| + \|\Phi'(q) - \Phi'(q')\|\big) + (M(q')-I)\big(\Lambda(q) - \Lambda(q')\big).\end{aligned}$$

The latter step holds since $M(q)$ is Lipschitz in $(\Phi(q), \Phi(q'))$ and $\|\Lambda(q)\|_1 \le O(1)$ from (5.88). Now, let v be the left Perron-Frobenius eigenvector of $M(q')$, so $v(M(q')-I) = 0$, normalized so that $v \succeq 0$ and $\|v\|_1 = 1$. Combining the previous displays implies that

$$(\Psi(q') - \Psi(q)) \cdot \langle v, w_2(q)\rangle = O\big(\|\Phi(q) - \Phi(q')\| + \|\Phi'(q) - \Phi'(q')\|\big).$$

Finally we show that $\langle v, w_2(q)\rangle$ is bounded away from 0. Indeed both vectors are entrywise positive, and $\|w_2(q)\|_1 \ge \Omega(1)$ while $\min_s v_s \ge \Omega(1)$. The latter statement holds for similar reasons to the right eigenvector properties of $\widetilde{M}$ explained above: for *any* $v \in \mathbb{R}^{\mathscr{S}}_{>0}$, the ratios $\frac{(vM)_s}{(vM)_{s'}}$ are uniformly bounded, and this ratio is simply $v_s/v_{s'}$ when v is the left Perron-Frobenius eigenvector. We conclude that

$$|\Psi(q) - \Psi(q')| \le O\big(\|\Phi(q) - \Phi(q')\| + \|\Phi'(q) - \Phi'(q')\|\big) \tag{5.162}$$

which ends this second detour.

Finishing the proof Having established (5.88) and (5.161), we return to showing that $\Lambda(q)$ is Lipschitz in $(\Phi(q), \Phi'(q))$. Fix a different pair

$$(\Phi(q'), \Phi'(q')) \neq (\Phi(q), \Phi'(q)).$$

Accordingly define $w_1(q'), w_2(q'), M(q'), V(q')$ and so on using $(\Phi(q'), \Phi'(q'))$. (Since we don't require admissibility but only its differential version (5.150), there is no loss of generality here; q' like q is just a place-holder variable so e.g. $\Phi(q) = \Phi(q')$ is possible.)

Then Lemma 5.C.11 implies:

$$\|(\widetilde{M}(q) - I)V(q) - (\widetilde{M}(q) - I)V(q')\|_1 \ge \Omega\big(\|V(q) - V(q')\|_1\big). \tag{5.163}$$

Using the reverse triangle inequality in the first step, we find the lower bound

$$\begin{aligned}\|(\widetilde{M}(q) - I)V(q) - (\widetilde{M}(q') - I)V(q')\|_1 &\ge \|(\widetilde{M}(q) - I)V(q) - (\widetilde{M}(q) - I)V(q')\|_1\\ &\quad - \|(\widetilde{M}(q) - I)V(q') - (\widetilde{M}(q') - I)V(q')\|_1\\ &\overset{(5.163)}{\ge} \Omega\big(\|V(q) - V(q')\|_1\big) - O\big(\|\widetilde{M}(q) - \widetilde{M}(q')\|_1\big)\\ &\ge \Omega\big(\|V(q) - V(q')\|_1\big) - O\big(\|\Phi(q) - \Phi(q')\|_1 + \|\Phi'(q) - \Phi'(q')\|_1\big).\end{aligned}$$

By (5.153), (5.154), (5.161) (5.162), and the simple estimate $\max\big(|w_1(q)_s|, |w_2(q)_s|\big) \leq O(\Phi'_s(q))$, the left-hand side above is upper bounded by

$$
\begin{aligned}
&\|(\widetilde{M}(q) - I)V(q) - (\widetilde{M}(q') - I)V(q')\|_1 \\
&= \|(\widetilde{M}(q) - I)\widetilde{V}(q) - (\widetilde{M}(q') - I)\widetilde{V}(q')\|_1 \\
&= \Big\| D(\Phi'(q))^{-1}\Big((M(q) - I)\Lambda(q)\Big) - D(\Phi'(q'))^{-1}\Big((M(q') - I)\Lambda(q')\Big)\Big\|_1 \\
&= \Big\| D(\Phi'(q))^{-1}\Big(w_1(q) + \Psi(q)w_2(q)\Big) - D(\Phi'(q'))^{-1}\Big(w_1(q') + \Psi(q')w_2(q')\Big)\Big\|_1 \\
&\leq O\big(\|\Phi(q) - \Phi(q')\|_1 + \|\Phi'(q) - \Phi'(q')\|_1\big).
\end{aligned}
$$

Rearranging the previous two displays implies that

$$
\|V(q) - V(q')\|_1 \leq O\big(\|\Phi(q) - \Phi(q')\|_1 + \|\Phi'(q) - \Phi'(q')\|_1\big). \tag{5.164}
$$

It remains to unwind the transformations to conclude the same for Λ. Mimicking (5.160) in the first step,

$$
\begin{aligned}
\Big|\big(V_s(q) - V_s(q')\big) - \big(\widetilde{V}_s(q) - \widetilde{V}_s(q')\big)\Big| &= \Big|\sum_s \lambda_s\big(\Phi'_s(q)V(q) - \Phi'_s(q')V(q')\big)\Big| \\
&\leq O\big(\|\Phi'(q)\| \cdot \|V(q) - V(q')\|\big) + O\big(\|\Phi'(q) - \Phi'(q')\| \cdot \|V(q')\|\big) \\
&\overset{(5.164),(5.159)}{\leq} O\big(\|\Phi(q) - \Phi(q')\|_1 + \|\Phi'(q) - \Phi'(q')\|_1\big) \\
&\quad + O\big(\|\Phi'(q) - \Phi'(q')\big).
\end{aligned}
$$

Combining the previous two displays, we conclude that

$$
\begin{aligned}
\|\widetilde{V}(q) - \widetilde{V}(q')\|_1 &\leq \|V(q) - V(q')\|_1 + \|(V(q) - V(q')) - (\widetilde{V}(q) - \widetilde{V}(q'))\| \\
&\leq O\big(\|\Phi(q) - \Phi(q')\|_1 + \|\Phi'(q) - \Phi'(q')\|_1\big).
\end{aligned}
$$

Finally since $\Lambda(q) = \widetilde{V}(q) \odot \Phi'(q)$ and $\|\widetilde{V}(q')\|_1, \|\Phi'(q)\|_1 \leq O(1)$, we obtain the desired:

$$
\begin{aligned}
\|\Lambda(q) - \Lambda(q')\|_1 &\leq O\big(\|\widetilde{V}(q) - \widetilde{V}(q')\|_1 \cdot \|\Phi'(q)\|_1\big) + O\big(\|\widetilde{V}(q')\|_1 \cdot \|\Phi'(q) - \Phi'(q')\|_1\big) \\
&\leq O\big(\|\widetilde{V}(q) - \widetilde{V}(q')\|_1 + \|\Phi'(q) - \Phi'(q')\|_1\big) \\
&\leq O\big(\|\Phi(q) - \Phi(q')\|_1 + \|\Phi'(q) - \Phi'(q')\|_1\big).
\end{aligned}
$$

This concludes the proof. □

Lemma 5.C.11. *Let $\mathcal{M} \subseteq \mathbb{R}_{\geq 0}^{\mathscr{S} \times \mathscr{S}}$ be a compact set of entry-wise non-negative matrices with unique Perron-Frobenius eigenvector $\vec{1}$ and associated eigenvalue 1.*

Then for all $v \in \mathbb{R}^{\mathscr{S}}$ with $\sum_{s \in \mathscr{S}} v_s = 0$, we have

$$
\begin{aligned}
\|((M - I)v)_+\|_1 &\geq \Omega_{\mathcal{M},r}(\|v\|_1), \\
\|((M - I)v)_-\|_1 &\geq \Omega_{\mathcal{M},r}(\|v\|_1).
\end{aligned}
$$

Proof. The two statements are equivalent under negation so we assume the first is false and derive a contradiction. If it is false, by taking a convergent sequence of approximate counterexamples $(M^i, v^i) \to (\widehat{M}, \widehat{v})$ with $M^i \in \mathcal{M}$ and $\|v^i\|_1 = 1$, we have:

1. $\widehat{M} \in \mathcal{M}$.
2. $\widehat{M}$ has Perron-Frobenius eigenvector $\vec{1}$ and eigenvalue 1.
3. $\sum_{s \in \mathscr{S}} \widehat{v}_s = 0$.
4. $\|\widehat{v}\|_1 = 1$.

5. $\widehat{M}\widehat{v} \preceq \widehat{v}$ (since $((\widehat{M}-I)\widehat{v})_+ = 0$).

Since $\widehat{M}$ has simple Perron-Frobenius eigenvalue 1, for $\widehat{M}\widehat{v} \preceq \widehat{v}$ to hold we must actually have $\widehat{M}\widehat{v} = \widehat{v}$. Therefore $\widehat{v} = \vec{1}/r$ is a multiple of the right Perron-Frobenius eigenvector, contradicting $\sum_{s\in\mathscr{S}} \widehat{v}_s = 0$. □

Lemma 5.C.12. *For $C > 0$, let $\mathcal{M}_C \subseteq \mathbb{R}_{\geq 0}^{\mathscr{S}\times\mathscr{S}}$ consist of all matrices M such that:*

1. *$M_{s,s'} \in [0, C]$ for all $s, s' \in \mathscr{S}$.*
2. *$M_{s,s'} \leq CM_{s'',s'}$ for all $s, s', s'' \in \mathscr{S}$.*
3. *$M\vec{1} = \vec{1}$.*
4. *$\sum_{s,s'\in\mathscr{S}} M_{s,s'} \geq 1/C$.*

Then $\mathcal{M} = \mathcal{M}_C$ satisfies the conditions of Lemma 5.C.11.

Proof. The only thing to show is that $\vec{1}$ is the **unique** right Perron-Frobenius eigenvector associated to the eigenvalue 1 of any $M \in \mathcal{M}_C$, even though M may include zero entries. Thus, suppose that $w \in \mathbb{R}^{\mathscr{S}}$ satisfies $Mw = w$; we will show that w has all equal entries. Let $S' \subseteq \mathscr{S}$ be the non-empty set of s' such that $M_{s,s'} > 0$ (which does not depend on s by definition of $\mathcal{M}_C$). Then letting M' and w' be the $S' \times S'$ and S'-dimensional restrictions of M and w, we have $M'w' = w'$. Since M' has strictly positive entries, we conclude that w' has all entries proportional. Hence for some $a \geq 0$, we have $w_s = a$ for all $s \in S'$. By definition of S' we obtain $w = Mw = Ma^{\mathscr{S}} = a^{\mathscr{S}}$. This concludes the proof. □

Chapter 6

Sampling from spherical spin glasses in total variation via algorithmic stochastic localization

Abstract – We consider the problem of algorithmically sampling from the Gibbs measure of a mixed p-spin spherical spin glass. We give a polynomial-time algorithm that samples from the Gibbs measure up to vanishing total variation error, for any model whose mixture satisfies

$$\xi''(s) < \frac{1}{(1-s)^2}, \qquad \forall s \in [0,1).$$

Our algorithm follows the algorithmic stochastic localization (SL) approach introduced in [AMS22] and in fact achieves a more difficult task. Namely, it (approximately) generates a sample path $(\boldsymbol{y}_t)_{t\geq 0}$ from the SL process. For large T, $\boldsymbol{y}_T/T$ is an approximate sample from the target Gibbs measure, and hence we solve the original problem as a special case.

We provide evidence that the condition $\xi''(s) < (1-s)^{-2}$ is sharp for the task of sampling SL sample paths. Furthermore, we prove that it is optimal within a class of generalized SL processes. Remarkably this threshold is close to the conjectured threshold for sampling from the original Gibbs measure. In particular, for the pure p-spin glasses it is within an absolute (p-independent) constant of the so-called shattering phase transition. Earlier work on related models was suboptimal by a factor diverging polynomially in p.

A key step of this approach is to estimate the mean of a sequence of tilted measures. We produce an improved estimator for this task by identifying a suitable correction to the TAP fixed point selected by approximate message passing (AMP). As a consequence, we improve the algorithm's guarantee over previous work, from normalized Wasserstein to total variation error. In particular, the new algorithm and analysis opens the way to perform inference about one-dimensional projections of the measure.

6.1 Introduction

Let $\gamma_2, \gamma_3, \ldots \geq 0$ satisfy $\sum_{p\geq 2} 2^p \gamma_p^2 < \infty$. The mixed p-spin glass Hamiltonian $H_N : \mathbb{R}^N \to \mathbb{R}$ is

$$H_N(\boldsymbol{\sigma}) = \sum_{p\geq 2} \frac{\gamma_p}{N^{(p-1)/2}} \sum_{i_1,\ldots,i_p=1}^{N} G_{i_1,\ldots,i_p} \sigma_{i_1} \cdots \sigma_{i_p}, \qquad G_{i_1,\ldots,i_p} \stackrel{i.i.d.}{\sim} \mathcal{N}(0,1). \tag{6.1}$$

Define the mixture function $\xi(s) = \sum_{p\geq 2} \gamma_p^2 s^p$, so that H_N is the Gaussian process with covariance

$$\mathbb{E}\, H_N(\boldsymbol{\sigma}^1) H_N(\boldsymbol{\sigma}^2) = N\xi\big(\langle \boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2 \rangle / N\big)\,.$$

The Gibbs measure of this model is the probability measure over the sphere $S_N = \{\boldsymbol{x} \in \mathbb{R}^N : \|\boldsymbol{x}\|_2^2 = N\}$ given by

$$\mu_{H_N}(\mathsf{d}\boldsymbol{\sigma}) = \frac{1}{Z_N} \exp(H_N(\boldsymbol{\sigma}))\ \mu_0(\mathsf{d}\boldsymbol{\sigma}), \qquad Z_N = \int_{S_N} \exp(H_N(\boldsymbol{\sigma}))\ \mu_0(\mathsf{d}\boldsymbol{\sigma})\,. \tag{6.2}$$

Here and below, μ_0 denotes the uniform probability measure on S_N. We will denote by $\boldsymbol{G} = (G_{i_1,\dots,i_p})_{p\geq 2, i_\ell \leq N}$ the vector of couplings that defines the Hamiltonian. We consider the problem of efficiently sampling from this Gibbs measure. For dist a distance on $\mathcal{P}(\mathbb{R}^N)$ (the set of probability measures over $\mathbb{R}^N$), we seek a computationally efficient algorithm that generates $\boldsymbol{\sigma}^{\mathsf{alg}}$ whose law μ^{alg} satisfies $\mathsf{dist}(\mu^{\mathsf{alg}}, \mu_{H_N}) = o_N(1)$, with high probability over H_N.

The most classical approach to sampling is to constructs a Markov chain that is reversible with respect to $\mu_{H_N}(\mathsf{d}\boldsymbol{\sigma})$, and for which a single step can be implemented efficiently. In the present context, such a Markov chain can be obtain by discretizing Langevin dynamics, see [Dal17, DM17, DCWY19] and Section 6.2.2.

An alternative approach has emerged recently. The basic idea is to generate (an approximation of) a sample path from the following Ito diffusion on $\mathbb{R}^N$:

$$\mathsf{d}\boldsymbol{y}_t = \boldsymbol{m}(\boldsymbol{y}_t, t)\,\mathsf{d}t + \mathsf{d}\boldsymbol{B}_t\,, \qquad \boldsymbol{y}_0 = \boldsymbol{0}\,, \tag{6.3}$$

where $(\boldsymbol{B}_t)_{t\geq 0}$ is a standard Brownian motion and $\boldsymbol{m}(\boldsymbol{y}, t) = \mathbb{E}[\boldsymbol{\sigma}|t\boldsymbol{\sigma} + \sqrt{t}\boldsymbol{g} = \boldsymbol{y}]$ (conditioning over $\boldsymbol{G}$ is implicit here), with the conditional expectation being taken with respect to $(\boldsymbol{\sigma}, \boldsymbol{g}) \sim \mu_{H_N} \otimes \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_N)$. The key remark (see Section 6.3.1) is that $(\boldsymbol{y}_t)_{t\geq 0}$ thus defined has the same distribution at $(t\boldsymbol{\sigma} + \boldsymbol{B}'_t)_{t\geq 0}$ (with $\boldsymbol{B}'_t$ a different Brownian motion) and therefore $\boldsymbol{y}_t/t$ converges to a sample from the desired measure. Of course, constructing an actual algorithm requires to discretize time and — crucially — to define an efficient algorithm that approximates the conditional mean $\boldsymbol{m}(\,\cdot\,, t)$ well enough.

This idea was introduced in [AMS22], which developed it as an algorithmic implementation of Eldan's stochastic localization (SL) technique for proving functional inequalities in high-dimensional probability [Eld13, Eld20b, Eld22]. Independently, the same process attracted considerable interest in machine learning, under the name of 'denoising diffusions' [SDWMG15, HJA20, SSDK+21]. In that context, the target distribution μ is unknown, but one is given access to samples $\boldsymbol{\sigma}_1, \dots, \boldsymbol{\sigma}_n$. These are used to learn an approximation of the drift $\boldsymbol{m}(\,\cdot\,, t)$ and hence generate trajectories from Eq. (6.3) (We refer to [Mon23b] for further discussion of this connection.)

The analysis of [AMS22, Cel24, AMS23b] establishes that algorithmic SL samples from the Gibbs measure of the Sherrington-Kirkpatrick model on the (more difficult) cube $\Sigma_N = \{-1, 1\}^N$, up to vanishing normalized Wasserstein error. That is, with probability $1 - o_N(1)$ over the Sherrington-Kirkpatrick Hamiltonian H_N, there is a coupling of μ_{H_N} and μ^{alg} such that for $(\boldsymbol{\sigma}, \boldsymbol{\sigma}^{\mathsf{alg}})$ drawn from this coupling,

$$\frac{1}{N} \mathop{\mathbb{E}}_{\boldsymbol{\sigma}, \boldsymbol{\sigma}^{\mathsf{alg}}} \|\boldsymbol{\sigma} - \boldsymbol{\sigma}^{\mathsf{alg}}\|_2^2 = o_N(1). \tag{6.4}$$

Remarkably, this guarantee holds throughout the 'replica symmetric' phase of the Sherrington-Kirkpatrick model [Tal10], i.e. for all $\beta < 1$ [Cel24]. Conversely [AMS22] proves that no randomized algorithm whose output is a Lipschitz continuous function of the Hamiltonian H_N (for each random seed) succeeds for $\beta > 1$. In particular, in this case, generating the whole SL path appears to not be fundamentally harder than sampling from the target measure μ_{H_N}. On the other hand, the Hamiltonian of the Sherrington-Kirkpatrick model is quadratic, and obtaining similar results for the general polynomial Hamiltonian of Eq. (6.1) poses entirely new challenges [AMS23b].

Despite the intense research activity in this area, it is fair to say that many fundamental questions remain open from a mathematical viewpoint. First, previous works on algorithmic SL produced algorithms that sampled from the target measure in the sense of vanishing normalized Wasserstein error. It is important to ask whether this approach can be upgraded to sample with vanishing total variation error. Second, generating an approximation of the whole SL trajectory $(\boldsymbol{y}_t)_{t\geq 0}$ is *a priori* a more challenging task than sampling from the target distribution μ. One may thus ask whether there is a fundamental gap between measures $\mu(\mathsf{d}\boldsymbol{\sigma})$ that can be sampled using general polynomial-time algorithms, and those that instead can be sampled via SL paths.

In this paper, we address both questions. We first affirmatively answer the first question. Our main result is an algorithm sampling with vanishing total variation error from the Gibbs measure of any spherical spin glass satisfying

$$\xi''(s) < \frac{1}{(1-s)^2}, \qquad \forall s \in [0,1). \tag{6.5}$$

As we will soon discuss, the condition (6.5) presents a fundamental barrier to algorithms for generating sample paths from SL. The key challenge in implementing the algorithmic SL approach is the construction of an efficient algorithm to approximate the mean of the measure $\mu_{H_N}(\mathsf{d}\boldsymbol{\sigma})$, as well as its conditional mean given Gaussian observations $\boldsymbol{y}_t$. The latter corresponds to the mean of a exponential tilt $\mu_{H_N,\boldsymbol{y}}(\mathsf{d}\boldsymbol{\sigma}) \propto \exp(\langle \boldsymbol{y}_t, \boldsymbol{\sigma}\rangle)\mu_{H_N}(\mathsf{d}\boldsymbol{\sigma})$. Approximating $\boldsymbol{m}(\boldsymbol{y})$ was achieved in [AMS22] by a variational approach that requires minimizing the so called Thouless-Anderson-Palmer (TAP) free energy [TAP77]. The same paper established that the resulting estimate satisfies (with high probability) $\|\boldsymbol{m}(\boldsymbol{y}) - \boldsymbol{m}^{\mathsf{TAP}}(\boldsymbol{y})\|^2 = o(N)$. (For the case of a measure supported over S_N, the function $\boldsymbol{m}(\,\cdot\,)$ does not depend on t, and we will therefore omit this argument.)

Note that $\|\boldsymbol{m}(\boldsymbol{y})\|^2 = \Theta(N)$, and therefore [AMS22] establishes the weakest non-trivial upper bound on $\|\boldsymbol{m}(\boldsymbol{y}) - \boldsymbol{m}^{\mathsf{TAP}}(\boldsymbol{y})\|^2$. This is the standard level of precision for estimates in spin glass theory both in physics and mathematics [MPV87, Tal10].

However, in order to obtain a sampling algorithm with guarantees in total variation distance, it is necessary to construct an efficient estimator $\widehat{\boldsymbol{m}}(\boldsymbol{y})$ satisfying $\|\boldsymbol{m}(\boldsymbol{y}) - \widehat{\boldsymbol{m}}(\boldsymbol{y})\|^2 = o(1)$. The construction and analysis of such an estimator is the main problem solved in the present paper.

In fact we prove the following:

1. The TAP estimator is significantly more accurate than what could be hoped from the analysis of [AMS22, AMS23b]. Namely, we prove that $\|\boldsymbol{m}(\boldsymbol{y}) - \boldsymbol{m}^{\mathsf{TAP}}(\boldsymbol{y})\|^2 = O(1)$.

2. We design a correction $\boldsymbol{\Delta}(\boldsymbol{y})$ to the TAP estimator that can be computed efficiently and such that, letting $\widehat{\boldsymbol{m}}(\boldsymbol{y}) = \boldsymbol{m}^{\mathsf{TAP}}(\boldsymbol{y}) + \boldsymbol{\Delta}(\boldsymbol{y})$, we achieve the desired accuracy $\|\boldsymbol{m}(\boldsymbol{y}) - \widehat{\boldsymbol{m}}(\boldsymbol{y})\|^2 = o(1)$.

To the best of our knowledge, neither of these results was known before, even at a heuristic level.

Turning to the second question on the general behavior of algorithms based on SL paths, we will show that the general situation is richer than for the Sherrington-Kirkpatrick model. Indeed, it is reasonable to conjecture that a necessary condition for algorithmic SL to succeed is that the conditional law of $\boldsymbol{\sigma}$ given $\boldsymbol{y}_t$ remains replica symmetric for all t. While one might naively expect that this conditional measure is simpler than the original one, statistical physics calculations suggest that this might not be the case [MRTS07, GDKZ24]. Namely, it can be the case that the original measure $\mu_{H_N}(\mathsf{d}\boldsymbol{\sigma})$ is replica symmetric and does not exhibit shattering (see [GJK23, AMS25, BJ24] and below), while the localized measure $\mu_{H_N}(\mathsf{d}\boldsymbol{\sigma}|\boldsymbol{y}_t)$ exhibits replica symmetry breaking for some t.

Within the spherical spin glass studied in this paper, the conjectured threshold for efficient sampling is $\xi'(q) < q/(1-q)$ for all $q \in (0,1)$, and shattering is believed to take place beyond this threshold. The conditional measure generated by the SL process at time t is a spherical model with mixture $\xi_t(q) = \xi(q) + tq$ and the condition for absence of shattering for all t turns out to be t is $\xi'(q) + t = q/(1-q)$ having only one solution in $[0,1)$ for all t. This is in turn equivalent to (6.5), under which our sampling algorithm succeeds.

We further show that the condition (6.5) is a fundamental barrier for algorithmic SL. We show that in the subset of the replica symmetric regime where (6.5) does not hold, a form of algorithmic SL with generalized side-channels does not succeed. (We also expect this algorithm to continue to fail beyond the replica symmetric regime, see Remark 6.2.4.) As the conjectured regime for efficient sampling is $\xi'(q) < q/(1-q)$ for all $q \in (0,1)$, which strictly contains the regime (6.5), this suggests a separation between sampling by general polynomial-time algorithms and by SL paths.

Remark 6.1.1. The threshold (6.5) is remarkably close to the threshold for shattering in the measure μ_{H_N}. For the special case of pure models $\xi(s) = \beta^2 s^p$, (6.5) holds for all $\beta < \beta_{\mathsf{SL}}(p)$, where we defined the stochastic localization inverse temperature as

$$\beta_{\mathsf{SL}}(p) := \frac{1}{2}\sqrt{\left(\frac{p}{p-1}\right)\left(\frac{p}{p-2}\right)^{p-2}}. \tag{6.6}$$

For large p we have $\beta_{\mathsf{SL}}(p) = e/2 + O(1/p)$. On the other hands the conjectured no-shattering condition $\xi'(q) < q/(1-q)$ holds for all $\beta < \beta_{\mathsf{sh}}(p)$ where

$$\beta_{\mathsf{sh}}(p) = \sqrt{\frac{(p-1)^{p-1}}{p(p-2)^{p-2}}} \,. \tag{6.7}$$

For large p, $\beta_{\mathsf{sh}}(p) = \sqrt{e} + O(1/p)$. In particular $\beta_{\mathsf{sh}}(p)/\beta_{\mathsf{SL}}(p) \le 2/\sqrt{e} \approx 1.213$ for all p. Further, the two thresholds coincide in mixed models when $\xi''(0)$ is sufficiently large.

6.1.1 Further background and related work

A substantial line of work in probability theory studies Langevin dynamics for the Gibbs measure (6.2). This is defined as the following diffusion on S_N

$$\mathsf{d}\boldsymbol{\sigma}_t = \left(\mathsf{proj}^{\perp}_{\boldsymbol{\sigma}_t}\nabla H_N(\boldsymbol{\sigma}_t) - \frac{N-1}{2N}\boldsymbol{\sigma}_t\right)\mathsf{d}t + \sqrt{2}\mathsf{proj}^{\perp}_{\boldsymbol{\sigma}_t}\mathsf{d}\boldsymbol{B}_t \,, \tag{6.8}$$

where $\boldsymbol{B}_t$ is a standard N-dimensional Brownian motion, and $\mathsf{proj}^{\perp}_{\boldsymbol{\sigma}_t}$ is the projector orthogonal to $\boldsymbol{\sigma}_t$. Langevin dynamics is a Markov process reversible for the measure μ_{H_N} of Eq. (6.2). As mentioned above, suitable discretizations of Langevin dynamics can be used to sample from μ_{H_N} [Dal17, DM17, DCWY19].

An asymptotically exact characterization of Langevin dynamics on short times horizons $t = O(1)$, in the high-dimensional limit $N \to \infty$, is provided by the so-called Cugliandolo-Kurchan or 'dynamical mean-firld theory' equations. These were studied first in physics [CHS93, CK93] and subsequently established rigorously in probability theory [BDG06]. Unfortunately, this approach does not give access to mixing times. On top of that, the Cugliandolo-Kurchan equations proved difficult to analyze rigorously except at sufficiently 'high temperature' (i.e. when $\xi(s) = \beta^2\xi_1(s)$, for a fixed ξ_1 and β small enough) [DGM07].

Based on a postulated asymptotic form of the Cugliandolo-Kurchan equations, as well as on thermodynamic calculations, physicists conjecture a phase transition in the mixing time of Langevin dynamics, when initialized uniformly at random [CHS93, CK93] . Namely, they expect the mixing time to be polynomial in N for

$$\xi'(q) < \frac{q}{1-q}, \qquad \forall q \in (0,1) \,. \tag{6.9}$$

and exponentially large in the opposite case, and more precisely when $\sup_{q\in(0,1)}(1-q)\xi'(q)/q > 1$. This is commonly referred to as the 'dynamical phase transition,' and corresponds to a phase transition in the geometry of the Gibbs measure, known as 'shattering phase transition.' In the homogeneous case $\xi(t) = \beta^2 t^p$, the above formula implies that the dynamical/shattering phase transition takes place at $\beta = \beta_{\mathsf{sh}}(p)$ given by Eq. (6.7).

We also recall that a second phase transition ('condensation' or 'static' or 'replica symmetry breaking') takes place at a lower temperature

$$\beta_c^2(p) = \inf_{s\in[0,1]}\left(\frac{1}{s^p}\log\left(\frac{1}{1-s}\right) - \frac{1}{s^{p-1}}\right) \,. \tag{6.10}$$

This corresponds to a non-analiticity of the free energy, and to the temperature at which the overlap stops concentrating [Che13]. For large p, we have $\beta_c(p) = \sqrt{\log p}(1 + o_p(1))$.

Towards the goal of proving the dynamical phase transition phenomenon, Ben Arous and Jagannath [BJ24] established that — for the homogeneous model — shattering takes place in a non-empty temperature interval, implying in particular $\beta_{\mathsf{sh}}(p) < \beta_c(p)$ strictly. A order-optimal bound was proven in [AMS25], who proved $\beta_{\mathsf{sh}}(p) \le C$ for a p-independent constant C.

A bolder version of the dynamical phase transition conjecture postulates that not only Langevin dynamics is slow, but indeed sampling is fundamentally hard beyond the shattering phase transition. Rigorous evidence was provided in [AMS25], which proves that 'stable algorithms' fail to sample from μ_{H_N} under shattering.

In the positive direction Gheissari and Jagannath [GJ19] proved that there exists $\underline{\beta}(p) > 0$ such that Langevin dynamics mixes rapidly for $\beta < \underline{\beta}(p)$. These authors also note that their proof technique extends to mixed models.

A closely related model is the Ising version of model (6.2), whereby the uniform measure μ_0 over the sphere S_N is replaced by the uniform measure over the hypercube $\{+1,-1\}^N$. A dynamical/shattering phase transition was conjectured in that setting as well [KT87], although at a different temperature. In this context, shattering for a non-empty interval of temperatures was proven in [GJK23], while mixing of Glauber dynamics at high temperature was proven in [ABXY24, AJK+24]. As for the spherical case, positive and negative results are separated by a large gap, indeed diverging with p.

The algorithmic stochastic localization approach was applied to Ising mixed p-spin spin classes in [AMS23b], which established the Wasserstein guarantee (6.4).

6.1.2 Notations

Throughout this paper, $\|\boldsymbol{\sigma}\|_N = \|\boldsymbol{\sigma}\|/\sqrt{N} = \sqrt{\boldsymbol{\sigma}^\top\boldsymbol{\sigma}/N}$ is the norm corresponding to the inner product $\langle\boldsymbol{\sigma}_1,\boldsymbol{\sigma}_2\rangle_N = \langle\boldsymbol{\sigma}_1,\boldsymbol{\sigma}_2\rangle/N = \boldsymbol{\sigma}_1^\top\boldsymbol{\sigma}_2/N$. There will be no confusion with the ℓ_p norm, which will not appear. Given a matrix $\boldsymbol{A}$, we denote by $\|\boldsymbol{A}\|_{\mathrm{F}}$ its Frobenius norm. For $\boldsymbol{m}\in\mathbb{R}^N$, measurable $I\subseteq\mathbb{R}$, and $\rho>0$, we define

$$\begin{aligned}\mathsf{Band}(\boldsymbol{m},I) &:= \{\boldsymbol{\sigma}\in S_N : \langle\boldsymbol{m},\boldsymbol{\sigma}\rangle_N \in I\}\ ,\\ \mathsf{B}_N(\boldsymbol{m},\rho) &:= \left\{\boldsymbol{x}\in\mathbb{R}^N : \|\boldsymbol{x}-\boldsymbol{m}\|_N \le \rho\right\}\ .\end{aligned}$$

We will occasionally abuse notations and write, for $q\in\mathbb{R}$, $\mathsf{Band}(\boldsymbol{m},q)$ instead of $\mathsf{Band}(\boldsymbol{m},\{q\})$.

We will often state that certain events occur with probability $1-e^{-cN}$. When we do, $c>0$ is an unspecified constant, which may change from line to line and may depend on all parameters other than N. We use p-lim to denote limit in probability.

We write $\boldsymbol{G}\sim\mathsf{GOE}(N)$ if $\boldsymbol{G}$ is a symmetric matrix with independent centered Gaussian entries on or above the diagonal with $G_{ii}\sim\mathcal{N}(0,2/N)$ and $G_{ij}\sim\mathcal{N}(0,1/N)$ for $i<j$.

Throughout the paper, the mixture ξ is fixed and various constants can depend on ξ but we will track this dependence. If ι is a small constant, we write $\iota' = o_\iota(1)$ if $|\iota'|\le h(\iota)$ where h is a function independent of N, such that $\lim_{\iota\to 0} h(\iota)=0$.

6.2 Main result

In this section we describe the sampling algorithm and state our main result. Throughout, we assume the model ξ satisfies (6.5).

6.2.1 Mean estimation of tilted measure

We first describe the main subroutine of our algorithm, which estimates the mean of the following exponentially tilted version of μ_{H_N}. For $\boldsymbol{y}\in\mathbb{R}^N$, define

$$\mu_{H_N,\boldsymbol{y}}(\mathsf{d}\boldsymbol{\sigma}) = \frac{1}{Z(\boldsymbol{y})}\exp\{H_N(\boldsymbol{\sigma})+\langle\boldsymbol{y},\boldsymbol{\sigma}\rangle\}\ \mu_0(\mathsf{d}\boldsymbol{\sigma})\,. \tag{6.11}$$

The tilt $\boldsymbol{y}$ will be generated by the outer loop of the algorithm described in Subsection 6.2.2, which implements a discretized version of the stochastic localization process. The outer loop also provides a time $t>0$, which this subroutine will take as input. The algorithm consists of three steps as outlined below. We defer the description of the correction $\boldsymbol{\Delta}(\boldsymbol{m})$ to Section 6.2.3.

(1) Let $\xi_t(s)=\xi(s)+ts$, and define the sequence $\{q_k : k\ge 0\}$ by $q_0=0$ and

$$q_{k+1} = \frac{\xi_t'(q_k)}{1+\xi_t'(q_k)}. \tag{6.12}$$

Starting from initialization $\boldsymbol{m}^{-1}=\boldsymbol{w}^0=\boldsymbol{0}$, run the approximate message passing (AMP) iteration

$$\boldsymbol{m}^k = (1-q_k)\boldsymbol{w}^k, \qquad \boldsymbol{w}^{k+1} = \nabla H_N(\boldsymbol{m}^k)+\boldsymbol{y}-(1-q_k)\xi''(q_k)\boldsymbol{m}^{k-1}, \tag{6.13}$$

for K_{AMP} iterations. Let $\boldsymbol{m}^{\mathsf{AMP}} = \boldsymbol{m}^{K_{\mathrm{AMP}}}$.

(2) Define

$$\theta(s) = \xi(1) - \xi(s) - (1-s)\xi'(s) \tag{6.14}$$

and the TAP free energy

$$\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m};\boldsymbol{y}) = H_N(\boldsymbol{m}) + \langle \boldsymbol{y}, \boldsymbol{m}\rangle + \frac{N}{2}\theta(\|\boldsymbol{m}\|_N^2) + \frac{N}{2}\log(1-\|\boldsymbol{m}\|_N^2). \tag{6.15}$$

Starting from $\boldsymbol{m}^{\mathsf{AMP}}$, run gradient ascent on $\mathcal{F}_{\mathsf{TAP}}(\cdot;\boldsymbol{y})$ for $K_{\mathsf{GD}}(N) := \lfloor K^*_{\mathsf{GD}} \log N \rfloor$ iterations, and let the resulting point be $\boldsymbol{m}^{\mathsf{GD}}$.

(3) Output $\boldsymbol{m}^{\mathsf{alg}} := \boldsymbol{m}^{\mathsf{GD}} + \boldsymbol{\Delta}(\boldsymbol{m}^{\mathsf{GD}})$, with $\boldsymbol{\Delta}(\boldsymbol{m})$ defined as in Section 6.2.3.

Pseudocode for the computation of $\boldsymbol{m}^{\mathsf{alg}}$ is provided in Algorithm 1.

Algorithm 1: Approximate mean computation

Input: H_N, $\boldsymbol{y} \in \mathbb{R}^N$, $t > 0$. Parameters: K_{AMP}, $K_{\mathsf{GD}}(N)$, $\eta > 0$
$\boldsymbol{m}^{-1} = \boldsymbol{w}^0 = \boldsymbol{0}$,
For $k = 0, \dots, K_{\mathsf{AMP}}$, run iteration (6.13)
Let $\boldsymbol{u}^0 = \boldsymbol{m}^{\mathsf{AMP}} = \boldsymbol{m}^{K_{\mathsf{AMP}}}$
for $k = 0, \dots, K_{\mathsf{GD}}(N) - 1$ **do**
  $\boldsymbol{u}^{k+1} = \boldsymbol{u}^k - \eta \nabla \mathcal{F}_{\mathsf{TAP}}(\boldsymbol{u}^k; \boldsymbol{y})$
end
Let $\boldsymbol{m}^{\mathsf{GD}} = \boldsymbol{u}^{K_{\mathsf{GD}}(N)}$
return $\boldsymbol{m}^{\mathsf{alg}}(H_N, \boldsymbol{y}, t) = \boldsymbol{m}^{\mathsf{GD}} + \boldsymbol{\Delta}(\boldsymbol{m}^{\mathsf{GD}})$

6.2.2 Stochastic localization sampling

We are now in position to describe the sampling algorithm, which uses Algorithm 1 as a subroutine. The main idea is to truncate the diffusion process (6.3) to the interval $[0, T]$, and to replace it by its Euler discretization (see Step 6 in Algorithm 2 below).

We will prove that, for T a sufficiently large constant, the tilted measure of Eq. (6.11), with $\boldsymbol{y} = \boldsymbol{y}_T$ is well approximated by a strongly log-concave measure. As a consequence, we can sample from it in total variation using standard approaches such as the Metropolis-adjusted Langevin algorithm, or MALA (see [CLA+21] and references therein). Formally, define

$$\boldsymbol{\sigma}_{\boldsymbol{y}}(\boldsymbol{\rho}) = \frac{\widehat{\boldsymbol{y}} + \boldsymbol{U}\boldsymbol{\rho}}{\sqrt{1+\|\boldsymbol{\rho}\|_N^2}}, \qquad\qquad \widehat{\boldsymbol{y}} = \frac{\boldsymbol{y}}{\|\boldsymbol{y}\|_N}, \tag{6.16}$$

where $\boldsymbol{U} \in \mathbb{R}^{N\times(N-1)}$ is an orthonormal basis of the orthogonal complement of $\boldsymbol{y}$, and

$$H^{\mathsf{proj}}_{N,\boldsymbol{y}}(\boldsymbol{\rho}) = H_{N,\boldsymbol{y}}(\boldsymbol{\sigma}_{\boldsymbol{y}}(\boldsymbol{\rho})) - \frac{N}{2}\log(1+\|\boldsymbol{\rho}\|_N^2). \tag{6.17}$$

Note that $\boldsymbol{\sigma}_{\boldsymbol{y}}$ is the inverse of the stereographic projection $\boldsymbol{T}_{\boldsymbol{y}}$ from $S_N \cap \{\boldsymbol{\sigma} : \langle \boldsymbol{\sigma}, \boldsymbol{y}\rangle > 0\}$ to the affine plane $\{\widehat{\boldsymbol{y}} + \boldsymbol{U}\boldsymbol{\rho} : \boldsymbol{\rho} \in \mathbb{R}^{N-1}\}$. We will see (Lemma 6.9.5) that the push-forward of $\mu_{H_N,\boldsymbol{y}}(\cdot|\langle \boldsymbol{\sigma}, \boldsymbol{y}\rangle > 0)$ under $\boldsymbol{T}_{\boldsymbol{y}}$ is precisely

$$\nu^{\mathsf{proj}}_{H_N,\boldsymbol{y}}(\mathsf{d}\boldsymbol{\rho}) = \frac{1}{\widehat{Z}(\boldsymbol{y})} \exp H^{\mathsf{proj}}_{N,\boldsymbol{y}}(\boldsymbol{\rho})\ \mathsf{d}\boldsymbol{\rho}. \tag{6.18}$$

Let $\varepsilon_0 = 0.1$ and $\varphi : [0, +\infty) \to [0, +\infty)$ be a twice continuously differentiable function satisfying $\varphi(x) = 0$ for $x \in [0, \varepsilon_0]$ and

$$\frac{1}{(1+x)^{3/2}} + \varphi'(x) \geq \varepsilon_0, \qquad\qquad \frac{1-2x}{(1+x)^{5/2}} + \varphi'(x) + 2x\varphi''(x) \geq \varepsilon_0 \tag{6.19}$$

for all $x \geq 0$. (Existence of such a function is shown in Fact 6.9.9.) Define the following measure on $\mathbb{R}^{N-1}$:

$$\widetilde{\nu}^{\mathsf{proj}}_{H_N,\boldsymbol{y}}(\mathsf{d}\boldsymbol{\rho}) = \frac{1}{\widetilde{Z}(\boldsymbol{y})} \exp \widetilde{H}^{\mathsf{proj}}_{N,\boldsymbol{y}}(\boldsymbol{\rho})\ \mathsf{d}\boldsymbol{\rho}, \qquad \widetilde{H}^{\mathsf{proj}}_{N,\boldsymbol{y}}(\boldsymbol{\rho}) = H^{\mathsf{proj}}_{N,\boldsymbol{y}}(\boldsymbol{\rho}) - \frac{TN}{2}\varphi(\|\boldsymbol{\rho}\|_N^2). \tag{6.20}$$

We will show that for sufficiently large T, $\widetilde{\nu}^{\mathsf{proj}}_{H_N,\boldsymbol{y}}$ is strongly log-concave (Proposition 6.9.8) and approximates $\nu^{\mathsf{proj}}_{H_N,\boldsymbol{y}}$ in total variation (Corollary 6.9.7). Thus, we may sample from it using MALA, and produce samples from $\mu_{H_N,\boldsymbol{y}}$ by pushing forward through $\boldsymbol{\sigma}_{\boldsymbol{y}}$.

Algorithm 2: SAMPLING

Input: H_N. Parameters: K_{AMP}, $K_{\mathsf{GD}}(N)$, η, $T > 0$, where T is a multiple of N^{-4}

Set $\delta = N^{-4}$, $L = T/\delta$
Set $\boldsymbol{y}^0 = \boldsymbol{0}$
for $\ell = 0, \dots, L-1$ **do**
  Let $\boldsymbol{m}^\ell = \boldsymbol{m}^{\mathsf{alg}}(H_N, \boldsymbol{y}^\ell, \ell\delta)$ be the output of Algorithm 1 on input $(H_N, \boldsymbol{y}^\ell, \ell\delta, K_{\mathsf{AMP}}(N), K_{\mathsf{GD}}, \eta)$
  Draw $\boldsymbol{w}^\ell \sim \mathcal{N}(0, \boldsymbol{I}_N)$ independent of everything else
  Set $\boldsymbol{y}^{\ell+1} = \boldsymbol{y}^\ell + \delta\boldsymbol{m}^\ell + \sqrt{\delta}\boldsymbol{w}^\ell$
end
Let $\widetilde{\nu}^{\mathsf{proj}}_{H_N,\boldsymbol{y}^L}$ be defined according to Eq. (6.20)
Use MALA to sample from $\boldsymbol{\rho}^{\mathsf{MALA}} \sim \nu^{\mathsf{MALA}}$, to accuracy $\mathrm{TV}(\nu^{\mathsf{MALA}}, \widetilde{\nu}^{\mathsf{proj}}_{H_N,\boldsymbol{y}^L}) \leq 1/N$
return $\boldsymbol{\sigma}_{\boldsymbol{y}^L}(\boldsymbol{\rho}^{\mathsf{MALA}})$

Theorem 6.2.1. *Suppose ξ satisfies (6.5). There exist constants $K_{\mathsf{AMP}}, K^*_{\mathsf{GD}}, \eta, T$ depending on ε and ξ such that running Algorithm 2 with parameters K_{AMP}, $K_{\mathsf{GD}}(N) = K^*_{\mathsf{GD}} \log N$, η, T, the following holds. With probability $1 - o_N(1)$ over H_N, $\mu^{\mathsf{alg}} = \mathcal{L}(\boldsymbol{\sigma}^{\mathsf{alg}})$ satisfies*

$$\mathrm{TV}(\mu^{\mathsf{alg}}, \mu_{H_N}) \leq o_N(1).$$

Further the complexity of the algorithm is upper bounded by $CN^4\,(N + \chi_{\nabla H}) \log N + \chi_{\text{log-conc}}$, where $\chi_{\nabla H}$ is the complexity of evaluating $\nabla H_N(\boldsymbol{m})$ at a point $\boldsymbol{m}$ with $\|\boldsymbol{m}\|_N \leq 1$, and $\chi_{\text{log-conc}}$ is the complexity of sampling from a 1-strongly log-concave measure in N dimension using MALA *to accuracy $1/N$ in total variation.*

Remark 6.2.2. The main result of [CLA+21] implies that, for a 'warm start' initialization $\chi_{\text{log-conc}}$ is of order $N^{3/2} \log N$. In the present case we do not have a good warm start, and obtain $\chi_{\text{log-conc}} \leq C \cdot N^{5/2}$. We believe this bound is suboptimal, but made no attempt at improving it.

6.2.3 The correction $\boldsymbol{\Delta}(\boldsymbol{m})$

We now describe the computation of the correction $\boldsymbol{\Delta}(\boldsymbol{m})$. Let $\mathsf{T}_{\boldsymbol{m}}$ be the $(N-1)$-dimensional subspace orthogonal to $\boldsymbol{m}$ and define $H_N(\,\cdot\,;\boldsymbol{m}) : \mathsf{T}_{\boldsymbol{m}} \to \mathbb{R}$ via $H_N(\boldsymbol{x};\boldsymbol{m}) := H_N(\boldsymbol{m}+\boldsymbol{x})$. We then define the tensors

$$\boldsymbol{A}^{(2)}(\boldsymbol{m}) := \nabla^2_{\boldsymbol{x}} H_N(\boldsymbol{0};\boldsymbol{m})\,, \quad \boldsymbol{A}^{(3)}(\boldsymbol{m}) := \nabla^3_{\boldsymbol{x}} H_N(\boldsymbol{0};\boldsymbol{m})\,. \tag{6.21}$$

These should be interpreted as tensors $\boldsymbol{A}^{(i)}(\boldsymbol{m}) \in \mathsf{T}_{\boldsymbol{m}}^{\otimes i}$. Let $\gamma_{*,N}(\boldsymbol{m})$ be the unique solution of

$$\begin{cases} \mathsf{Tr}\big((\gamma_{*,N}\boldsymbol{I}_{N-1} - \boldsymbol{A}^{(2)}(\boldsymbol{m}))^{-1}\big) = N \cdot \Big(1 - \|\boldsymbol{m}\|^2/N\Big) & , \\ \gamma_{*,N} > \lambda_{\max}(\boldsymbol{A}^{(2)}(\boldsymbol{m})) & . \end{cases} \tag{6.22}$$

Here $\boldsymbol{I}_{N-1}$ denotes the identity matrix acting on $\mathsf{T}_{\boldsymbol{m}}$, and the inverse is over quadratic forms on $\mathsf{T}_{\boldsymbol{m}}$.

Then we define

$$\Delta_i(\boldsymbol{m}) = \frac{1}{2}\langle \boldsymbol{A}^{(3)}(\boldsymbol{m}), \boldsymbol{Q}(\boldsymbol{m}) \otimes \boldsymbol{Q}(\boldsymbol{m})_{i,\cdot}\rangle = \frac{1}{2}\sum_{a,b,c=1}^{N} A^{(3)}_{abc}(\boldsymbol{m})Q_{ia}(\boldsymbol{m})Q_{bc}(\boldsymbol{m})\,, \tag{6.23}$$

$$\boldsymbol{Q}(\boldsymbol{m}) := \big(\gamma_{*,N}(\boldsymbol{m})\boldsymbol{I}_{N-1} - \boldsymbol{A}^{(2)}(\boldsymbol{m})\big)^{-1}\,. \tag{6.24}$$

It is useful to make two additional remarks about the evaluation of $\boldsymbol{\Delta}(\boldsymbol{m})$:

1. For any fixed $\boldsymbol{m}$, $\boldsymbol{A}^{(2)}(\boldsymbol{m}) \overset{d}{=} \sqrt{\xi''(\|\boldsymbol{m}\|_N^2)\cdot \frac{N-1}{N}}\,\boldsymbol{W}$, for $\boldsymbol{W} \sim \mathsf{GOE}(N-1)$. It turns out that, although $\boldsymbol{m}^{\mathsf{TAP}}$ is itself random, this nonetheless gives the correct asymptotics for $\gamma_{*,N}(\boldsymbol{m}^{\mathsf{TAP}})$. Let $q_* = q_*(t)$ be the solution to $\frac{q_*}{1-q_*} = \xi_t'(q_*)$, existence and uniqueness of which is shown in Fact 6.4.2. We will show (see Proposition 6.4.4) that typically $\|\boldsymbol{m}^{\mathsf{TAP}}\|_N^2 = q_* + o_N(1)$, and (see Lemma 6.6.22) $\gamma_{*,N}(\boldsymbol{m}^{\mathsf{TAP}}) = \gamma_* + o_N(1)$, for $\gamma_* = (1-q_*)^{-1} + (1-q_*)\xi''(q_*)$. For the computation of $\boldsymbol{\Delta}$, we can replace $\gamma_{*,N}$ by γ_* with negligible error.

2. The tensors $\boldsymbol{A}^{(2)}(\boldsymbol{m})$ and $\boldsymbol{A}^{(3)}(\boldsymbol{m})$ can be written as explicit linear functions of the couplings $\boldsymbol{g}$, and hence can be computed efficiently without need to take any numerical derivative.

6.2.4 Fundamental limits of algorithmic SL, replica symmetry breaking, and shattering

It is useful to compare condition (6.5) with the condition for (absence of) shattering, and replica symmetry breaking:

- As mentioned above (cf. Eq. (6.9)), it is conjectured [CHS93, CS95, BCKM98] that shattering is absent if and only if

$$\xi'(q) < \frac{q}{1-q}\,, \quad \forall q \in (0,1)\,. \tag{6.25}$$

  This is implied by the condition under which our algorithm succeeds, namely Eq. (6.5), by integrating once.

- The tight condition for replica symmetry was identified in [Tal06a, Proposition 2.3].

$$\xi(q) + q + \log(1-q) \le 0, \qquad \forall q \in [0,1) \tag{6.26}$$

  Note that this holds under (6.25) by integrating once, and hence under (6.5).

In this section, we prove that the condition (6.5) is necessary not only for Algorithm 2 to succeed, but indeed for a broader class of stochastic localization schemes that we next introduce. This points at a fundamental gap between such schemes and the possible computational limit for sampling, a fact that was suggested in [GDKZ24] and, in a related context, in [MRTS07].

By the key remark below (6.3), the process $\boldsymbol{y}_t$ generated by (6.3) consists of observations of some $\boldsymbol{\sigma} \sim \mu_{H_N}$ through a progressively less noisy Gaussian channel. A natural generalization of this process outputs observations of $\boldsymbol{\sigma}, \boldsymbol{\sigma}^{\otimes 2}, \boldsymbol{\sigma}^{\otimes 3}, \dots$ through Gaussian channels of varying signal strengths, and can similarly be converted to a sampling algorithm.

Consider any $J \in \mathbb{N}$ and continuously differentiable, coordinate-wise increasing $\tau : [0,+\infty) \to [0,+\infty)^J$, normalized to $\|\tau(t)\|_1 = t$ for all $t \in [0,+\infty)$, and such that $\lim_{t\to\infty} \tau_j(t) = \infty$ for at least one odd $j \le J$. For each $j \le J$, let $(\boldsymbol{B}_t^j)_{t\ge 0}$ be a standard Brownian motion in $(\mathbb{R}^N)^{\otimes j}$. Let $(\vec{\boldsymbol{y}}_t)_{t\ge 0} = (\boldsymbol{y}_t^1, \dots, \boldsymbol{y}_t^J) \in \mathbb{R}^N \times \cdots \times (\mathbb{R}^N)^{\otimes J}$ be given by the Ito diffusion

$$\mathsf{d}\boldsymbol{y}_t^j = \tau_j'(t)\boldsymbol{m}_j(\vec{\boldsymbol{y}}_t, t)\,\mathsf{d}t + \tau_j'(t)^{1/2}\,\mathsf{d}\boldsymbol{B}_t^j, \qquad \vec{\boldsymbol{y}}_0 = \boldsymbol{0}, \tag{6.27}$$

where, with expectation over $\boldsymbol{\sigma} \sim \mu_{H_N}$ and $\boldsymbol{G}^j \sim \mathcal{N}(0, \boldsymbol{I}_N^{\otimes j})$,

$$\boldsymbol{m}_j(\vec{\boldsymbol{y}}_t, t) = \mathbb{E}[\boldsymbol{\sigma}^{\otimes j} | \tau_i(t)\boldsymbol{\sigma}^{\otimes i} + \tau_i(t)^{1/2}\boldsymbol{G}^i = \boldsymbol{y}_t^i, \forall 1 \le i \le J]. \tag{6.28}$$

The process (6.3) corresponds to the case $J = 1$. As in that case, a sampling algorithm can be constructed from Eq. (6.27) by discretizing time and approximating the calculation of$\boldsymbol{m}_j(\vec{\boldsymbol{y}}_t, t)$ (see Remark 6.2.5 below).

For $\boldsymbol{A} \in (\mathbb{R}^N)^{\otimes j}$ and $1 \le \ell \le j$, let $\boldsymbol{A}^{(\ell)}$ be the tensor obtained by rotating coordinates by $i \pmod j$, that is

$$\boldsymbol{A}^{(\ell)}_{i_1,\dots,i_j} = \boldsymbol{A}_{i_{\ell+1},\dots,i_j,i_1,\dots,i_\ell}.$$

Then, for $\boldsymbol{B} \in (\mathbb{R}^N)^{\otimes j-1}$, let $(\boldsymbol{A}, \boldsymbol{B})_{\mathsf{sym}} \in \mathbb{R}^N$ be the vector satisfying

$$\langle \boldsymbol{v}, (\boldsymbol{A}, \boldsymbol{B})_{\mathsf{sym}} \rangle = \sum_{\ell=1}^{j} \langle \boldsymbol{B} \otimes \boldsymbol{v}, \boldsymbol{A}^{(\ell)} \rangle$$

for all $\boldsymbol{v} \in \mathbb{R}^N$. Let

$$\check{\xi}_t(s) = \xi(s) + \sum_{j=1}^{J} \tau_j(t) s^j$$

and define sequence $\{\check{q}_k : k \ge 0\}$ by $q_0 = 0$ and

$$\check{q}_{k+1} = \frac{\check{\xi}_t'(q_k)}{1 + \check{\xi}_t'(q_k)} \tag{6.29}$$

Finally define an AMP iteration analogous to (6.13) by

$$\check{\boldsymbol{m}}^k = (1 - \check{q}_j)\check{\boldsymbol{w}}^k, \qquad \check{\boldsymbol{w}}^{k+1} = \nabla H_N(\check{\boldsymbol{m}}^k) + \sum_{j=1}^{J} \frac{1}{N^{j-1}} ((\check{\boldsymbol{m}}^k)^{\otimes j-1}, \boldsymbol{y}_t^j)_{\mathsf{sym}} - (1 - \check{q}_k)\check{\xi}''(\check{q}_k)\check{\boldsymbol{m}}^{k-1}. \tag{6.30}$$

The next theorem is proved in Section 6.10, under the following condition which is a strict form of (6.26).

$$\xi''(0) < 1, \qquad \xi(q) + q + \log(1 - q) < 0, \quad \forall q \in (0, 1). \tag{6.31}$$

Theorem 6.2.3. *Suppose that (6.31) holds and that there exists $q \in [0, 1)$ such that $\xi''(q) > \frac{1}{(1-q)^2}$. There exists a positive measure set $\mathcal{I} \subseteq [0, +\infty)$ such that for all $t \in \mathcal{I}$ the following holds. There exists $1 \le j \le J$ such that $\tau_j'(t) > 0$ and, for $\vec{\boldsymbol{y}}_t$ generated from (6.27),*

$$\lim_{k\to\infty} \liminf_{N\to\infty} \mathbb{E} \frac{1}{N^j} \left\| (\check{\boldsymbol{m}}^k)^{\otimes j} - \boldsymbol{m}_j(\vec{\boldsymbol{y}}_t, t) \right\|_2^2 > 0.$$

Remark 6.2.4. In this theorem we assume Eq. (6.31) to hold, but note that this an artifact of our proof technique. Indeed efficient sampling is believed to be impossible beyond the threshold (6.31). Indeed [AMS25] implies that 'stable' algorithms fail under replica symmetry breaking.

Remark 6.2.5. As alluded to above, we can define a natural analog of Algorithm 1 for this generalized setting, which computes an estimator $\check{\boldsymbol{m}}^{\mathsf{alg}}$ for $\boldsymbol{m}_1(\vec{\boldsymbol{y}}_t, t)$. For some $K_{\mathsf{AMP}} \in \mathbb{N}$, the point $\check{\boldsymbol{m}}^{K_{\mathsf{AMP}}}$ is the result of the first phase of this algorithm. The output $\check{\boldsymbol{m}}^{\mathsf{alg}}$ of this algorithm satisfies $\|\check{\boldsymbol{m}}^{K_{\mathsf{AMP}}} - \check{\boldsymbol{m}}^{\mathsf{alg}}\|_N \to 0$ as $K_{\mathsf{AMP}} \to \infty$; see Theorem 6.4.1 and Proposition 6.4.4 below, which show this for Algorithm 1 when (6.5) holds.

The analog of Algorithm 2 simulates the SDE (6.27) via an Euler discretization, estimating each $\boldsymbol{m}_j(\vec{\boldsymbol{y}}_t, t)$ with $(\check{\boldsymbol{m}}^{\mathsf{alg}})^{\otimes j}$. Theorem 6.2.3 shows that for a interval of t of positive measure, this algorithm fails for a tensor order j relevant to the Euler discretization.

6.3 Preliminaries

In this section we provide further background. The contents of Subsections 6.3.1 and 6.3.2 are known and we often refer to [AMS22, Sections 3 and 4.1] for proofs. Subsection 6.3.4 introduces a lemma about conditioning a Gaussian process on a random vector: this is a fairly standard but crucial technical tool.

6.3.1 Stochastic localization

Fix a realization of H_N. The stochastic localization process is defined by the SDE (6.3), which has unique strong solutions provided $\boldsymbol{y} \mapsto \boldsymbol{m}(\boldsymbol{y}, t)$ is Lipschitz continuous. Note that, for $\mu_{H_N, \boldsymbol{y}_t}$ as in (6.11), $\boldsymbol{m}$ is the mean

$$\boldsymbol{m}(\boldsymbol{y}, t) = \int \boldsymbol{\sigma} \; \mu_{H_N, \boldsymbol{y}}(\mathsf{d}\boldsymbol{\sigma}).$$

Therefore Lipschitz continuity is implied by $\sup_{\boldsymbol{y}} \|\mathsf{Cov}(\mu_{H_N,\boldsymbol{y}})\|_{\mathsf{op}} < \infty$ which always holds since $\mu_{H_n,\boldsymbol{y}}$ is supported on a compact set.

As already mentioned in the introduction, we have the following facts (see for instance [AM22]).

Proposition 6.3.1. *Let $(\boldsymbol{y}_t)_{t \geq 0}$ be the unique solution of the SDE (6.3). Then there exists a standard Brownian motion $\boldsymbol{B}'_t$ independent of $\boldsymbol{\sigma} \sim \mu_{H_N}$, such that, for all t, $\boldsymbol{y}_t = t\boldsymbol{\sigma} + \boldsymbol{B}'_t$.*

Further, $\mathbb{E}\mathsf{Cov}(\mu_{H_N,\boldsymbol{y}_t}) \preceq \boldsymbol{I}_N \,/\, t$. In particular $\mu_{H_N,\boldsymbol{y}_t} \Rightarrow \delta_{\boldsymbol{\sigma}}$ almost surely as $t \to \infty$.

6.3.2 Planted model and contiguity

Recall that μ_0 denotes the uniform probability measure on S_N. Further, let $\mathscr{H}_N$ be the space of Hamiltonians H_N (i.e. continuous functions $H_N : S_N \to \mathbb{R}$ endowed with the uniform convergence topology and the induced Borel sigma-algebra) and $\mu_{\mathsf{null}} \in \mathcal{P}(\mathscr{H}_N)$ be the law induced on H_N by Eq. (6.1). Define the planted measure $\mu_{\mathsf{pl}} \in \mathcal{P}(S_N \times \mathscr{H}_N)$ by

$$\mu_{\mathsf{pl}}(\mathsf{d}\boldsymbol{x}, \mathsf{d}H_N) := \frac{1}{Z_{\mathsf{pl}}} \exp\left\{H_N(\boldsymbol{x})\right\} \, \mathsf{d}\mu_0(\boldsymbol{x}) \mathsf{d}\mu_{\mathsf{null}}(H_N).$$

For $H_N \in \mathscr{H}_N$, define the partition function

$$Z(H_N) := \int \exp\left\{H_N(\boldsymbol{\sigma})\right\} \, \mu_0(\mathsf{d}\boldsymbol{\sigma}) \, .$$

Lemma 6.3.2 (Proved in Section 6.8)**.** *Suppose ξ satisfies (6.31). Let $W \sim \mathcal{N}(-\frac{1}{2}\sigma^2, \sigma^2)$, where $\sigma^2 = -\frac{1}{2}\log(1 - \xi''(0))$. As $N \to \infty$, for $H_N \sim \mu_{\mathsf{null}}$, the Radon-Nykodym derivative of μ_{pl} with respect to μ_{null} is*

$$\frac{\mathsf{d}\mu_{\mathsf{pl}}}{\mathsf{d}\mu_{\mathsf{null}}}(H_N) = \frac{Z(H_N)}{\mathbb{E}\, Z(H_N)} \xrightarrow{d} \exp(W).$$

Remark 6.3.3. In most of this paper, we are interested in ξ satisfying the condition (6.5), which implies (6.31) by integrating twice. However, the proof of Theorem 6.2.3 in Section 6.10 only assumes ξ satisfies (6.31), so we state this lemma with the more general condition.

For any $T > 0$, let $\mathbb{P}, \mathbb{Q} \in \mathcal{P}(S_N \times \mathscr{H}_N \times C([0,T], \mathbb{R}^N))$ be the laws of $(\boldsymbol{\sigma}, H_N, (\boldsymbol{y}_t)_{t \in [0,T]})$, generated as follows.

- Under $\mathbb{Q}$:

  $$H_N \sim \mu_{\mathsf{null}}, \quad \boldsymbol{\sigma} \sim \mu_{H_N}, \quad \boldsymbol{y}_t = t\boldsymbol{\sigma} + \boldsymbol{B}_t \, , \tag{6.32}$$

  for $\boldsymbol{B}_t$ a standard Brownian motion independent of $\boldsymbol{\sigma}, H_N$. By Proposition 6.3.1, an equivalent description of this distribution is: $H_N \sim \mu_{\mathsf{null}}$, $(\boldsymbol{y}_t)_{t \geq 0}$ given by the SDE (6.3) and $\boldsymbol{\sigma} = \lim_{t \to \infty} \boldsymbol{y}_t / t$.

- Under $\mathbb{P}$:

  $$(H_N, \boldsymbol{\sigma}) \sim \mu_{\mathsf{pl}}, \qquad \boldsymbol{y}_t = t\boldsymbol{\sigma} + \boldsymbol{B}_t \, , \tag{6.33}$$

  for $\boldsymbol{B}_t$ a standard Brownian motion independent of $\boldsymbol{\sigma}, H_N$. As before, we can equivalently generate first H_N, then $(\boldsymbol{y}_t)_{t \geq 0}$ given by the SDE (6.3) and finally $\boldsymbol{\sigma}$.

  The joint distribution of $(H_N, \boldsymbol{\sigma}) \sim \mu_{\mathsf{pl}}$ can be described in two equivalent ways. In the first one, we generate first H_N and then $\boldsymbol{\sigma}$ conditional on H_N:

  $$H_N \sim \mu_{\mathsf{pl}}(\mathsf{d}H_N) = \frac{Z(H_N)}{\mathbb{E}Z(H_N)} \mu_{\mathsf{null}}(\mathsf{d}H_N) \, , \qquad \boldsymbol{\sigma} \sim \mu_{H_N} \, . \tag{6.34}$$

In the second, we generate first $\boldsymbol{\sigma}$ and then H_N:

$$\boldsymbol{\sigma} \sim \mu_0\,, \qquad H_N \sim \mu_{\mathsf{pl}}(\mathsf{d}H_N|\boldsymbol{\sigma}) \,\propto\, e^{H_N(\boldsymbol{\sigma})}\mu_{\mathsf{null}}(\mathsf{d}H_N). \tag{6.35}$$

A short calculation shows that $H_N \sim \mu_{\mathsf{pl}}(\cdot|\boldsymbol{x})$ is given by

$$H_N(\boldsymbol{\sigma}) = N\xi(\langle \boldsymbol{x}, \boldsymbol{\sigma}\rangle_N) + \widetilde{H}_N(\boldsymbol{\sigma}), \tag{6.36}$$

where $\widetilde{H}_N \sim \mu_{\mathsf{null}}$. The above definition has the following immediate consequence.

Proposition 6.3.4 ([AMS22, Proposition 4.2])**.** *For all $T \geq 0$,*

$$\frac{\mathsf{d}\mathbb{P}}{\mathsf{d}\mathbb{Q}}(\boldsymbol{\sigma}, H_N, (\boldsymbol{y}_t)_{t\in[0,T]}) = \frac{Z(H_N)}{\mathbb{E}\, Z(H_N)}.$$

As a consequence of Lemma 6.3.2 and Proposition 6.3.4, Le Cam's first lemma implies the following.

Corollary 6.3.5. *The measures $\mathbb{P}$ and $\mathbb{Q}$ are mutually contiguous. That is, for any sequence of events $\mathcal{E}_N$, $\mathbb{P}(\mathcal{E}_N) \to 0$ if and only if $\mathbb{Q}(\mathcal{E}_N) \to 0$.*

Thus it suffices to analyze our algorithm under the planted distribution $\mathbb{P}$.

6.3.3 Basic regularity estimate

For a tensor $\boldsymbol{A} \in (\mathbb{R}^N)^{\otimes k}$, define the operator norm

$$\|\boldsymbol{A}\|_{\mathsf{op},N} = \frac{1}{N} \sup_{\|\boldsymbol{\sigma}^1\|_N,\dots,\|\boldsymbol{\sigma}^k\|_N\leq 1} |\langle \boldsymbol{A}, \boldsymbol{\sigma}^1 \otimes \cdots \otimes \boldsymbol{\sigma}^k\rangle|.$$

Notice that this normalization is different from the standard injective norm $\|\cdot\|_{\mathsf{inj}}$ in that $\|\boldsymbol{A}\|_{\mathsf{op},N} = N^{(k-2)/2}\|\boldsymbol{A}\|_{\mathsf{inj}}$.

Proposition 6.3.6 ([HS25, Proposition 2.3])**.** *There exists a sequence of constants $(C_k)_{k\geq 0}$ independent of N for which the following holds. Define the event*

$$K_N := \left\{ \sup_{\|\boldsymbol{\sigma}\|_N\leq 1} \|\nabla^k H_N(\boldsymbol{\sigma})\|_{\mathsf{op},N} \leq C_k \ \ \forall k \geq 0 \right\}.$$

Then $\mathbb{P}(K_N) \geq 1 - e^{-cN}$.

6.3.4 Conditioning lemma

Lemma 6.3.7. *Let $D \subseteq \mathbb{R}^N$ be an open set and $\mathcal{F} : D \to \mathbb{R}$ be a (not necessarily centered) C^2 Gaussian process on a probability space $(\Omega, \Sigma, \mathbb{P})$. Let X be a random variable on (Ω, Σ) taking values in $[0, 1]$, and $\boldsymbol{m}_0$ be a random vector on the same space taking values in $\mathbb{R}^N$. For $\varepsilon, c_{\mathsf{spec}}, c_{\mathsf{op}} > 0$ satisfying $\varepsilon \leq c_{\mathsf{spec}}^2/10c_{\mathsf{op}}$, define $U_{\boldsymbol{m}_0} := \mathsf{B}_N(\boldsymbol{m}_0, 5\varepsilon/c_{\mathsf{spec}})$ and the events*

$$\begin{aligned}
\mathcal{G}(\varepsilon, c_{\mathsf{spec}}) &:= \left\{ \|\nabla \mathcal{F}(\boldsymbol{m}_0)\|_N \leq \varepsilon\,, \quad \nabla^2 \mathcal{F}(\boldsymbol{m}_0) \preceq -c_{\mathsf{spec}} \boldsymbol{I}_n \right\},\\
\mathcal{H}(c_{\mathsf{op}}) &:= \left\{ \sup_{\boldsymbol{m}\in D} \|\nabla^2 \mathcal{F}(\boldsymbol{m})\|_{\mathsf{op},N} \leq c_{\mathsf{op}}, \quad \sup_{\boldsymbol{m}\in D} \|\nabla^3 \mathcal{F}(\boldsymbol{m})\|_{\mathsf{op},N} \leq c_{\mathsf{op}} \right\},\\
\mathcal{E}_{\mathsf{cond}} &:= \mathcal{G}(\varepsilon, c_{\mathsf{spec}}) \cap \mathcal{H}(c_{\mathsf{op}}) \cap \{\|\boldsymbol{m}_0\|_N \leq 1\} \cap \{U_{\boldsymbol{m}_0} \subseteq D\}\,.
\end{aligned}$$

Finally, assume $\boldsymbol{m} \mapsto \mathbb{E}\, \nabla \mathcal{F}(\boldsymbol{m})$ is continuous and $\lambda_{\min}(\mathsf{Cov}(\nabla \mathcal{F}(\boldsymbol{m})))$ is bounded away from 0 uniformly over $\boldsymbol{m} \in D$. Then, with $\varphi_{\nabla \mathcal{F}(\boldsymbol{m})}$ the probability density of $\nabla \mathcal{F}(\boldsymbol{m})$ w.r.t. Lebesgue measure on $\mathbb{R}^N$ and d^N denoting integration against this measure,

$$\mathbb{E}(X\mathbf{1}\{\mathcal{E}_{\mathsf{cond}}\}) = \int_D \mathbb{E}\left[|\det \nabla^2 \mathcal{F}(\boldsymbol{m})| X \mathbf{1}\{\mathcal{E}_{\mathsf{cond}} \cap \{\boldsymbol{m} \in U_{\boldsymbol{m}_0}\}\} \big| \nabla \mathcal{F}(\boldsymbol{m}) = \mathbf{0}\right] \varphi_{\nabla \mathcal{F}(\boldsymbol{m})}(\mathbf{0})\ \mathsf{d}^N \boldsymbol{m}.$$

Proof. On event $\mathcal{E}_{\mathsf{cond}}$, for all $\boldsymbol{m} \in U_{\boldsymbol{m}_0}$ we have

$$\lambda_{\max}(\nabla^2\mathcal{F}(\boldsymbol{m})) \le \lambda_{\max}(\nabla^2\mathcal{F}(\boldsymbol{m}_0)) + c_{\mathsf{op}}\|\boldsymbol{m}-\boldsymbol{m}_0\|_N \le -c_{\mathsf{spec}} + \frac{5\varepsilon c_{\mathsf{op}}}{c_{\mathsf{spec}}} \le -\frac{1}{2}c_{\mathsf{spec}}. \tag{6.37}$$

Since $\|\nabla\mathcal{F}(\boldsymbol{m}_0)\|_N \le \varepsilon$, there is exactly one solution to $\nabla\mathcal{F}(\boldsymbol{m}_*) = \mathbf{0}$ in $U_{\boldsymbol{m}_0}$, which is measurable on (Ω, Σ) and furthermore lies in $\mathsf{B}_N(\boldsymbol{m}_0, 4\varepsilon/c_{\mathsf{spec}})$. The strong concavity (6.37) implies that $\nabla\mathcal{F}$ is injective on $U_{\boldsymbol{m}_0}$ and its image contains a neighborhood of $\mathbf{0}$. By the area formula, for sufficiently small $\iota > 0$,

$$1 = \frac{1}{|\mathsf{B}_N(\mathbf{0},\iota)|}\int_{U_{\boldsymbol{m}_0}} |\det\nabla^2\mathcal{F}(\boldsymbol{m})|\mathbf{1}\{\|\nabla\mathcal{F}(\boldsymbol{m})\|_N \le \iota\}\ \mathsf{d}^N\boldsymbol{m}.$$

Multiplying by $X\mathbf{1}\{\mathcal{E}_{\mathsf{cond}}\}$ and taking expectations of both sides by Fubini yields

$$\begin{aligned}
&\mathbb{E}(X\mathbf{1}\{\mathcal{E}_{\mathsf{cond}}\}) \\
&= \frac{1}{|\mathsf{B}_N(\mathbf{0},\iota)|}\int_D \mathbb{E}\left[|\det\nabla^2\mathcal{F}(\boldsymbol{m})|X\mathbf{1}\{\mathcal{E}_{\mathsf{cond}}\cap\{\boldsymbol{m}\in U_{\boldsymbol{m}_0}\}\cap\{\|\nabla\mathcal{F}(\boldsymbol{m})\|_N\le\iota\}\}\right]\ \mathsf{d}^N\boldsymbol{m} \\
&= \int_D \mathbb{E}\left[|\det\nabla^2\mathcal{F}(\boldsymbol{m})|X\mathbf{1}\{\mathcal{E}_{\mathsf{cond}}\cap\{\boldsymbol{m}\in U_{\boldsymbol{m}_0}\}\}\big|\|\nabla\mathcal{F}(\boldsymbol{m})\|_N\le\iota\right]\frac{\mathbb{P}(\|\nabla\mathcal{F}(\boldsymbol{m})\|_N\le\iota)}{|\mathsf{B}_N(\mathbf{0},\iota)|}\ \mathsf{d}^N\boldsymbol{m}.
\end{aligned}$$

Note that on $\mathcal{E}_{\mathsf{cond}}$, $|\det\nabla^2\mathcal{F}(\boldsymbol{m})| \le c_{\mathsf{op}}^N$. Since $\mathcal{E}_{\mathsf{cond}}$ is contained in the event $\|\boldsymbol{m}_0\|_N \le 1$, $\{\boldsymbol{m}\in U_{\boldsymbol{m}_0}\}$ can only occur for $\boldsymbol{m}$ on a bounded set. Since $\lambda_{\min}(\mathsf{Cov}(\nabla\mathcal{F}(\boldsymbol{m})))$ is bounded away from 0, $\varphi_{\nabla\mathcal{F}(\boldsymbol{m})}$ is bounded, and thus so is $\mathbb{P}(\|\nabla\mathcal{F}(\boldsymbol{m})\|_N \le \iota)/|\mathsf{B}_N(\mathbf{0},\iota)|$. Therefore the integral in the last display is dominated by a bounded integrable function. Continuity of $\mathbb{E}\,\nabla\mathcal{F}(\boldsymbol{m})$ implies that $\varphi_{\nabla\mathcal{F}(\boldsymbol{m})}(\boldsymbol{z})$ is continuous in $\boldsymbol{z}$ in a neighborhood of $\mathbf{0}$. We take the $\iota\to 0$ limit of the last display by dominated convergence to conclude. □

6.4 Analysis of mean computation algorithm

The next several sections are devoted to the analysis of Algorithm 1. We fix $t \in [0,T]$ and consider $(\boldsymbol{x}, H_N, (\boldsymbol{y}_t)_{t\ge0}) \in S_N \times \mathscr{H}_N \times C([0,T],\mathbb{R}^N)$ distributed according to the planted law $\mathbb{P}$ defined in Eq. (6.33). Define

$$\begin{aligned}
H_{N,t}(\boldsymbol{\sigma}) &= H_N(\boldsymbol{\sigma}) + \langle\boldsymbol{y}_t,\boldsymbol{\sigma}\rangle \\
&= N\xi(\langle\boldsymbol{x},\boldsymbol{\sigma}\rangle_N) + \widetilde{H}_N(\boldsymbol{\sigma}) + \langle\boldsymbol{y}_t,\boldsymbol{\sigma}\rangle.
\end{aligned} \tag{6.38}$$

where we recall $\widetilde{H}_N(\boldsymbol{\sigma}) \sim \mu_{\mathsf{null}}$. The tilted measure $\mu_t = \mu_{H_N,\boldsymbol{y}_t}$ defined in (6.11) has the form

$$\mu_t(\mathsf{d}\boldsymbol{\sigma}) = \frac{1}{Z}\exp H_{N,t}(\boldsymbol{\sigma})\ \mu_0(\mathsf{d}\boldsymbol{\sigma}).$$

Let $\boldsymbol{m}_t$ be the mean of μ_t. The main result of our analysis is the following.

Theorem 6.4.1. *Under condition (6.5), there exist parameters $(K_{\mathsf{AMP}}, K^*_{\mathsf{GD}}, \eta)$ depending only on (ξ, t) such that the point $\boldsymbol{m}^{\mathsf{alg}}$ output by Algorithm 1 on input $(H_N, \boldsymbol{y}_t)$, with parameters K_{AMP}, $K_{\mathsf{GD}}(N) = K^*_{\mathsf{GD}}\log N$, η satisfies*

$$\mathbb{E}\,\|\boldsymbol{m}^{\mathsf{alg}} - \boldsymbol{m}_t\|_N^2 = o(N^{-1}).$$

Recall that we defined $\xi_t(q) = \xi(q) + tq$.

Fact 6.4.2. *For any $t \in [0,\infty)$, there is a unique solution $q_* = q_*(t) \in [0,1)$ to*

$$\xi_t'(q) = \frac{q}{1-q}. \tag{6.39}$$

Proof. Define $f(q) = \xi_t'(q) - \frac{q}{1-q}$. Since $f(0) = t > 0$ and $\lim_{q\to1^-} f(q) = -\infty$, there is at least one solution. As

$$\frac{\mathsf{d}}{\mathsf{d}q}\left(\xi_t'(q) - \frac{q}{1-q}\right) = \xi''(q) - \frac{1}{(1-q)^2} \overset{(6.5)}{<} 0,$$

this solution is unique. □

Henceforth let q_* denote this solution. It will also be useful to rewrite (6.38) as

$$H_{N,t}(\boldsymbol{\sigma}) = N\xi_t(\langle \boldsymbol{x}, \boldsymbol{\sigma}\rangle_N) + \widetilde{H}_{N,t}(\boldsymbol{\sigma}), \tag{6.40}$$

where

$$\widetilde{H}_{N,t}(\boldsymbol{\sigma}) = \widetilde{H}_N(\boldsymbol{\sigma}) + \langle \boldsymbol{B}_t, \boldsymbol{\sigma}\rangle \tag{6.41}$$

is a spin glass with mixture ξ_t. In the proofs below, we will switch between these two representations of $H_{N,t}$ as convenient.

The first step of our analysis characterizes the limiting performance of the AMP iteration (6.13), on $(H_N, \boldsymbol{y}_t)$ generated from the planted process (6.33). Recall the TAP free energy $\mathcal{F}_{\mathsf{TAP}}$ introduced in (6.15). With the notation (6.40), we can write

$$\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m}) = N\xi_t(\langle \boldsymbol{x}, \boldsymbol{m}\rangle_N) + \widetilde{H}_{N,t}(\boldsymbol{m}) + \frac{N}{2}\theta(\|\boldsymbol{m}\|_N^2) + \frac{N}{2}\log(1 - \|\boldsymbol{m}\|_N^2).$$

Proposition 6.4.3. *For any $\iota > 0$, there exists $k_0 \in \mathbb{N}$, depending only on (ξ, t, ι), such that for any fixed k, $k \geq k_0$ the following holds with probability $1 - e^{-cN}$. The AMP iterate $\boldsymbol{m}^k$ satisfies*

$$|\langle \boldsymbol{x}, \boldsymbol{m}^k\rangle_N - q_*|, |\langle \boldsymbol{m}^k, \boldsymbol{m}^k\rangle_N - q_*| \leq \iota \tag{6.42}$$

and

$$\left\|\nabla \mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m}^k)\right\|_N, \left\|\nabla \widetilde{H}_{N,t}(\boldsymbol{m}^k) + \xi_t'(q_*)\boldsymbol{x} - \left((1-q_*)\xi''(q_*) + \frac{1}{1-q_*}\right)\boldsymbol{m}^k\right\|_N \leq \iota. \tag{6.43}$$

Moreover, with $I = I(\iota) = [q_* - \iota, q_* + \iota]$,

$$\mu_t(\mathsf{Band}(\boldsymbol{m}^k, I) \cap \mathsf{Band}(\boldsymbol{x}, I)) \geq 1 - e^{-cN}. \tag{6.44}$$

The proof of this proposition is presented in Section 6.5. For $\iota > 0$, define

$$\mathcal{S}_\iota := \left\{\boldsymbol{m} \in \mathbb{R}^N : |\langle \boldsymbol{m}, \boldsymbol{x}\rangle_N - q_*|, |\langle \boldsymbol{m}, \boldsymbol{m}\rangle_N - q_*| \leq \iota\right\}. \tag{6.45}$$

Proposition 6.4.4. *There exist $C_{\max}^{\mathsf{spec}} > C_{\min}^{\mathsf{spec}} > 0$ and $L > 0$ such that, for any sufficiently small $\iota > 0$, there is an event $\mathcal{E}_0$ with probability $1 - e^{-cN}$, on which the following holds.*

(a) The event K_N from Proposition 6.3.6 holds.

(b) $\mathcal{F}_{\mathsf{TAP}}$ has a unique critical point $\boldsymbol{m}^{\mathsf{TAP}}$ in $\mathcal{S}_\iota$, which further satisfies

$$\mathsf{spec}(\nabla^2 \mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m}^{\mathsf{TAP}})) \subseteq [-C_{\max}^{\mathsf{spec}}, -C_{\min}^{\mathsf{spec}}]. \tag{6.46}$$

(c) For K_{AMP} large enough (depending on ι), we have $\boldsymbol{m}^{\mathsf{AMP}} \in \mathcal{S}_{\iota/2}$ and $\|\boldsymbol{m}^{\mathsf{AMP}} - \boldsymbol{m}^{\mathsf{TAP}}\|_N \leq \iota/2$.

Note that under (a), there exists c_{op} such that $\left\|\nabla^2 \mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m})\right\|_{\mathsf{op},N}, \left\|\nabla^3 \mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m})\right\|_{\mathsf{op},N} \leq c_{\mathsf{op}}$ uniformly over $\boldsymbol{m} \in \mathcal{S}_\iota$, for all sufficiently small $\iota > 0$. Let

$$\varepsilon = \min\left(\frac{\iota c_{\mathsf{op}}}{10}, \frac{(C_{\min}^{\mathsf{spec}})^2}{40 c_{\mathsf{op}}}\right). \tag{6.47}$$

Let $\mathcal{E} = \mathcal{E}_0 \cap \{\|\nabla \mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m}^{\mathsf{AMP}})\|_N \leq \varepsilon\}$. (For K_{AMP} large enough, this holds with probability $1 - e^{-cN}$ by Proposition 6.4.3.) We further have:

(d) For any $\delta > 0$ there exists $C_\delta > 0$ such that the following holds. For any random variable X with $0 \leq X \leq 1$ almost surely,

$$\mathbb{E}[X\mathbf{1}\{\mathcal{E}\}] \leq C_\delta \sup_{\boldsymbol{m} \in \mathcal{S}_\iota} \mathbb{E}\left[X^{1+\delta}\mathbf{1}\{\mathcal{E}\}\middle|\nabla \mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m}) = \mathbf{0}\right]^{1/(1+\delta)}.$$

Proposition 6.4.5. *For sufficiently small $\iota > 0$, with probability $1 - e^{-cN}$, the event $\mathcal{E}$ from Proposition 6.4.4 holds and:*

(a) For $I = I(\iota)$ as above, we have

$$\mu_t(\mathsf{Band}(\boldsymbol{m}^{\text{TAP}}, I) \cap \mathsf{Band}(\boldsymbol{x}, I)) \geq 1 - e^{-cN}.$$

(b) For η small enough and K^*_{GD} large enough, we have $\|\boldsymbol{m}^{\mathsf{GD}} - \boldsymbol{m}^{\text{TAP}}\|_N \leq N^{-10}$.

(c) For any $\boldsymbol{m}_1, \boldsymbol{m}_2 \in \mathsf{B}_N(\boldsymbol{m}^{\text{TAP}}, \iota)$, we have $\|\boldsymbol{\Delta}(\boldsymbol{m}_1) - \boldsymbol{\Delta}(\boldsymbol{m}_2)\|_N \leq \frac{L}{N}\|\boldsymbol{m}_1 - \boldsymbol{m}_2\|_N$.

The proofs of the last two propositions are given in Section 6.6.

For $\iota > 0$, define the truncated magnetization

$$\widetilde{\boldsymbol{m}}_\iota(\boldsymbol{m}) = \frac{\int_{\mathsf{Band}(\boldsymbol{m}, I(\iota)) \cap \mathsf{Band}(\boldsymbol{x}, I(\iota))} \boldsymbol{\sigma} \exp(H_{N,t}(\boldsymbol{\sigma}))\ \mu_0(\mathrm{d}\boldsymbol{\sigma})}{\int_{\mathsf{Band}(\boldsymbol{m}, I(\iota)) \cap \mathsf{Band}(\boldsymbol{x}, I(\iota))} \exp(H_{N,t}(\boldsymbol{\sigma}))\ \mu_0(\mathrm{d}\boldsymbol{\sigma})}.$$

Proposition 6.4.6. *Let $\boldsymbol{\Delta}(\,\cdot\,)$ be defined as in Section 6.2.3. Then, for sufficiently small $\iota, \delta > 0$, we have*

$$\sup_{\boldsymbol{m} \in \mathcal{S}_\iota} \mathbb{E}\left[\|\boldsymbol{m} + \boldsymbol{\Delta}(\boldsymbol{m}) - \widetilde{\boldsymbol{m}}_{2\iota}(\boldsymbol{m})\|_N^{2+\delta} \middle| \nabla \mathcal{F}_{\text{TAP}}(\boldsymbol{m}) = \boldsymbol{0}\right] \leq N^{-(1+\delta)}.$$

The proof of this proposition is given in Section 6.7.

Proof of Theorem 6.4.1. Let $\mathcal{E}_1$ be the intersection of $\mathcal{E}$ from Proposition 6.4.4 and the event in Proposition 6.4.5. On $\mathcal{E}_1$, the point $\boldsymbol{m}^{\text{TAP}}$ is well-defined and we can write

$$\begin{aligned}
\boldsymbol{m}^{\mathsf{alg}} - \boldsymbol{m}_t &= \boldsymbol{m}^{\mathsf{GD}} + \boldsymbol{\Delta}(\boldsymbol{m}^{\mathsf{GD}}) - \boldsymbol{m}_t \\
&= (\boldsymbol{m}^{\mathsf{GD}} - \boldsymbol{m}^{\text{TAP}}) + (\boldsymbol{\Delta}(\boldsymbol{m}^{\mathsf{GD}}) - \boldsymbol{\Delta}(\boldsymbol{m}^{\text{TAP}})) + (\widetilde{\boldsymbol{m}}_{2\iota}(\boldsymbol{m}^{\text{TAP}}) - \boldsymbol{m}_t) \\
&\quad + (\boldsymbol{m}^{\text{TAP}} + \boldsymbol{\Delta}(\boldsymbol{m}^{\text{TAP}}) - \widetilde{\boldsymbol{m}}_{2\iota}(\boldsymbol{m}^{\text{TAP}}))\,,
\end{aligned}$$

whence

$$\begin{aligned}
\|\boldsymbol{m}^{\mathsf{alg}} - \boldsymbol{m}_t\|_N^2 \leq{}& 4\|\boldsymbol{m}^{\mathsf{GD}} - \boldsymbol{m}^{\text{TAP}}\|_N^2 + 4\|\boldsymbol{\Delta}(\boldsymbol{m}^{\mathsf{GD}}) - \boldsymbol{\Delta}(\boldsymbol{m}^{\text{TAP}})\|_N^2 + 4\|\widetilde{\boldsymbol{m}}_{2\iota}(\boldsymbol{m}^{\text{TAP}}) - \boldsymbol{m}_t\|_N^2 \\
&+ 4\|\boldsymbol{m}^{\text{TAP}} + \boldsymbol{\Delta}(\boldsymbol{m}^{\text{TAP}}) - \widetilde{\boldsymbol{m}}_{2\iota}(\boldsymbol{m}^{\text{TAP}})\|_N^2
\end{aligned}$$

The following also holds on $\mathcal{E}_1$. By Proposition 6.4.5(b) and 6.4.5(c), for some constant C (changing from line to line below),

$$\|\boldsymbol{m}^{\mathsf{GD}} - \boldsymbol{m}^{\text{TAP}}\|_N^2, \|\boldsymbol{\Delta}(\boldsymbol{m}^{\mathsf{GD}}) - \boldsymbol{\Delta}(\boldsymbol{m}^{\text{TAP}})\|_N^2 \leq CN^{-20}.$$

By Proposition 6.4.4(a), the complement of $\mathsf{Band}(\boldsymbol{m}^{\text{TAP}}, I) \cap \mathsf{Band}(\boldsymbol{x}, I)$ accounts for a e^{-cN} fraction of the Gibbs measure. Because the spins $\boldsymbol{\sigma}$ are bounded, this implies

$$\|\widetilde{\boldsymbol{m}}_{2\iota}(\boldsymbol{m}^{\text{TAP}}) - \boldsymbol{m}_t\|_N^2 \leq e^{-cN}.$$

Therefore, on $\mathcal{E}_1$, for all sufficiently large N

$$\|\boldsymbol{m}^{\mathsf{alg}} - \boldsymbol{m}_t\|_N^2 \leq CN^{-20} + 4\|\boldsymbol{m}^{\text{TAP}} + \boldsymbol{\Delta}(\boldsymbol{m}^{\text{TAP}}) - \widetilde{\boldsymbol{m}}_{2\iota}(\boldsymbol{m}^{\text{TAP}})\|_N^2.$$

Thus

$$\begin{aligned}
\mathbb{E}[\|\boldsymbol{m}^{\mathsf{alg}} - \boldsymbol{m}_t\|_N^2] &\leq \mathbb{P}(\mathcal{E}_1^c) + \mathbb{E}[\|\boldsymbol{m}^{\mathsf{alg}} - \boldsymbol{m}_t\|_N^2 \mathbf{1}\{\mathcal{E}_1\}] \\
&\leq CN^{-20} + 4\,\mathbb{E}\left[\|\boldsymbol{m}^{\text{TAP}} + \boldsymbol{\Delta}(\boldsymbol{m}^{\text{TAP}}) - \widetilde{\boldsymbol{m}}_{2\iota}(\boldsymbol{m}^{\text{TAP}})\|_N^2 \mathbf{1}\{\mathcal{E}_1\}\right] \\
&\leq CN^{-20} + 4C_{\delta/2} \sup_{\boldsymbol{m} \in \mathcal{S}_\iota} \mathbb{E}\left[\|\boldsymbol{m} + \boldsymbol{\Delta}(\boldsymbol{m}) - \widetilde{\boldsymbol{m}}_{2\iota}(\boldsymbol{m})\|_N^{2+\delta} \mathbf{1}\{\mathcal{E}_1\} \middle| \nabla \mathcal{F}_{\text{TAP}}(\boldsymbol{m}) = \boldsymbol{0}\right]^{1/(1+\delta/2)} \\
&\leq CN^{-20} + 4C_{\delta/2} N^{-(1+\delta)/(1+\delta/2)} = o(N^{-1}).
\end{aligned}$$

In the second-last line, we applied Proposition 6.4.4(d), noting that on $\mathcal{E}_1$ and conditioned on $\nabla \mathcal{F}_{\text{TAP}}(\boldsymbol{m}) = \boldsymbol{0}$, we have $\boldsymbol{m}^{\text{TAP}} = \boldsymbol{m}$ almost surely. The last line is Proposition 6.4.6.

□

6.5 Analysis of AMP iteration: proof of Proposition 6.4.3

6.5.1 State evolution limit

We first prove (6.42) and (6.43) using the state evolution result of [Bol14, BM11, JM13]. Recalling the change of notation (6.40), the AMP iteration (6.13) can be rewritten as $\boldsymbol{m}^{-1} = \boldsymbol{w}^0 = \boldsymbol{0}$,

$$\begin{aligned}
\boldsymbol{m}^k &= (1-q_k)\boldsymbol{w}^k, \\
\boldsymbol{w}^{k+1} &= \nabla H_{N,t}(\boldsymbol{m}^k) - (1-q_k)\xi''(q_k)\boldsymbol{m}^{k-1} \\
&= \nabla \widetilde{H}_{N,t}(\boldsymbol{m}^k) + \xi_t'(\langle \boldsymbol{x}, \boldsymbol{m}^k\rangle_N)\boldsymbol{x} - (1-q_k)\xi''(q_k)\boldsymbol{m}^{k-1}.
\end{aligned} \tag{6.48}$$

Here and below, the sequence $(q_k)_{k\geq 0}$ is defined as per Eq. (6.12).

Set $\gamma_0 = \Sigma_{0,i} = \Sigma_{i,0} = 0$ for all $i \geq 0$, and define the following recurrence. Sample $X \sim \mathcal{N}(0,1)$ and, for $k \geq 0$,

$$(G_1, \dots, G_k) \sim \mathcal{N}(0, \Sigma_{\leq k}), \qquad W_i = G_i + \gamma_i X.$$

Then, let

$$\gamma_{k+1} = \xi_t'((1-q_k)\gamma_k) \tag{6.49}$$

$$\Sigma_{k+1,j+1} = \xi_t'\left((1-q_k)(1-q_j)\,\mathbb{E}[W_k W_j]\right). \tag{6.50}$$

The following proposition is an immediate consequence of [AMS21, Proposition 3.1], which generalizes to the tensor case [BM11, Theorem 1].

Proposition 6.5.1. *For any $k \geq 0$, the empirical distribution of the AMP iterates' coordinates converges in W_2 in probability:*

$$\frac{1}{N}\sum_{i=1}^N \delta_{x_i, w_i^1, \dots, w_i^k} \xrightarrow{W_2} \mathcal{L}(X, W_1, \dots, W_k).$$

(In words, the left-hand side is the probability distribution on $\mathbb{R}^{k+1}$ that puts mass $1/N$ on each point $(x_i, w_i^1, \dots, w_i^k)$, for $i \in [N]$.)

Lemma 6.5.2. *For all $k, j \geq 0$, we have $\Sigma_{k,j} = \gamma_{k\wedge j} = \frac{q_{k\wedge j}}{1-q_{k\wedge j}}$.*

Proof. We first prove by induction that $\gamma_k = \frac{q_k}{1-q_k}$. For $k=0$ this is clear, and then by induction

$$\gamma_{k+1} = \xi_t'(q_k) = \frac{q_{k+1}}{1-q_{k+1}}.$$

Similarly, by induction

$$(1-q_k)(1-q_j)\,\mathbb{E}[W_k W_j] = (1-q_k)(1-q_j)\left(\Sigma_{k,j} + \gamma_k\gamma_j\right) = (1-q_{k\vee j})q_{k\wedge j} + q_k q_j = q_{k\wedge j},$$

and thus

$$\Sigma_{k+1,j+1} = \xi_t'(q_{k\wedge j}) = \frac{q_{k\wedge j+1}}{1-q_{k\wedge j+1}}.$$

□

Lemma 6.5.3. *As $k \to \infty$, we have $q_k \to q_*$.*

Proof. Since the function $f(q) = \frac{\xi_t'(q)}{1+\xi_t'(q)}$ is increasing, with $f(0) > 0$, $f(1) < 1$, q_k must converge to a solution of $q = f(q)$. This rearranges to $\xi_t'(q) = \frac{q}{1-q}$, which has unique solution q_* by Fact 6.4.2. □

Proposition 6.5.4. *With probability $1-e^{-cN}$, (6.42) and (6.43) hold for all $k \geq k_0$.*

Proof. Let $\simeq$ denote equality up to an additive error $o_{P,N}(1)$ (a term vanishing in probability as $N \to \infty$). By Proposition 6.5.1,

$$\langle \boldsymbol{x}, \boldsymbol{m}^k \rangle_N = (1-q_k)\langle \boldsymbol{x}, \boldsymbol{w}^k \rangle_N \simeq (1-q_k)\gamma_k = q_k. \tag{6.51}$$

Moreover,

$$\langle \boldsymbol{m}^k, \boldsymbol{m}^k \rangle_N = (1-q_k)^2\langle \boldsymbol{w}^k, \boldsymbol{w}^k \rangle_N \simeq (1-q_k)^2\left(\Sigma_{k,k} + \gamma_k^2\right) = q_k. \tag{6.52}$$

By Lemma 6.5.3, for all k large enough we have $|q_k - q_*| \le \iota/3$, whence (6.42) holds with high probability. Rearranging the AMP iteration gives

$$\begin{aligned}\nabla \widetilde{H}_{N,t}(\boldsymbol{m}^k) &= -\xi_t'(\langle \boldsymbol{x}, \boldsymbol{m}^k \rangle_N)\boldsymbol{x} + \boldsymbol{w}^{k+1} + (1-q_k)\xi''(q_k)\boldsymbol{m}^{k-1} \\ &= -\xi_t'(\langle \boldsymbol{x}, \boldsymbol{m}^k \rangle_N)\boldsymbol{x} + \frac{1}{1-q_{k+1}}\boldsymbol{m}^{k+1} + (1-q_k)\xi''(q_k)\boldsymbol{m}^{k-1},\end{aligned} \tag{6.53}$$

By Proposition 6.5.1, Lemma 6.5.2, and Lemma 6.5.3, we have

$$\lim_{k\to\infty} \operatorname*{p-lim}_{N\to\infty} \|\boldsymbol{m}^{k+1} - \boldsymbol{m}^k\| = 0\,, \tag{6.54}$$

$$\lim_{k\to\infty} \operatorname*{p-lim}_{N\to\infty} \|\boldsymbol{w}^{k+1} - \boldsymbol{w}^k\| = 0\,. \tag{6.55}$$

and therefore, by Eq. (6.53),

$$\lim_{k\to\infty} \operatorname*{p-lim}_{N\to\infty} \left\|\nabla \widetilde{H}_{N,t}(\boldsymbol{m}^k) + \xi_t'(q_*)\boldsymbol{x} + \left(\frac{1}{1-q_*} + (1-q_*)\xi''(q_*)\right)\boldsymbol{m}^k\right\|_N = 0\,.$$

As

$$\nabla \mathcal{F}_{\textsf{TAP}}(\boldsymbol{m}) = \nabla \widetilde{H}_{N,t}(\boldsymbol{m}) + \xi_t'(\langle \boldsymbol{x}, \boldsymbol{m} \rangle_N)\boldsymbol{x} + \left(\frac{1}{1-\|\boldsymbol{m}\|_N^2} + (1-\|\boldsymbol{m}\|_N^2)\xi''(\|\boldsymbol{m}\|_N^2)\right)\boldsymbol{m}, \tag{6.56}$$

equations (6.51), (6.52) further imply

$$\lim_{k\to\infty} \operatorname*{p-lim}_{N\to\infty} \|\nabla \mathcal{F}_{\textsf{TAP}}(\boldsymbol{m}^k)\|_N = 0\,.$$

Thus, for large enough k, (6.43) holds with high probability.

To improve these assertions to $1 - e^{-cN}$ probability, note that by [HS25, Section 8], the AMP iterate $\boldsymbol{m}^k$ is, on an event $\mathcal{E}_{\textsf{Lip}}$ with probability $1 - e^{-cN}$, a $O(1)$-Lipschitz function of the disorder Gaussians in $\widetilde{H}_{N,t}$. By Kirszbraun's extension theorem, there is a measurable, $O(1)$-Lipschitz function $\widetilde{\boldsymbol{m}}^k$ of the disorder which agrees with $\boldsymbol{m}^k$ on $\mathcal{E}_{\textsf{Lip}}$. Thus $\langle \boldsymbol{x}, \widetilde{\boldsymbol{m}}^k \rangle_N$ and $\langle \widetilde{\boldsymbol{m}}^k, \widetilde{\boldsymbol{m}}^k \rangle_N$ are $O(N^{-1/2})$-Lipschitz in the disorder. By Gaussian concentration of measure

$$|\langle \boldsymbol{x}, \widetilde{\boldsymbol{m}}^k \rangle_N - \mathbb{E}\langle \boldsymbol{x}, \widetilde{\boldsymbol{m}}^k \rangle_N|, |\langle \widetilde{\boldsymbol{m}}^k, \widetilde{\boldsymbol{m}}^k \rangle_N - \mathbb{E}\langle \widetilde{\boldsymbol{m}}^k, \widetilde{\boldsymbol{m}}^k \rangle_N| \le \iota/3$$

with probability $1 - e^{-cN}$. Since $\boldsymbol{m}^k = \widetilde{\boldsymbol{m}}^k$ on $\mathcal{E}_{\textsf{Lip}}$, (6.42) holds with probability $1 - e^{-cN}$.

By Proposition 6.3.6, $\boldsymbol{m} \mapsto \nabla \widetilde{H}_{N,t}(\boldsymbol{m})$ is also $O(1)$-Lipschitz over $\|\boldsymbol{m}\|_N \le 1$ with probability $1 - e^{-cN}$. A similar argument shows that (6.43) holds with probability $1 - e^{-cN}$. □

6.5.2 Overlap with AMP iterates

The following proposition constitutes the first half of the proof of Eq. (6.44).

Proposition 6.5.5. *Let $\iota > 0$ and $I = I(\iota)$. With probability $1 - e^{-cN}$, for all $k \ge k_0$ (with k_0 a sufficiently large constant depending on (ξ, t, ι)),*

$$\mu_t(\textsf{Band}(\boldsymbol{m}^k, I)) \ge 1 - e^{-cN}.$$

To prove Proposition 6.5.5, we will combine Lemma 6.5.7 below, which identifies a band on which the Gibbs measure μ_t concentrates, with a self-reduction argument. We return to the earlier representation (6.38) of $H_{N,t}$, which we reproduce below.

$$
\begin{aligned}
H_{N,t}(\boldsymbol{\sigma}) &= H_N(\boldsymbol{\sigma}) + \langle \boldsymbol{y}_t, \boldsymbol{\sigma}\rangle, \qquad \text{where} \\
H_N(\boldsymbol{\sigma}) &= N\xi(\langle \boldsymbol{x}, \boldsymbol{\sigma}\rangle_N) + \widetilde{H}_N(\boldsymbol{\sigma}), \\
\boldsymbol{y}_t &= t\boldsymbol{x} + \sqrt{t}\boldsymbol{g}, \qquad \boldsymbol{g} \sim \mathcal{N}(0, \boldsymbol{I}_N).
\end{aligned}
$$

Let $\langle \cdot \rangle$ denote average with respect to $\boldsymbol{\sigma} \sim \mu_t$. The following fact is a restatement of Bayes theorem: sampling $\boldsymbol{x}$ and then $\boldsymbol{y}_t$ is equivalent to sampling $\boldsymbol{y}_t$ and then $\boldsymbol{x}$ from the posterior. In the context of statistical physics, this is known as 'Nishimori's property.'

Fact 6.5.6. *For any bounded measurable f, $\mathbb{E}\, f(\boldsymbol{x}, \boldsymbol{y}_t) = \mathbb{E}\langle f(\boldsymbol{\sigma}, \boldsymbol{y}_t)\rangle$.*

Lemma 6.5.7. *Let $\iota > 0$ be arbitrary. With probability $1 - e^{-cN}$,*

$$
\Big| \|\boldsymbol{y}_t\|_N^2 - t^2 - t \Big| \le \iota, \quad |\langle \boldsymbol{x}, \boldsymbol{y}_t\rangle_N - t| \le \iota, \quad \mu_t(\mathsf{Band}(\boldsymbol{y}_t, [t-\iota, t+\iota])) \ge 1 - e^{-cN}.
$$

Proof. Clearly $\|\boldsymbol{y}_t\|_N^2 \simeq t^2 + t$ and $\langle \boldsymbol{x}, \boldsymbol{y}_t\rangle_N \simeq t$, so the first two conclusions follow by standard concentration arguments. By Fact 6.5.6,

$$
\mathbb{E}\left\langle \mathbf{1}\left\{\langle \boldsymbol{\sigma}, \boldsymbol{y}_t\rangle_N \notin [t-\iota, t+\iota]\right\}\right\rangle = \mathbb{P}\left(\langle \boldsymbol{x}, \boldsymbol{y}_t\rangle_N \notin [t-\iota, t+\iota]\right) \le e^{-cN}.
$$

By Markov's inequality,

$$
\mathbb{P}\left\{ \left\langle \mathbf{1}\left\{\langle \boldsymbol{\sigma}, \boldsymbol{y}_t\rangle_N \notin [t-\iota, t+\iota]\right\}\right\rangle \ge e^{-cN/2} \right\} \le e^{-cN/2}.
$$

This implies the final conclusion after adjusting c. $\square$

We next introduce a self-reduction property of models obtained by restriction to a certain band. Define

$$
U = \left\{ \boldsymbol{\sigma} \in \mathbb{R}^N : \langle \boldsymbol{\sigma}, \boldsymbol{y}_t\rangle_N = 0 \right\}.
$$

Recall that $(q_k)_{k\ge 0}$ is defined by Eq. (6.12), and in particular $q_1 = t/(1+t)$. Let $\widehat{\boldsymbol{y}}_t = \boldsymbol{y}_t/\|\boldsymbol{y}_t\|_N$ and $r = \sqrt{q_1}$. Consider the Hamiltonian on $\boldsymbol{\rho} \in U$ defined by

$$
\widehat{H}(\boldsymbol{\rho}) = H_N(r\widehat{\boldsymbol{y}}_t + \sqrt{1-r^2}\boldsymbol{\rho}) - H_N(r\widehat{\boldsymbol{y}}_t).
$$

Further define

$$
\xi_{(1)}(s) = \xi(q_1 + (1-q_1)s) - \xi(q_1).
$$

Let $r_1 = \langle \boldsymbol{x}, \widehat{\boldsymbol{y}}_t\rangle_N$ and define $\boldsymbol{x}^\perp \in U$ by $\boldsymbol{x} = r_1\widehat{\boldsymbol{y}}_t + \sqrt{1-r_1^2}\boldsymbol{x}^\perp$. Note that conditionally on $(\boldsymbol{y}_t, r_1)$, $\boldsymbol{x}^\perp$ is a uniformly random vector in $U \cap S_N$. Also define the Hamiltonian

$$
\widehat{H}'(\boldsymbol{\rho}) = N\xi_{(1)}(\langle \boldsymbol{x}^\perp, \boldsymbol{\rho}\rangle_N) + \widetilde{H}'(\boldsymbol{\rho}),
$$

where $\widetilde{H}'$ is a Gaussian process on U with covariance

$$
\mathbb{E}\, \widetilde{H}'(\boldsymbol{\rho}^1)\widetilde{H}'(\boldsymbol{\rho}^2) = N\xi_{(1)}(\langle \boldsymbol{\rho}^1, \boldsymbol{\rho}^2\rangle_N).
$$

Note that $\widehat{H}'$ is of the form (6.40), with one fewer dimension and $\xi_{(1)}$ in place of ξ_t.

Proposition 6.5.8 (Self-reduction)**.** *There exists a constant C such that the following holds. Let $\iota > 0$. Let S be the $(\boldsymbol{y}_t, r_1)$-measurable event*

$$
\Big| \|\boldsymbol{y}_t\|_N - \sqrt{t(1+t)} \Big|, |\langle \boldsymbol{x}, \widehat{\boldsymbol{y}}_t\rangle_N - \sqrt{q_1}| \le \iota. \tag{6.57}
$$

Then $\mathbb{P}(S) \geq 1 - e^{-cN}$ *and for any* $(\boldsymbol{y}_t, r_1) \in S$ *the following holds. There is a coupling* $\mathcal{C}$ *of* $\mathcal{L}(\widehat{H}|\boldsymbol{y}_t, r_1)$ *and* $\mathcal{L}(\widehat{H}')$ *such that almost surely,*

$$\begin{aligned} \frac{1}{N} \sup_{\boldsymbol{\rho} \in U \cap S_N} |\widehat{H}(\boldsymbol{\rho}) - \widehat{H}'(\boldsymbol{\rho})| &\leq C\iota\,, \\ \sup_{\boldsymbol{\rho} \in U \cap S_N} \left\| \nabla_U \widehat{H}(\boldsymbol{\rho}) - \nabla_U \widehat{H}'(\boldsymbol{\rho}) \right\|_N &\leq C\iota. \end{aligned} \tag{6.58}$$

Proof. Suppose the event in Lemma 6.5.7 holds. Then, using $q_1 = t/(1+t)$,

$$r_1 = \frac{\langle \boldsymbol{x}, \boldsymbol{y}_t \rangle_N}{\|\boldsymbol{y}_t\|_N} = \frac{t + O(\iota)}{\sqrt{t(1+t)} + O(\iota)} = \sqrt{q_1} + O(\iota).$$

This proves $\mathbb{P}(S) \geq 1 - e^{-cN}$, after adjusting ι by a constant factor. Now suppose $(\boldsymbol{y}_t, r_1) \in S$. We have $\widehat{H}(\boldsymbol{\rho}) = \widehat{H}_1(\boldsymbol{\rho}) + \widehat{H}_2(\boldsymbol{\rho})$, where

$$\begin{aligned} \widehat{H}_1(\boldsymbol{\rho}) &= N \left\{ \xi \left(R \left(r\widehat{\boldsymbol{y}}_t + \sqrt{1-r^2}\boldsymbol{\rho}, r_1\widehat{\boldsymbol{y}}_t + \sqrt{1-r_1^2}\boldsymbol{x}^\perp \right) \right) - \xi \left(R \left(r\widehat{\boldsymbol{y}}_t, r_1\widehat{\boldsymbol{y}}_t + \sqrt{1-r_1^2}\boldsymbol{x}^\perp \right) \right) \right\}, \\ \widehat{H}_2(\boldsymbol{\rho}) &= \left\{ \widetilde{H}_N \left(r\widehat{\boldsymbol{y}}_t + \sqrt{1-r^2}\boldsymbol{\rho} \right) - \widetilde{H}_N(r\widehat{\boldsymbol{y}}_t) \right\}. \end{aligned}$$

The first summand simplifies as

$$\widehat{H}_1(\boldsymbol{\rho}) = N \left\{ \xi \left(rr_1 + \sqrt{(1-r^2)(1-r_1^2)} \langle \boldsymbol{\rho}, \boldsymbol{x}^\perp \rangle_N \right) - \xi(rr_1) \right\} = N\xi_{(1)}(\langle \boldsymbol{\rho}, \boldsymbol{x}^\perp \rangle_N) + N \cdot O(\iota).$$

The second summand is a Gaussian process on U with covariance

$$\mathbb{E}\, \widehat{H}_2(\boldsymbol{\rho}^1)\widehat{H}_2(\boldsymbol{\rho}^2) = N \left(\xi(r^2 + (1-r^2)\langle \boldsymbol{\rho}^1, \boldsymbol{\rho}^2 \rangle_N) - \xi(r^2) \right) = N\xi_{(1)}(\langle \boldsymbol{\rho}^1, \boldsymbol{\rho}^2 \rangle_N).$$

Thus we can couple $\widehat{H}_2$ and $\widetilde{H}'$ so that $\widehat{H}_2 = \widetilde{H}'$ almost surely. □

Define $\widehat{q}_0 = 0$ and, similarly to (6.12),

$$\widehat{q}_{k+1} = \frac{\xi'_{(1)}(\widehat{q}_k)}{1 + \xi'_{(1)}(\widehat{q}_k)}.$$

Lemma 6.5.9. *For all* $k \geq 0$*, we have* $q_1 + (1-q_1)\widehat{q}_k = q_{k+1}$*.*

Proof. We induct on k. The base case $k = 0$ is trivial. Recalling $q_1 = \frac{t}{1+t}$, the inductive step follows from

$$\begin{aligned} q_1 + (1-q_1)\widehat{q}_{k+1} &= q_1 + (1-q_1)\frac{\xi'_{(1)}(\widehat{q}_k)}{1+\xi'_{(1)}(\widehat{q}_k)} = 1 - (1-q_1)\left(1 - \frac{\xi'_{(1)}(\widehat{q}_k)}{1+\xi'_{(1)}(\widehat{q}_k)}\right) \\ &= 1 - \frac{1-q_1}{1+(1-q_1)\xi'(q_{k+1})} = 1 - \frac{1}{1+t+\xi'(q_{k+1})} \\ &= \frac{\xi'_t(q_{k+1})}{1+\xi'_t(q_{k+1})} = q_{k+2}. \end{aligned}$$

□

Define the AMP iteration, analogous to (6.48), on the reduced model $\widehat{H}'$, by $\widehat{\boldsymbol{m}}^{-1} = \widehat{\boldsymbol{w}}^0 = \boldsymbol{0}$ and

$$\widehat{\boldsymbol{m}}^k = (1-\widehat{q}_k)\widehat{\boldsymbol{w}}^k, \qquad \widehat{\boldsymbol{w}}^{k+1} = \nabla_U \widehat{H}'(\widehat{\boldsymbol{m}}^k) - (1-\widehat{q}_k)\xi''_{(1)}(\widehat{q}_k)\widehat{\boldsymbol{m}}^{k-1}.$$

Note that $\widehat{\boldsymbol{m}}^k, \widehat{\boldsymbol{w}}^k \in U$.

Proposition 6.5.10 (Self-reduction of AMP iterates)**.** *Let $\iota > 0$. Suppose $(\boldsymbol{y}_t, r_1) \in S$ for S as in Proposition 6.5.8, and couple $\mathcal{L}(\widehat{H}|\boldsymbol{y}_t, r_1)$ and $\widehat{H}'$ as in that proposition. Then (conditionally on $\boldsymbol{y}_t, r_1$) with probability $1 - e^{-cN}$, for all $1 \le k \le O(1)$,*

$$\|\boldsymbol{m}^k - \widetilde{\boldsymbol{m}}^k\|_N \le O(\iota), \qquad \text{where} \qquad \widetilde{\boldsymbol{m}}^{k+1} = \sqrt{q_1}\widehat{\boldsymbol{y}}_t + \sqrt{1-q_1}\widehat{\boldsymbol{m}}^k. \tag{6.59}$$

Proof. We induct on the claim that (6.59) holds for all $1 \le k \le K$. First, we have

$$\boldsymbol{m}^1 = (1 - q_1)\boldsymbol{y}_t = \frac{\boldsymbol{y}_t}{1+t}, \qquad \widetilde{\boldsymbol{m}}^1 = \sqrt{q_1}\widehat{\boldsymbol{y}}_t = \sqrt{\frac{t}{1+t}}\widehat{\boldsymbol{y}}_t. \tag{6.60}$$

For $(\boldsymbol{y}_t, r_1) \in S$, we have $|\|\boldsymbol{y}_t\|_N - \sqrt{t(1+t)}| \le \iota$, and thus

$$\|\boldsymbol{m}^1 - \widetilde{\boldsymbol{m}}^1\|_N = \left|\frac{\sqrt{\|\boldsymbol{y}_t\|_N}}{1+t} - \sqrt{\frac{t}{1+t}}\right| \le \frac{\iota}{1+t}.$$

This proves the base case $K = 1$. Suppose (6.59) holds for $1 \le k \le K$. By Proposition 6.5.1, for all $1 \le j, k \le K+1$,

$$\langle \boldsymbol{m}^j, \boldsymbol{m}^k \rangle_N \to_p q_{j\wedge k}, \qquad \langle \widehat{\boldsymbol{m}}^j, \widehat{\boldsymbol{m}}^k \rangle_N \to_p \widehat{q}_{j\wedge k},$$

and thus, by Lemma 6.5.9,

$$\langle \widetilde{\boldsymbol{m}}^j, \widetilde{\boldsymbol{m}}^k \rangle_N \to_p q_1 + (1-q_1)\widehat{q}_{(j-1)\wedge(k-1)} = q_k.$$

Because AMP iterates are Lipschitz in the disorder (see the proof of Proposition 6.5.4), on an event with probability $1 - e^{-cN}$,

$$\langle \boldsymbol{m}^j, \boldsymbol{m}^k \rangle_N, \langle \widetilde{\boldsymbol{m}}^j, \widetilde{\boldsymbol{m}}^k \rangle_N \in [q_{j\wedge k} - \iota, q_{j\wedge k} + \iota] \tag{6.61}$$

for all $1 \le j, k \le K+1$. Since $\boldsymbol{m}^1$ is a multiple of $\boldsymbol{y}_t = \nabla H_{N,t}(\boldsymbol{0})$,

$$\boldsymbol{m}^{K+1} \in \mathsf{span}(\boldsymbol{m}^1, \dots, \boldsymbol{m}^K, \nabla H_{N,t}(\boldsymbol{m}^K)) = \mathsf{span}(\boldsymbol{m}^1, \dots, \boldsymbol{m}^K, \nabla_U H_N(\boldsymbol{m}^K)).$$

As

$$\widehat{\boldsymbol{m}}^K \in \mathsf{span}(\widehat{\boldsymbol{m}}^1, \dots, \widehat{\boldsymbol{m}}^{K-1}, \nabla_U \widehat{H}'(\widehat{\boldsymbol{m}}^{K-1})),$$

we have

$$\widetilde{\boldsymbol{m}}^{K+1} \in \mathsf{span}(\widetilde{\boldsymbol{m}}^1, \dots, \widetilde{\boldsymbol{m}}^K, \nabla_U \widehat{H}'(\widehat{\boldsymbol{m}}^{K-1})).$$

Note that $\sqrt{1-q_1}\nabla_U H_N(\widetilde{\boldsymbol{m}}^K) = \nabla_U \widehat{H}(\widehat{\boldsymbol{m}}^{K-1})$. Thus (on an event where ∇H_N is $O(1)$-Lipschitz, and the event in Proposition 6.5.8, both of which are probability $1 - e^{-cN}$)

$$\begin{aligned}\left\|\sqrt{1-q_1}\nabla_U H_N(\boldsymbol{m}^K) - \nabla_U \widehat{H}'(\widehat{\boldsymbol{m}}^{K-1})\right\|_N &\le \sqrt{1-q_1}\left\|\nabla_U H_N(\boldsymbol{m}^K) - \nabla_U H_N(\widetilde{\boldsymbol{m}}^K)\right\|_N \\ &\quad + \left\|\nabla_U \widehat{H}(\widehat{\boldsymbol{m}}^{K-1}) - \nabla_U \widehat{H}'(\widehat{\boldsymbol{m}}^{K-1})\right\|_N = O(\iota).\end{aligned}$$

This and (6.61) imply $\|\boldsymbol{m}^{K+1} - \widetilde{\boldsymbol{m}}^{K+1}\|_N = O(\iota)$, completing the induction. □

Proposition 6.5.11. *For all $\iota > 0$ and $k \ge 1$ fixed, the following holds. Let*

$$V_k(\iota) = \left\{\boldsymbol{\sigma} \in S_N : |\langle \boldsymbol{\sigma}, \boldsymbol{m}^j \rangle_N - q_j| \le \iota, \quad \forall 1 \le j \le k\right\}.$$

Then, with probability $1 - e^{-cN}$,

$$\mu_t(V_k(\iota)) \ge 1 - e^{-cN}.$$

Proof. We induct on k. By Lemma 6.5.7, with probability $1 - e^{-cN}$,

$$\mu_t(\mathsf{Band}(\boldsymbol{y}_t, [t - \iota, t + \iota])) \ge 1 - e^{-cN}. \tag{6.62}$$

As calculated in (6.60), $\boldsymbol{m}^1 = \boldsymbol{y}_t/(1+t)$, so $\boldsymbol{\sigma} \in \mathsf{Band}(\boldsymbol{y}_t, [t-\iota, t+\iota])$ if and only if

$$\langle \boldsymbol{\sigma}, \boldsymbol{m}^1 \rangle_N = \frac{t}{1+t} + O(\iota) = q_1 + O(\iota).$$

This proves the base case $k = 1$ after adjusting ι by a constant factor.

For the inductive step, let ι_1 be suitably small in ι. Let S_1 be the event (6.57) with right-hand side ι_1. By Proposition 6.5.8, $(\boldsymbol{y}_t, r_1) \in S_1$ with probability $1 - e^{-cN}$. Condition on any such $(\boldsymbol{y}_t, r_1)$. Along with (6.62), this implies

$$\mu_t(\mathsf{Band}(\widehat{\boldsymbol{y}}_t, [\sqrt{q_1} - C\iota_1, \sqrt{q_1} + C\iota_1])) \geq 1 - e^{-cN}$$

for suitable C. For $r_2 \in [\sqrt{q_1} - C\iota_1, \sqrt{q_1} + C\iota_1]$, let $\widehat{\mu}_t^{r_2}$ be the Gibbs measure on $U \cap S_N$ given by

$$\widehat{\mu}_t^{r_2} = Q_{\#}\mu_t(\cdot | \langle \boldsymbol{\sigma}, \widehat{\boldsymbol{y}}_t \rangle_N = r_2), \qquad \text{where} \qquad Q(\boldsymbol{\sigma}) = \frac{P^{\perp}_{\widehat{\boldsymbol{y}}_t}(\boldsymbol{\sigma})}{\|P^{\perp}_{\widehat{\boldsymbol{y}}_t}(\boldsymbol{\sigma})\|_N}.$$

Note that $\widehat{\mu}_t^{\sqrt{q_1}}$ is the Gibbs measure on $U \cap S_N$ corresponding to Hamiltonian $\widehat{H}$. Couple $\widehat{H}$ and $\widehat{H}'$ as in Proposition 6.5.8, and let $\widehat{\mu}_t'$ be the Gibbs measure on $U \cap S_N$ corresponding to Hamiltonian $\widehat{H}'$.

By the inductive hypothesis **applied to Hamiltonian $\widehat{H}'$ and mixture** $\xi_{(1)}$, with probability $1-e^{-cN}$, $\widehat{\mu}_t'(\widehat{V}_k(\iota)) \geq 1 - e^{-cN}$, where

$$\widehat{V}_k(\iota) = \left\{ \boldsymbol{\rho} \in U \cap S_N : |\langle \boldsymbol{\rho}, \widehat{\boldsymbol{m}}^j \rangle_N - \widehat{q}_j| \leq \iota, \quad \forall 1 \leq j \leq k \right\}.$$

By Proposition 6.5.8,

$$\frac{1}{N} \sup_{\boldsymbol{\rho} \in U \cap S_N} |\widehat{H}(\boldsymbol{\rho}) - \widehat{H}'(\boldsymbol{\rho})| \leq \iota_1.$$

For ι_1 small enough in ι, this implies

$$\widehat{\mu}_t^{\sqrt{q_1}}(\widehat{V}_k(2\iota)) \geq 1 - e^{-cN}.$$

By Lipschitz continuity of $H_{N,t}$, for ι_1 small enough in ι, we have

$$\widehat{\mu}_t^{r_2}(\widehat{V}_k(3\iota)) \geq 1 - e^{-cN}, \qquad \forall r_2 \in [\sqrt{q_1} - C\iota_1, \sqrt{q_1} + C\iota_1].$$

This implies $\mu_t(\widetilde{V}_{k+1}(4\iota)) \geq 1 - e^{-cN}$, where

$$\widetilde{V}_k(\iota) = \left\{ \boldsymbol{\sigma} \in S_N : |\langle \boldsymbol{\sigma}, \widetilde{\boldsymbol{m}}^j \rangle_N - q_j| \leq \iota, \quad \forall 1 \leq j \leq k \right\}.$$

However, by Proposition 6.5.10, with probability $1 - e^{-cN}$, $\|\boldsymbol{m}^j - \widetilde{\boldsymbol{m}}^j\|_N \leq \iota$ for all $1 \leq j \leq k+1$. On this event, $\widetilde{V}_{k+1}(4\iota) \subseteq V_k(5\iota)$. Thus $\mu_t(V_k(5\iota)) \geq 1 - e^{-cN}$. This completes the induction, upon adjusting ι. □

Proof of Proposition 6.5.5. Let

$$V_k^+(\iota) = \left\{ \boldsymbol{\sigma} \in S_N : |\langle \boldsymbol{\sigma}, \boldsymbol{m}^k \rangle_N - q_k| \leq \iota \right\},$$

so clearly $V_k^+(\iota) \supseteq V_k(\iota)$. By Proposition 6.5.4, for all $k \geq k_0$ we have $|q_k - q_*| \leq \iota$. Thus

$$\mathsf{Band}(\boldsymbol{m}^k, [q_* - 2\iota, q_* + 2\iota]) \supseteq V_k^+(\iota).$$

By Proposition 6.5.11, with probability $1 - e^{-cN}$,

$$\mu_t(\mathsf{Band}(\boldsymbol{m}^k, [q_* - 2\iota, q_* + 2\iota])) \geq 1 - e^{-cN}.$$

The result follows by adjusting ι. □

6.5.3 Overlap with planted signal

The following proposition completes the proof of (6.44).

Proposition 6.5.12. *Let $\iota > 0$ and $I = I(\iota)$. With probability $1 - e^{-cN}$,*

$$\mu_t(\mathsf{Band}(\boldsymbol{x}, I)) \geq 1 - e^{-cN}.$$

Lemma 6.5.13. *The function*

$$f(q) = \xi_t(q) + q + \log(1-q)$$

is maximized over $[0,1]$ uniquely at $q = q_*$.

Proof. We calculate

$$f'(q) = \xi_t'(q) - \frac{q}{1-q}, \qquad f''(q) = \xi''(q) - \frac{1}{(1-q)^2}.$$

By (6.39), f is stationary at q_*. By (6.5), it is concave on $[0,1)$. □

We will use the following replica-symmetric upper bound on the free energy. Let $\widehat{H}_N$ be the Hamiltonian a spherical spin glass with mixture $\widehat{\xi}$, which may contain a degree-1 term (i.e., possibly $\widehat{\xi}'(0) > 0$).

Define the partition function

$$\widehat{Z}_N = \int_{S_N} \exp\left\{\widehat{H}_N(\boldsymbol{\sigma})\right\} \mu_0(\mathsf{d}\boldsymbol{\sigma}). \tag{6.63}$$

Proposition 6.5.14. *For any $u \in [0,1)$, we have*

$$\operatorname*{p-lim}_{N\to\infty} \frac{1}{N} \log \widehat{Z}_N \leq \frac{1}{2}\left(\widehat{\xi}(1) - \widehat{\xi}(u) + \frac{u}{1-u} + \log(1-u)\right). \tag{6.64}$$

Furthermore, equality holds if

$$g(s) = \int_0^s \left(\widehat{\xi}'(r) - \frac{r}{(1-u)^2}\right) \, \mathsf{d}r \tag{6.65}$$

is maximized over $s \in [0,u]$ at $s = u$, and $\widehat{\xi}_u(s) = \widehat{\xi}(u + (1-u)s) - \widehat{\xi}(u) - (1-u)\widehat{\xi}'(u)s$ satisfies

$$\widehat{\xi}_u(s) + s + \log(1-s) \leq 0 \tag{6.66}$$

for all $s \in [0,1)$.

Proof. The bound (6.64) is the spherical Parisi formula [Tal06a, Theorem 1.1] with order parameter δ_u. The equality condition follows from the extremality condition [Tal06a, Proposition 2.1]. □

Let $H_{N,t}$ be as in (6.40). Let ψ_N denote the probability density of z_1, where $\boldsymbol{z}$ is a sample from the uniform Haar measure on the unit sphere $\mathbb{S}^{N-1}$. It is known that

$$\psi_N(q) = \frac{1}{Z_{N,\psi}}(1-q^2)^{(N-3)/2}, \qquad q \in [-1,1] \tag{6.67}$$

for some normalizing constant $Z_{N,\psi}$. For $q \in [-1,1]$, define

$$Z(q) = \int_{\mathsf{Band}(\boldsymbol{x},q)} \exp\left\{H_{N,t}(\boldsymbol{\sigma})\right\} \, \mathsf{d}\mu_{(q)}(\boldsymbol{\sigma}), \tag{6.68}$$

where $\mu_{(q)}$ is the uniform measure on $\mathsf{Band}(\boldsymbol{x},q)$, normalized to $\mu_{(q)}(\mathsf{Band}(\boldsymbol{x},q)) = \psi_N(q)$. Note that

$$\int_{-1}^1 Z(q) \, \mathsf{d}q = \int_{S_N} \exp\left\{H_{N,t}(\boldsymbol{\sigma})\right\} \, \mathsf{d}\mu_0(\boldsymbol{\sigma}).$$

Proposition 6.5.15. *For any fixed $q \in (-1,1)$,*

$$\operatorname*{p-lim}_{N\to\infty} \frac{1}{N} \log Z(q) \le \frac{1}{2}\Big(\xi_t(1) + \xi_t(|q|) + |q| + \log(1-|q|)\Big). \tag{6.69}$$

Equality holds for $q = q_*$, and does not hold for any $q < 0$.

Proof. Consider first $q \in [0,1]$. On $\mathsf{Band}(\boldsymbol{x}, q)$, if we write $\boldsymbol{\sigma} = q\boldsymbol{x} + \sqrt{1-q^2}\boldsymbol{\rho}$, where $\langle \boldsymbol{x}, \boldsymbol{\rho}\rangle = 0$, then the random part

$$\widehat{H}_{N,q}(\boldsymbol{\rho}) := \widetilde{H}_{N,t}(\boldsymbol{\sigma}) - \widetilde{H}_{N,t}(q\boldsymbol{x}) = \widetilde{H}_{N,t}(q\boldsymbol{x} + \sqrt{1-q^2}\boldsymbol{\rho})$$

is a spin glass with one fewer dimension and mixture ξ replaced by

$$\widehat{\xi}(s) = \xi_t(q^2 + (1-q^2)s) - \xi_t(q^2).$$

Then,

$$\operatorname*{p-lim}_{N\to\infty} \frac{1}{N} \log Z(q) = \xi_t(q) + \frac{1}{2}\log(1-q^2) + \operatorname*{p-lim}_{N\to\infty} \frac{1}{N} \log \widehat{Z}_{N,q}, \tag{6.70}$$

where $\widehat{Z}_{N,q}$ is the free energy of the spin glass with Hamiltonian $\widehat{H}_{N,q}$. By Proposition 6.5.14 with $u = \frac{q}{1+q}$,

$$\operatorname*{p-lim}_{N\to\infty} \frac{1}{N} \log \widehat{Z}_{N,q} \le \frac{1}{2}\left(\xi_t(1) - \xi_t(q) + q - \log(1+q)\right). \tag{6.71}$$

Combining with (6.70) proves (6.69). For $q < 0$, (6.70) still holds. Since $\xi_t(q) < \xi_t(|q|)$, and the remaining terms on the right-hand side of (6.70) depend on q only through $|q|$, (6.69) holds with strict inequality.

To show that equality holds in (6.69) for $q = q_*$, we will verify that (6.71) holds with equality. Let $u_* = \frac{q_*}{1+q_*}$. Then

$$\frac{\mathsf{d}\widehat{\xi}}{\mathsf{d}u}(u_*) = (1-q_*^2)\xi_t'(q_*) \stackrel{(6.39)}{=} q_*(1+q_*) = \frac{u_*}{(1-u_*)^2},$$

while

$$\frac{\mathsf{d}^2\widehat{\xi}}{\mathsf{d}u^2}(u_*) = (1-q_*^2)^2\xi''(q_*) \stackrel{(6.5)}{<} (1+q_*)^2 = \frac{1}{(1-u_*)^2}.$$

Thus, for g in (6.65), $g'(u_*) = 0$ and $g''(u_*) < 0$. However, over $s \in [0, u_*]$,

$$g'(s) = \widehat{\xi}'(s) - \frac{1}{(1-u_*)^2}$$

is convex because $\widehat{\xi}'$ is convex. So, $g''(s) < 0$ for all $s \in [0, u_*]$, which implies $g'(s) \ge 0$ for all $s \in [0, u_*]$. It follows that $g(s)$ is maximized over $s \in [0, u_*]$ at u_*, verifying (6.65). Since

$$\begin{aligned}\widehat{\xi}_{u_*}(s) &= \xi_t(q_* + (1-q_*)s) - \xi_t(q_*) - (1-q_*)\xi_t'(q_*)s \\ &\stackrel{(6.39)}{=} \xi_t(q_* + (1-q_*)s) - \xi_t(q_*) - q_*s,\end{aligned}$$

we have

$$\begin{aligned}\widehat{\xi}_{u_*}(s) + s + \log(1-s) = &\Big\{\xi_t(q_* + (1-q_*)s) + (q_* + (1-q_*)s) + \log\left[1 - (q_* + (1-q_*)s)\right]\Big\} \\ &- \Big\{\xi_t(q_*) + q_* + \log(1-q_*)\Big\} \le 0,\end{aligned}$$

where the final inequality is by Lemma 6.5.13. This verifies (6.66) and completes the proof. ☐

Proof of Proposition 6.5.12. Fix $\iota > 0$ arbitrarily (independent of N). We will choose $\upsilon = \upsilon(\iota)$ a sufficiently small constant to verify the derivations below. Let

$$q_k^+ = q_* + \iota + k\upsilon, \qquad q_k^- = q_* - \iota - k\upsilon,$$

and let k^+ (resp. k^-) be the largest integer such that $q_{k^+}^+ \le 1$ (resp. $q_{k^-}^- \ge -1$). Let

$$J = \{q_{k^-}^-, \dots, q_1^-, q_1^+, \dots, q_{k^+}^+\}.$$

Define $h(q) = \frac{1}{2}(\xi_t(1) + \xi_t(|q|) + |q| + \log(1-|q|))$ to be the right-hand side of (6.69). Consider the event:

- K_N from Proposition 6.3.6 holds,
- $\frac{1}{N}\log Z(q_*) \ge h(q_*) - \upsilon$,
- $\frac{1}{N}\log Z(q) \le h(q) + \upsilon$ for all $q \in J$.

This holds with probability $1 - e^{-cN}$ by concentration properties of $Z(q)$. Further let

$$\begin{aligned}
Z_0 &= \int_{S_N} \mathbf{1}\{\langle \boldsymbol{\sigma}, \boldsymbol{x}\rangle_N \in [q_* - \upsilon, q_* + \upsilon]\}\big\{\exp H_{N,t}(\boldsymbol{\sigma})\big\}\mathsf{d}\mu_0(\boldsymbol{\sigma}) = \int_{q_*-\upsilon}^{q_*+\upsilon} Z(q)\ \mathsf{d}q,\\
Z_k^+ &= \int_{S_N} \mathbf{1}\{\langle \boldsymbol{\sigma}, \boldsymbol{x}\rangle_N \in [q_k^+, q_k^+ + \upsilon]\}\big\{\exp H_{N,t}(\boldsymbol{\sigma})\big\}\mathsf{d}\mu_0(\boldsymbol{\sigma}) = \int_{q_k^+}^{q_k^+ +\upsilon} Z(q)\ \mathsf{d}q,\\
Z_k^- &= \int_{S_N} \mathbf{1}\{\langle \boldsymbol{\sigma}, \boldsymbol{x}\rangle_N \in [q_k^- - \upsilon, q_k^-]\}\big\{\exp H_{N,t}(\boldsymbol{\sigma})\big\}\mathsf{d}\mu_0(\boldsymbol{\sigma}) = \int_{q_k^- -\upsilon}^{q_k^-} Z(q)\ \mathsf{d}q.
\end{aligned}$$

Since K_N holds, $H_{N,t}(\boldsymbol{\sigma})$ is $O(1)$-Lipschitz, and thus

$$Z_0 \ge Z(q_*)e^{-o_\upsilon(1)N}, \qquad Z_k^+ \le Z(q_k^+)e^{o_\upsilon(1)N}, \qquad Z_k^- \le Z(q_k^-)e^{o_\upsilon(1)N}.$$

Here and below, $o_\upsilon(1)$ denotes a term independent of N that vanishes as $\upsilon \to 0$. So

$$\frac{1}{N}\log\int_{S_N} \mathbf{1}\{\langle \boldsymbol{\sigma}, \boldsymbol{x}\rangle_N \in [q_* - \iota, q_* + \iota]\}\exp H_{N,t}(\boldsymbol{\sigma})\mathsf{d}\mu_0(\boldsymbol{\sigma}) \ge \frac{1}{N}\log Z_0 \ge h(q_*) - o_\upsilon(1)$$

while

$$\begin{aligned}
\frac{1}{N}\log\int_{S_N} \mathbf{1}\{\langle \boldsymbol{\sigma}, \boldsymbol{x}\rangle_N \notin [q_* - \iota, q_* + \iota]\} &\le \frac{1}{N}\log\left(\sum_{k=0}^{k^+} Z_k^+ + \sum_{k=0}^{k^-} Z_k^-\right)\\
&\le \max_{q\in J} h(q) + o_\upsilon(1).
\end{aligned}$$

By Lemma 6.5.13, for υ small enough,

$$h(q_*) - o_\upsilon(1) > \max_{q\in J} h(q) + o_\upsilon(1)$$

and thus $\mu_t([q_* - \iota, q_* + \iota]) \ge 1 - e^{-cN}$. □

Proof of Proposition 6.4.3. Follows from Propositions 6.5.4, 6.5.5, and 6.5.12. □

6.6 Description of TAP fixed point: proof of Proposition 6.4.4

6.6.1 Existence and uniqueness of TAP fixed point

We say that $\boldsymbol{m}$ is a ι-approximate critical point of $\mathcal{F}_{\mathsf{TAP}}$ if $\|\nabla\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m})\|_N \le \iota$. In this subsection we show the following result.

Proposition 6.6.1. *There exist $C_{\max}^{\mathsf{spec}} > C_{\min}^{\mathsf{spec}} > 0$ such that, for sufficiently small $\iota > 0$, the following holds with probability $1 - e^{-cN}$.*

- *(a) $\mathcal{F}_{\mathsf{TAP}}$ has a unique critical point $\boldsymbol{m}^{\mathsf{TAP}}$ in $\mathcal{S}_\iota$, which further satisfies (6.46).*
- *(b) There exists $\iota' = o_\iota(1)$ such that any ι-approximate critical point $\boldsymbol{m} \in \mathcal{S}_\iota$ of $\mathcal{F}_{\mathsf{TAP}}$ satisfies $\|\boldsymbol{m} - \boldsymbol{m}^{\mathsf{TAP}}\|_N \le \iota'$.*

The proof of this proposition depends on an understanding of the landscape of $\widetilde{H}_{N,t}$ restricted to $\mathcal{S}_0$, given in Proposition 6.6.2 below (recall that $\widetilde{H}_{N,t}$ is the centered version of the Hamiltonian $H_{N,t}$, cf. Eqs. (6.40) and (6.41)). Note that $\mathcal{S}_0$ is an affine transformation of the sphere S_{N-2}; we will view it as a Riemannian manifold. We first recall notions of Riemannian gradient and Hessian. For $\boldsymbol{m} \in \mathcal{S}_0$, let

$$\boldsymbol{m}^\perp = \frac{\boldsymbol{m} - q_* \boldsymbol{x}}{\sqrt{q_*(1-q_*)}},$$

so that $\langle \boldsymbol{x}, \boldsymbol{m}^\perp \rangle_N = 0$ and $\|\boldsymbol{m}^\perp\|_N = 1$. The Riemannian gradient and radial derivative of $\widetilde{H}_{N,t}$ are

$$\nabla_{\mathsf{sp}} \widetilde{H}_{N,t}(\boldsymbol{m}) = P^\perp_{\mathsf{span}(\boldsymbol{m},\boldsymbol{x})} \nabla \widetilde{H}_{N,t}(\boldsymbol{m}), \qquad \partial_{\mathsf{rad}} \widetilde{H}_{N,t}(\boldsymbol{m}) = \langle \boldsymbol{m}^\perp, \nabla \widetilde{H}_{N,t}(\boldsymbol{m}) \rangle / \sqrt{N}.$$

In the below calculations, it will be convenient to work with the following rescaled radial derivative, whose typical maximum is $O(1)$:

$$\widetilde{\partial}_{\mathsf{rad}} \widetilde{H}_{N,t}(\boldsymbol{m}) = \partial_{\mathsf{rad}} \widetilde{H}_{N,t}(\boldsymbol{m}) / \sqrt{N} = \langle \boldsymbol{m}^\perp, \nabla \widetilde{H}_{N,t}(\boldsymbol{m}) \rangle_N .$$

Similarly to above, we say $\boldsymbol{m} \in \mathcal{S}_0$ is a **Riemannian critical point** of $\widetilde{H}_{N,t}$ if $\nabla_{\mathsf{sp}} \widetilde{H}_{N,t}(\boldsymbol{m}) = \boldsymbol{0}$, and an **$\iota$-approximate Riemannian critical point** if $\|\nabla_{\mathsf{sp}} \widetilde{H}_{N,t}(\boldsymbol{m})\|_N \le \iota$. Further define the tangential and Riemannian Hessian (these will be used in the next subsection)

$$\begin{aligned}
\nabla^2_{\mathsf{tan}} \widetilde{H}_{N,t}(\boldsymbol{m}) &= P^\perp_{\mathsf{span}(\boldsymbol{m},\boldsymbol{x})} \nabla^2 \widetilde{H}_{N,t}(\boldsymbol{m}) P^\perp_{\mathsf{span}(\boldsymbol{m},\boldsymbol{x})}, \\
\nabla^2_{\mathsf{sp}} \widetilde{H}_{N,t}(\boldsymbol{m}) &= \nabla^2_{\mathsf{tan}} \widetilde{H}_{N,t}(\boldsymbol{m}) - \frac{\widetilde{\partial}_{\mathsf{rad}} \widetilde{H}_{N,t}(\boldsymbol{m})}{\sqrt{q_*(1-q_*)}} P^\perp_{\mathsf{span}(\boldsymbol{m},\boldsymbol{x})}.
\end{aligned}$$

Proposition 6.6.2. *There exist $C^{\mathsf{spec}}_{\max} > C^{\mathsf{spec}}_{\min} > 0$ such that for any $\iota > 0$, the following holds with probability $1 - e^{-cN}$.*

(a) $\widetilde{H}_{N,t}$ has exactly two Riemannian critical points $\boldsymbol{m}_\pm$ on $\mathcal{S}_0$, and their (rescaled) radial derivatives satisfy

$$\left| \widetilde{\partial}_{\mathsf{rad}} \widetilde{H}_{N,t}(\boldsymbol{m}_\pm) \mp \sqrt{\frac{q_*}{1-q_*}} \left(1 + (1-q_*)^2 \xi''(q_*)\right) \right| \le \iota. \tag{6.72}$$

Moreover, there exists $\iota' = o_\iota(1)$ such that all ι-approximate Riemannian critical points $\boldsymbol{m}$ on $\mathcal{S}_0$ satisfy $\|\boldsymbol{m} - \boldsymbol{m}_\pm\|_N \le \iota'$ for some choice of sign $\pm$.

(b) The point $\boldsymbol{m}_+$ is an ι-approximate critical point of $\mathcal{F}_{\mathsf{TAP}}$ (i.e. $\|\nabla \mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m})\|_N \le \iota$).

(c) The point $\boldsymbol{m}_+$ satisfies

$$\mathsf{spec}(\nabla^2 \mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m}_+)) \subseteq [-C^{\mathsf{spec}}_{\max}, -C^{\mathsf{spec}}_{\min}].$$

We will prove this proposition in Subsection 6.6.2. We first show Proposition 6.6.1 given Proposition 6.6.2.

Lemma 6.6.3. *For sufficiently small $\iota > 0$, with probability $1 - e^{-cN}$, $\mathcal{F}_{\mathsf{TAP}}$ has a unique critical point $\boldsymbol{m}$ in the region $\|\boldsymbol{m} - \boldsymbol{m}_+\|_N \le \iota$, which further satisfies (6.46).*

Proof. Throughout this proof, assume the event K_N from Proposition 6.3.6 holds, which occurs with probability $1 - e^{-cN}$. By Proposition 6.6.2(c), with probability $1 - e^{-cN}$, $\boldsymbol{m}_+$ is well-defined and

$$\mathsf{spec}(\nabla^2 \mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m}_+)) \subseteq [-C^{\mathsf{spec}}_{\max}, -C^{\mathsf{spec}}_{\min}].$$

On K_N, the maps $\boldsymbol{m} \mapsto \lambda_{\max}(\nabla^2 \mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m}))$ and $\boldsymbol{m} \mapsto \lambda_{\min}(\nabla^2 \mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m}_+))$ are $O(1)$-Lipschitz (over $\|\boldsymbol{m}\|_N \le 1 - \varepsilon$, for any $\varepsilon > 0$). Thus, for suitably small ι,

$$\mathsf{spec}(\nabla^2 \mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m})) \subseteq \left[-2C^{\mathsf{spec}}_{\max}, -\frac{1}{2} C^{\mathsf{spec}}_{\min}\right] \qquad \forall \|\boldsymbol{m} - \boldsymbol{m}_+\|_N \le \iota. \tag{6.73}$$

Let υ be suitably small in ι. By Proposition 6.6.2(b), with probability $1 - e^{-cN}$, $\|\nabla \mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m}_+)\|_N \le \upsilon$. Combined with (6.73), this implies $\mathcal{F}_{\mathsf{TAP}}$ has a unique critical point $\boldsymbol{m}$ in the region $\|\boldsymbol{m} - \boldsymbol{m}_+\|_N \le \iota$. By (6.73), this critical point also satisfies (6.46), upon adjusting the constants $C^{\mathsf{spec}}_{\min}, C^{\mathsf{spec}}_{\max}$. $\square$

Lemma 6.6.4. *For any sufficiently small $\iota > 0$, there exists $\iota' = o_\iota(1)$ such that with probability $1 - e^{-cN}$, all ι-approximate critical points $\boldsymbol{m} \in \mathcal{S}_\iota$ of $\mathcal{F}_{\mathsf{TAP}}$ satisfy $\|\boldsymbol{m} - \boldsymbol{m}_+\|_N \le \iota'$.*

Proof. Suppose K_N holds. Let $\boldsymbol{m} \in \mathcal{S}_\iota$ be an ι-approximate critical point of $\mathcal{F}_{\mathsf{TAP}}$, and let $\widetilde{\boldsymbol{m}}$ be the nearest point in $\mathcal{S}_0$ to $\boldsymbol{m}$, so that $\|\boldsymbol{m} - \widetilde{\boldsymbol{m}}\|_N \le 2\iota$. On K_N, the map $\boldsymbol{m} \mapsto \nabla \mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m})$ is $O(1)$-Lipschitz. Thus $\widetilde{\boldsymbol{m}}$ is a $O(\iota)$-approximate critical point of $\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m})$, i.e.

$$\left\| \nabla \widetilde{H}_{N,t}(\widetilde{\boldsymbol{m}}) + \xi_t'(q_*)\boldsymbol{x} - \left((1-q_*)\xi''(q_*) + \frac{1}{1-q_*} \right) \widetilde{\boldsymbol{m}} \right\|_N \le O(\iota). \tag{6.74}$$

Thus $\|\nabla_{\mathsf{sp}} \widetilde{H}_{N,t}(\widetilde{\boldsymbol{m}})\|_N \le O(\iota)$. By Proposition 6.6.2(a), there exists $\iota' = o_\iota(1)$ such that on an event with probability $1 - e^{-cN}$, $\|\widetilde{\boldsymbol{m}} - \boldsymbol{m}_\pm\|_N \le \iota'/2$ for some choice of sign $\pm$. We now show the sign must be $+$. By (6.74),

$$\begin{aligned}
\widetilde{\partial}_{\mathsf{rad}} \widetilde{H}_{N,t}(\widetilde{\boldsymbol{m}}) &= \frac{1}{\sqrt{q_*(1-q_*)}} R\left(\widetilde{\boldsymbol{m}} - q_*\boldsymbol{x}, -\xi_t'(q_*)\boldsymbol{x} + \left((1-q_*)\xi''(q_*) + \frac{1}{1-q_*} \right) \widetilde{\boldsymbol{m}} \right) + O(\iota) \\
&= \sqrt{\frac{q_*}{1-q_*}} \left(1 + (1-q_*)^2 \xi''(q_*) \right) + O(\iota).
\end{aligned} \tag{6.75}$$

If we had $\|\widetilde{\boldsymbol{m}} - \boldsymbol{m}_-\|_N \le \iota'/2$, then Eq. (6.72) and Lipschitzness of $\boldsymbol{m} \mapsto \nabla \mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m})$ would imply

$$\widetilde{\partial}_{\mathsf{rad}} \widetilde{H}_{N,t}(\widetilde{\boldsymbol{m}}) = -\sqrt{\frac{q_*}{1-q_*}} \left(1 + (1-q_*)^2 \xi''(q_*) \right) + O(\iota'),$$

which contradicts (6.75) for small enough ι. Thus $\|\widetilde{\boldsymbol{m}} - \boldsymbol{m}_+\|_N \le \iota'/2$. Recalling $\|\boldsymbol{m} - \widetilde{\boldsymbol{m}}\|_N \le 2\iota$ implies the conclusion. □

Proof of Proposition 6.6.1. By Lemma 6.6.3, ($\boldsymbol{m}_+$ is well-defined and) there is a unique critical point of $\mathcal{F}_{\mathsf{TAP}}$ in the region $\|\boldsymbol{m} - \boldsymbol{m}_+\|_N \le \iota$, which also satisfies (6.46). Let $\boldsymbol{m}^{\mathsf{TAP}}$ denote this point.

Let $\iota' = o_\iota(1)$ be given by Lemma 6.6.4. For ι sufficiently small, Lemma 6.6.3 also implies that $\boldsymbol{m}^{\mathsf{TAP}}$ is the unique critical point of $\mathcal{F}_{\mathsf{TAP}}$ in the region $\|\boldsymbol{m} - \boldsymbol{m}_+\|_N \le \iota'$.

By Lemma 6.6.4, all ι-approximate critical points $\boldsymbol{m} \in \mathcal{S}_\iota$ of $\mathcal{F}_{\mathsf{TAP}}$ satisfy $\|\boldsymbol{m} - \boldsymbol{m}_+\|_N \le \iota'$. In particular all critical points are in this region, and thus $\boldsymbol{m}^{\mathsf{TAP}}$ is the unique critical point. This proves part (a). Furthermore, for ι-approximate critical points $\boldsymbol{m} \in \mathcal{S}_\iota$,

$$\|\boldsymbol{m} - \boldsymbol{m}^{\mathsf{TAP}}\|_N \le \|\boldsymbol{m} - \boldsymbol{m}_+\|_N + \|\boldsymbol{m}^{\mathsf{TAP}} - \boldsymbol{m}_+\|_N \le 2\iota'.$$

This proves part (b) upon adjusting ι'. □

6.6.2 Characterization of Riemannian critical points: proof of Proposition 6.6.2

The proof builds on a sequence of recent results on **topological trivialization** in spherical spin glasses [FLD14, Fyo15, BČNS22, HS23c].

Proof of Proposition 6.6.2(a). For $\boldsymbol{m} \in \mathcal{S}_0$, we may write $\boldsymbol{m} = q_*\boldsymbol{x} + \sqrt{q_*(1-q_*)}\boldsymbol{\tau}$, where $\langle \boldsymbol{x}, \boldsymbol{\tau} \rangle_N = 0$ and $\|\boldsymbol{\tau}\|_N = 1$. Let

$$\widehat{H}(\boldsymbol{\tau}) = \widetilde{H}_{N,t}(q_*\boldsymbol{x} + \sqrt{q_*(1-q_*)}\boldsymbol{\tau}) - \widetilde{H}_{N,t}(q_*\boldsymbol{x}).$$

This is a spin glass (in 1 fewer dimension) with mixture

$$\widetilde{\xi}(s) = \xi_t(q_*^2 + q_*(1-q_*)s) - \xi_t(q_*^2). \tag{6.76}$$

Note that

$$\widetilde{\xi}'(1) = q_*(1-q_*)\xi_t'(q_*) \overset{(6.39)}{=} q_*^2, \qquad \widetilde{\xi}''(1) = q_*^2(1-q_*)^2\xi''(q_*) \overset{(6.5)}{<} q_*^2.$$

Thus $\widetilde{\xi}'(1) > \widetilde{\xi}''(1)$, which is the condition for topological trivialization identified in [Fyo15, Equation 64], see also [BČNS22, Theorem 1.1]. Thus, with high probability, $\widehat{H}$ has exactly two critical points $\boldsymbol{\tau}_\pm$, which have radial derivative

$$\widetilde{\partial}_{\mathsf{rad}}\widehat{H}(\boldsymbol{\tau}_\pm) = \pm\left(\sqrt{\widetilde{\xi}'(1)} + \frac{\widetilde{\xi}''(1)}{\sqrt{\widetilde{\xi}'(1)}}\right) + O(\iota) = \pm q_*\left(1+(1-q_*)^2\xi''(1)\right) + O(\iota).$$

By [HS23c, Theorem 1.6], this actually holds with probability $1-e^{-cN}$. On this event, $\widetilde{H}_{N,t}$ has exactly two Riemannian critical points $\boldsymbol{m}_\pm$ on $\mathcal{S}_0$, which have radial derivative

$$\widetilde{\partial}_{\mathsf{rad}}\widetilde{H}_{N,t}(\boldsymbol{m}_\pm) = \frac{1}{\sqrt{q_*(1-q_*)}}\cdot\widetilde{\partial}_{\mathsf{rad}}\widehat{H}(\boldsymbol{\tau}^\pm) = \pm\sqrt{\frac{q_*}{1-q_*}}\left(1+(1-q_*)^2\xi''(1)\right) + O(\iota).$$

The estimate (6.72) holds by adjusting ι. The claim about approximate critical points also follows from [HS23c, Theorem 1.6], which shows that all approximate critical points are close to exact critical points. □

We will prove parts (b) and (c) by slightly modifying the calculation in [Fyo15, BČNS22]. This calculation is based on the Kac–Rice formula, which we now recall. Let Crt denote the set of Riemannian critical points of $\widetilde{H}_{N,t}$ on $\mathcal{S}_0$ and $\mu_{\mathcal{S}_0}$ denote the $(N-2)$-dimensional Hausdorff measure on $\mathcal{S}_0$. The Kac–Rice Formula [Ric44, Kac48] (see [AT09] for a textbook treatment), applied to $\nabla\widetilde{H}_{N,t}$ on the Riemannian manifold $\mathcal{S}_0$, states that for any (random) measurable set $\mathcal{T}\subseteq\mathcal{S}_0$,

$$\mathbb{E}\,|\mathsf{Crt}\cap\mathcal{T}| = \int_{\mathcal{S}_0}\mathbb{E}\left[|\det\nabla^2_{\mathsf{sp}}\widetilde{H}_{N,t}(\boldsymbol{m})|\mathbf{1}\{\boldsymbol{m}\in\mathcal{T}\}\middle|\nabla_{\mathsf{sp}}\widetilde{H}_{N,t}(\boldsymbol{m})=\boldsymbol{0}\right]\varphi_{\nabla_{\mathsf{sp}}\widetilde{H}_{N,t}(\boldsymbol{m})}(\boldsymbol{0})\,\mathsf{d}\mu_{\mathcal{S}_0}(\boldsymbol{m}). \tag{6.77}$$

Here φ_X denotes the probability density of the random variable X, and $\nabla^2_{\mathsf{sp}}\widetilde{H}_{N,t}(\boldsymbol{m})$ is understood as a $(N-2)\times(N-2)$ matrix. The following fact is standard, see, e.g., [AB13, Lemma 1].

Fact 6.6.5. *For any $\boldsymbol{m}\in\mathcal{S}_0$, the random variables $\partial_{\mathsf{rad}}\widetilde{H}_{N,t}(\boldsymbol{m})$, $\nabla_{\mathsf{sp}}\widetilde{H}_{N,t}(\boldsymbol{m})$, and $\nabla^2_{\mathsf{tan}}\widetilde{H}_{N,t}(\boldsymbol{m})$ are independent and Gaussian. Moreover, with $\boldsymbol{G}\sim\mathsf{GOE}(N-2)$, we have*

$$\nabla^2_{\mathsf{tan}}\widetilde{H}_{N,t}(\boldsymbol{m}) \overset{d}{=} \sqrt{\xi''(q_*)\frac{N-2}{N}}\boldsymbol{G}.$$

We defer the proof of the following lemma to Subsection 6.6.5.

Lemma 6.6.6. *Let $\boldsymbol{G}\sim\mathsf{GOE}(N)$. For any $t\ge 1$, $r>2$, there exists $C_{r,t}>0$, uniform for r in compact subsets of $(2,+\infty)$, such that*

$$\mathbb{E}\left[|\det(r\boldsymbol{I}-\boldsymbol{G})|^t\right]^{1/t}\le C_{r,t}\,\mathbb{E}\left[|\det(r\boldsymbol{I}-\boldsymbol{G})|\right].$$

Proposition 6.6.7. *We have $\mathbb{E}\,|\mathsf{Crt}| = 2+o_N(1)$.*

Proof. As shown in the proof of Proposition 6.6.2(a) above, after reparametrizing $\mathcal{S}_0$ to a sphere of radius $\sqrt{N}$, the restriction of $\widetilde{H}_{N,t}$ to $\mathcal{S}_0$ is a spherical spin glass in one fewer dimension with mixture $\widetilde{\xi}$ (6.76), which satisfies $\widetilde{\xi}'(1)>\widetilde{\xi}''(1)$. The claim follows from [Fyo15, Equation 64] or [BČNS22, Theorem 1.2]. □

We will use (6.77) through the following lemma. Let

$$r_* = \sqrt{\frac{q_*}{1-q_*}}\left(1+(1-q_*)^2\xi''(q_*)\right). \tag{6.78}$$

Lemma 6.6.8. *Let $\iota>0$ be sufficiently small, $I_\iota=[r_*-\iota,r_*+\iota]$, and*

$$\mathcal{T}_\iota = \left\{\boldsymbol{m}\in\mathcal{S}_0 : \widetilde{\partial}_{\mathsf{rad}}\widetilde{H}_{N,t}(\boldsymbol{m})\in I_\iota\right\}.$$

There exists a constant $C>0$ (independent of ι) such that for any measurable $\mathcal{T}\subseteq\mathcal{T}_\iota$,

$$\mathbb{E}\,|\mathsf{Crt}\cap\mathcal{T}|\le C\sup_{\boldsymbol{m}\in\mathcal{S}_0}\sup_{r\in I_\iota}\mathbb{P}\left[\boldsymbol{m}\in\mathcal{T}\middle|\nabla_{\mathsf{sp}}\widetilde{H}_{N,t}(\boldsymbol{m})=\boldsymbol{0},\widetilde{\partial}_{\mathsf{rad}}\widetilde{H}_{N,t}(\boldsymbol{m})=r\right]^{1/2}.$$

Proof. By Fact 6.6.5, $\widetilde{\partial}_{\mathsf{rad}}\widetilde{H}_{N,t}(\boldsymbol{m})$ is independent of $\nabla_{\mathsf{sp}}\widetilde{H}_{N,t}(\boldsymbol{m})$. Explicitly integrating $\widetilde{\partial}_{\mathsf{rad}}\widetilde{H}_{N,t}(\boldsymbol{m})$ in (6.77) gives

$$\begin{aligned}\mathbb{E}\,|\mathsf{Crt}\cap\mathcal{T}| = \int_{\mathcal{S}_0}\int_{I_\iota}\mathbb{E}\left[|\det\nabla^2_{\mathsf{sp}}\widetilde{H}_{N,t}(\boldsymbol{m})|\mathbf{1}\{\boldsymbol{m}\in\mathcal{T}\}\middle|\nabla_{\mathsf{sp}}\widetilde{H}_{N,t}(\boldsymbol{m})=\mathbf{0},\widetilde{\partial}_{\mathsf{rad}}\widetilde{H}_{N,t}(\boldsymbol{m})=r\right]\\ \times\varphi_{\widetilde{\partial}_{\mathsf{rad}}\widetilde{H}_{N,t}(\boldsymbol{m})}(r)\varphi_{\nabla_{\mathsf{sp}}\widetilde{H}_{N,t}(\boldsymbol{m})}(\mathbf{0})\,\mathsf{d}r\,\,\mathsf{d}\mu_{\mathcal{S}_0}(\boldsymbol{m}).\end{aligned}$$

By Cauchy-Schwarz,

$$\begin{aligned}&\mathbb{E}\left[|\det\nabla^2_{\mathsf{sp}}\widetilde{H}_{N,t}(\boldsymbol{m})|\mathbf{1}\{\boldsymbol{m}\in\mathcal{T}\}\middle|\nabla_{\mathsf{sp}}\widetilde{H}_{N,t}(\boldsymbol{m})=\mathbf{0},\widetilde{\partial}_{\mathsf{rad}}\widetilde{H}_{N,t}(\boldsymbol{m})=r\right]\\ &\le\mathbb{E}\left[|\det\nabla^2_{\mathsf{sp}}\widetilde{H}_{N,t}(\boldsymbol{m})|^2\middle|\nabla_{\mathsf{sp}}\widetilde{H}_{N,t}(\boldsymbol{m})=\mathbf{0},\widetilde{\partial}_{\mathsf{rad}}\widetilde{H}_{N,t}(\boldsymbol{m})=r\right]^{1/2}\\ &\qquad\times\mathbb{P}\left[\boldsymbol{m}\in\mathcal{T}\middle|\nabla_{\mathsf{sp}}\widetilde{H}_{N,t}(\boldsymbol{m})=\mathbf{0},\widetilde{\partial}_{\mathsf{rad}}\widetilde{H}_{N,t}(\boldsymbol{m})=r\right]^{1/2}.\end{aligned}$$

By Fact 6.6.5, conditional on $\nabla_{\mathsf{sp}}\widetilde{H}_{N,t}(\boldsymbol{m})=\mathbf{0}$, $\widetilde{\partial}_{\mathsf{rad}}\widetilde{H}_{N,t}(\boldsymbol{m})=r$,

$$\begin{aligned}\nabla^2_{\mathsf{sp}}\widetilde{H}_{N,t}(\boldsymbol{m})&\overset{d}{=}\sqrt{\xi''(q_*)\frac{N-2}{N}}\boldsymbol{G}-\frac{r}{\sqrt{q_*(1-q_*)}}\boldsymbol{I}\\ &=\sqrt{\xi''(q_*)\frac{N-2}{N}}\left(\boldsymbol{G}-\sqrt{\frac{N}{N-2}}\frac{r}{\sqrt{q_*(1-q_*)\xi''(q_*)}}\boldsymbol{I}\right).\end{aligned}$$

In light of (6.5),

$$\frac{r_*}{\sqrt{q_*(1-q_*)\xi''(q_*)}}=(1-q_*)\xi''(q_*)^{1/2}+\frac{1}{(1-q_*)\xi''(q_*)^{1/2}}>2,$$

and thus, for $r\in I_\iota$ and ι suitably small,

$$\sqrt{\frac{N}{N-2}}\frac{r}{\sqrt{q_*(1-q_*)\xi''(q_*)}}>2.$$

By Lemma 6.6.6, for some $C>0$,

$$\begin{aligned}&\mathbb{E}\left[|\det\nabla^2_{\mathsf{sp}}\widetilde{H}_{N,t}(\boldsymbol{m})|^2\middle|\nabla_{\mathsf{sp}}\widetilde{H}_{N,t}(\boldsymbol{m})=\mathbf{0},\widetilde{\partial}_{\mathsf{rad}}\widetilde{H}_{N,t}(\boldsymbol{m})=r\right]^{1/2}\\ &=\mathbb{E}\left[\left|\det\sqrt{\xi''(q_*)\frac{N-2}{N}}\left(\boldsymbol{G}-\sqrt{\frac{N}{N-2}}\frac{r}{\sqrt{q_*(1-q_*)\xi''(q_*)}}\boldsymbol{I}\right)\right|^2\right]^{1/2}\\ &\le C\,\mathbb{E}\left[\left|\det\sqrt{\xi''(q_*)\frac{N-2}{N}}\left(\boldsymbol{G}-\sqrt{\frac{N}{N-2}}\frac{r}{\sqrt{q_*(1-q_*)\xi''(q_*)}}\boldsymbol{I}\right)\right|\right]\\ &=C\,\mathbb{E}\left[|\det\nabla^2_{\mathsf{sp}}\widetilde{H}_{N,t}(\boldsymbol{m})|\middle|\nabla_{\mathsf{sp}}\widetilde{H}_{N,t}(\boldsymbol{m})=\mathbf{0},\widetilde{\partial}_{\mathsf{rad}}\widetilde{H}_{N,t}(\boldsymbol{m})=r\right].\end{aligned}\tag{6.79}$$

Combining, we find

$$\begin{aligned}\mathbb{E}\,|\mathsf{Crt}\cap\mathcal{T}|\le C\sup_{\boldsymbol{m}\in\mathcal{S}_0}\sup_{r\in I_\iota}\mathbb{P}\left[\boldsymbol{m}\in\mathcal{T}\middle|\nabla_{\mathsf{sp}}\widetilde{H}_{N,t}(\boldsymbol{m})=\mathbf{0},\widetilde{\partial}_{\mathsf{rad}}\widetilde{H}_{N,t}(\boldsymbol{m})=r\right]^{1/2}\\ \times\int_{\mathcal{S}_0}\int_{I_\iota}\mathbb{E}\left[|\det\nabla^2_{\mathsf{sp}}\widetilde{H}_{N,t}(\boldsymbol{m})|\middle|\nabla_{\mathsf{sp}}\widetilde{H}_{N,t}(\boldsymbol{m})=\mathbf{0},\widetilde{\partial}_{\mathsf{rad}}\widetilde{H}_{N,t}(\boldsymbol{m})=r\right]\\ \times\varphi_{\widetilde{\partial}_{\mathsf{rad}}\widetilde{H}_{N,t}(\boldsymbol{m})}(r)\varphi_{\nabla_{\mathsf{sp}}\widetilde{H}_{N,t}(\boldsymbol{m})}(\mathbf{0})\,\mathsf{d}r\,\,\mathsf{d}\mu_{\mathcal{S}_0}(\boldsymbol{m}).\end{aligned}$$

By the Kac–Rice formula, the last integral is the expected number of Riemannian critical points $\boldsymbol{m}$ of $\widetilde{H}_{N,t}$ with radial derivative $\widetilde{\partial}_{\mathsf{rad}}\widetilde{H}_{N,t}(\boldsymbol{m})\in I_\iota$. This is upper bounded by $\mathbb{E}\,|\mathsf{Crt}|=2+o_N(1)$, by Proposition 6.6.7. □

Proposition 6.6.9. *There exist $C^{\mathsf{spec}}_{\max} > C^{\mathsf{spec}}_{\min} > 0$ such that for all sufficiently small $\iota > 0$, there exists $\iota' = h(\iota) = o_\iota(1)$ such that the following holds. For any $\boldsymbol{m} \in \mathcal{S}_0$ define the events*

$$E_1(\boldsymbol{m}, \iota') := \{\|\nabla \mathcal{F}_{\text{TAP}}(\boldsymbol{m})\|_N \le \iota'\}, \qquad E_2(\boldsymbol{m}) := \left\{\mathsf{spec}(\nabla^2 \mathcal{F}_{\text{TAP}}(\boldsymbol{m})) \subseteq [-C^{\mathsf{spec}}_{\max}, -C^{\mathsf{spec}}_{\min}]\right\}.$$

Then,

$$\inf_{r \in I_\iota} \mathbb{P}\left[E_1(\boldsymbol{m}, \iota') \cap E_2(\boldsymbol{m}) \middle| \nabla_{\mathsf{sp}} \widetilde{H}_{N,t}(\boldsymbol{m}) = \boldsymbol{0}, \widetilde{\partial}_{\mathsf{rad}} \widetilde{H}_{N,t}(\boldsymbol{m}) = r\right] \ge 1 - e^{-cN}.$$

Here the constant c is uniform over $\boldsymbol{m} \in \mathcal{S}_0$.

We prove this proposition in the next subsection. Assuming it, we first complete the proof of Proposition 6.6.2.

Proof of Proposition 6.6.2(b)(c). Let υ be small enough that $\max(\upsilon, h(\upsilon)) \le \iota$, for the h from Proposition 6.6.9. Also let $C^{\mathsf{spec}}_{\max}, C^{\mathsf{spec}}_{\min}$ be given by this proposition. Let $\mathcal{T} \subseteq \mathcal{S}_0$ be the set of points $\boldsymbol{m}$ such that

- $\widetilde{\partial}_{\mathsf{rad}} \widetilde{H}_{N,t}(\boldsymbol{m}) \in I_\upsilon$, and
- $E_1(\boldsymbol{m}, \iota) \cap E_2(\boldsymbol{m})$ does not hold.

Thus $\mathcal{T} \subseteq \mathcal{T}_\upsilon$. By Lemma 6.6.8 and Proposition 6.6.9 (with υ for ι)

$$\begin{aligned}
\mathbb{E}\,|\mathsf{Crt} \cap \mathcal{T}| &\le C \sup_{\boldsymbol{m} \in \mathcal{S}_0} \sup_{r \in I_\upsilon} \mathbb{P}\left[(E_1(\boldsymbol{m}, \iota) \cap E_2(\boldsymbol{m}))^c \middle| \nabla_{\mathsf{sp}} \widetilde{H}_{N,t}(\boldsymbol{m}) = \boldsymbol{0}, \widetilde{\partial}_{\mathsf{rad}} \widetilde{H}_{N,t}(\boldsymbol{m}) = r\right]^{1/2} \\
&\le e^{-cN}.
\end{aligned}$$

Thus, with probability $1 - e^{-cN}$, there do not exist points $\boldsymbol{m} \in \mathcal{S}_0$ such that $\widetilde{\partial}_{\mathsf{rad}} \widetilde{H}_{N,t}(\boldsymbol{m}) \in I_\upsilon$ and $E_1(\boldsymbol{m}, \iota) \cap E_2(\boldsymbol{m})$ does not hold.

However, by Proposition 6.6.2(a) with υ in place of ι, $\widetilde{\partial}_{\mathsf{rad}} \widetilde{H}_{N,t}(\boldsymbol{m}_+) \in I_\upsilon$ with probability $1 - e^{-cN}$. Thus $E_1(\boldsymbol{m}_+, \iota) \cap E_2(\boldsymbol{m}_+)$ holds, completing the proof. □

6.6.3 Approximate stationarity and local concavity of $\mathcal{F}_{\text{TAP}}$: proof of Proposition 6.6.9

Lemma 6.6.10. *Let $\boldsymbol{m} \in \mathcal{S}_0$ and $r \in I_\iota$. Conditional on $\nabla_{\mathsf{sp}} \widetilde{H}_{N,t}(\boldsymbol{m}) = \boldsymbol{0}$ and $\widetilde{\partial}_{\mathsf{rad}} \widetilde{H}_{N,t}(\boldsymbol{m}) = r$, $\langle \boldsymbol{x}, \nabla \widetilde{H}_{N,t}(\boldsymbol{m})\rangle$ is Gaussian with mean $q_*(1-q_*)\xi''(q_*) + O(\iota)$ and variance $O(N^{-1})$.*

Proof. All the random variables considered are jointly Gaussian, so it suffices to compute the conditional mean and variance. A short linear-algebraic calculation shows

$$\mathbb{E}\left[\langle \boldsymbol{x}, \nabla \widetilde{H}_{N,t}(\boldsymbol{m})\rangle_N \middle| \nabla_{\mathsf{sp}} \widetilde{H}_{N,t}(\boldsymbol{m}), \widetilde{\partial}_{\mathsf{rad}} \widetilde{H}_{N,t}(\boldsymbol{m})\right] = \frac{q_*^{3/2}(1-q_*)^{1/2}\xi''(q_*)}{\xi_t'(q_*) + q_*(1-q_*)\xi''(q_*)} \widetilde{\partial}_{\mathsf{rad}} \widetilde{H}_{N,t}(\boldsymbol{m}).$$

Thus

$$\begin{aligned}
&\mathbb{E}\left[\langle \boldsymbol{x}, \nabla \widetilde{H}_{N,t}(\boldsymbol{m})\rangle_N \middle| \nabla_{\mathsf{sp}} \widetilde{H}_{N,t}(\boldsymbol{m}) = \boldsymbol{0}, \widetilde{\partial}_{\mathsf{rad}} \widetilde{H}_{N,t}(\boldsymbol{m}) = r\right] \\
&= \frac{q_*^{3/2}(1-q_*)^{1/2}\xi''(q_*)}{\xi_t'(q_*) + q_*(1-q_*)\xi''(q_*)} r_* + O(\iota) \\
&\overset{(6.39),(6.78)}{=} \frac{q_*^{3/2}(1-q_*)^{1/2}\xi''(q_*)}{\frac{q_*}{1-q_*} + q_*(1-q_*)\xi''(q_*)} \cdot \sqrt{\frac{q_*}{1-q_*}} \left(1 + (1-q_*)^2\xi''(q_*)\right) + O(\iota) \\
&= q_*(1-q_*)\xi''(q_*) + O(\iota).
\end{aligned}$$

Before any conditioning, $\langle \boldsymbol{x}, \nabla \widetilde{H}_{N,t}(\boldsymbol{m})\rangle_N$ is Gaussian with variance $O(N^{-1})$, and conditioning only reduces variance. □

Proposition 6.6.11. *Let $\boldsymbol{m} \in \mathcal{S}_0$ and $r \in I_\iota$. Conditional on $\nabla_{\mathsf{sp}} \widetilde{H}_{N,t}(\boldsymbol{m}) = \boldsymbol{0}$ and $\widetilde{\partial}_{\mathsf{rad}} \widetilde{H}_{N,t}(\boldsymbol{m}) = r$, $E_1(\boldsymbol{m}, \iota')$ holds with probability $1 - e^{-cN}$, for some $\iota' = o_\iota(1)$.*

Proof. By Lemma 6.6.10, with conditional probability $1 - e^{-cN}$,

$$|\langle \boldsymbol{x}, \nabla \widetilde{H}_{N,t}(\boldsymbol{m})\rangle_N - q_*(1-q_*)\xi''(q_*)| \le O(\iota).$$

Suppose this event holds. Since $\nabla_{\mathsf{sp}} \widetilde{H}_{N,t}(\boldsymbol{m}) = \boldsymbol{0}$,

$$\begin{aligned}
\nabla \widetilde{H}_{N,t}(\boldsymbol{m}) &= \widetilde{\partial}_{\mathsf{rad}} \widetilde{H}_{N,t}(\boldsymbol{m}) \frac{\boldsymbol{m} - q_* \boldsymbol{x}}{\sqrt{q_*(1-q_*)}} + \langle \boldsymbol{x}, \nabla \widetilde{H}_{N,t}(\boldsymbol{m})\rangle \boldsymbol{x} \\
&= \sqrt{\frac{q_*}{1-q_*}} \left(1 + (1-q_*)^2 \xi''(q_*)\right) \frac{\boldsymbol{m} - q_* \boldsymbol{x}}{\sqrt{q_*(1-q_*)}} + q_*(1-q_*)\xi''(q_*)\boldsymbol{x} + O(\iota)\boldsymbol{x} + O(\iota)\boldsymbol{m} \\
&= -\xi_t'(q_*)\boldsymbol{x} + \left(\frac{1}{1-q_*} + (1-q_*)\xi''(q_*)\right)\boldsymbol{m} + O(\iota)\boldsymbol{x} + O(\iota)\boldsymbol{m}.
\end{aligned}$$

Since

$$\nabla \mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m}) = \nabla \widetilde{H}_{N,t}(\boldsymbol{m}) + \xi_t'(q_*)\boldsymbol{x} - \left(\frac{1}{1-q_*} + (1-q_*)\xi''(q_*)\right)\boldsymbol{m},$$

it follows that $\|\nabla \mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m})\|_N \le O(\iota)$. □

The next lemma is a linear-algebraic calculation of the conditional law given $\nabla \widetilde{H}_{N,t}(\boldsymbol{m})$ of $\nabla^2 \widetilde{H}_{N,t}(\boldsymbol{m})$, now as a Hessian in $\mathbb{R}^N$ rather than a Riemannian Hessian in $\mathcal{S}_0$. While $\boldsymbol{m} \in \mathcal{S}_0$ for the proofs in the current subsection, we will not assume this for use in Fact 6.6.18 below.

Lemma 6.6.12. *Let $\boldsymbol{m} \in \mathbb{R}^N$ with $\|\boldsymbol{m}\|_N^2 = q_{\boldsymbol{m}} < 1$. Conditional on $\nabla \widetilde{H}_{N,t}(\boldsymbol{m}) = \boldsymbol{z}$, we have*

$$\nabla^2 \widetilde{H}_{N,t}(\boldsymbol{m}) \stackrel{d}{=} \frac{\xi''(q_{\boldsymbol{m}})}{\xi_t'(q_{\boldsymbol{m}})} \cdot \frac{\boldsymbol{m}\boldsymbol{z}^\top + \boldsymbol{z}\boldsymbol{m}^\top}{N} + \frac{\langle \boldsymbol{m}, \boldsymbol{z}\rangle_N}{\xi_t'(q_{\boldsymbol{m}}) + q_{\boldsymbol{m}}\xi''(q_{\boldsymbol{m}})} \left(\xi^{(3)}(q_{\boldsymbol{m}}) - \frac{2\xi''(q_{\boldsymbol{m}})^2}{\xi_t'(q_{\boldsymbol{m}})}\right) \frac{\boldsymbol{m}\boldsymbol{m}^\top}{N} + \boldsymbol{M},$$

where $\boldsymbol{M}$ is the following symmetric random matrix. Let $(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_N)$ be an orthonormal basis of $\mathbb{R}^N$ with $\boldsymbol{e}_1 = \boldsymbol{m}/\|\boldsymbol{m}\|_2$, and to reduce notation let $\boldsymbol{M}(i,j) = \langle \boldsymbol{M}\boldsymbol{e}_i, \boldsymbol{e}_j\rangle$. Then the random variables $\{\boldsymbol{M}(i,j) : 1 \le i \le j \le N\}$ are independent centered Gaussians with variance

$$\mathbb{E}\,\boldsymbol{M}(i,j)^2 = N^{-1} \times \begin{cases} (\textit{irrelevant } O(1)) & 1 = i = j \\ \xi''(q_{\boldsymbol{m}}) + q_{\boldsymbol{m}}\xi^{(3)}(q_{\boldsymbol{m}}) - \frac{q_{\boldsymbol{m}}\xi''(q_{\boldsymbol{m}})^2}{\xi_t'(q_{\boldsymbol{m}})} & 1 = i < j \\ 2\xi''(q_{\boldsymbol{m}}) & 1 < i = j \\ \xi''(q_{\boldsymbol{m}}) & 1 < i < j \end{cases} \tag{6.80}$$

Remark 6.6.13. The covariance calculation in the proof of Lemma 6.6.12 implies $\xi''(q_{\boldsymbol{m}}) + q_{\boldsymbol{m}}\xi^{(3)}(q_{\boldsymbol{m}}) - \frac{q_{\boldsymbol{m}}\xi''(q_{\boldsymbol{m}})^2}{\xi_t'(q_{\boldsymbol{m}})} \ge 0$, but this can also be seen directly by Cauchy-Schwarz:

$$\begin{aligned}
\left(\xi''(q_{\boldsymbol{m}}) + q_{\boldsymbol{m}}\xi^{(3)}(q_{\boldsymbol{m}})\right)\xi_t'(q_{\boldsymbol{m}}) &\ge q_{\boldsymbol{m}} \left(\sum_{p\ge 2} p(p-1)^2 \gamma_p^2 (q_{\boldsymbol{m}})^{p-2}\right)\left(\sum_{p\ge 2} p\gamma_p^2 (q_{\boldsymbol{m}})^{p-2}\right) \\
&\ge q_{\boldsymbol{m}} \left(\sum_{p\ge 2} p(p-1)\gamma_p^2 (q_{\boldsymbol{m}})^{p-2}\right)^2 = q_{\boldsymbol{m}}\xi''(q_{\boldsymbol{m}})^2.
\end{aligned}$$

Proof. It suffices to compute the conditional mean and covariance. Let $\boldsymbol{u}^1, \boldsymbol{u}^2 \in \mathbb{S}^{N-1}$. Then

$$\mathbb{E}\left[\langle \nabla^2 \widetilde{H}_{N,t}(\boldsymbol{m})\boldsymbol{u}^1, \boldsymbol{u}^2\rangle \middle| \nabla \widetilde{H}_{N,t}(\boldsymbol{m})\right] = \langle \boldsymbol{v}, \nabla \widetilde{H}_{N,t}(\boldsymbol{m})\rangle$$

for $\boldsymbol{v} = \boldsymbol{v}(\boldsymbol{u}^1, \boldsymbol{u}^2, \boldsymbol{m})$ such that for all $\boldsymbol{w} \in \mathbb{R}^N$,

$$\langle \nabla^2 \widetilde{H}_{N,t}(\boldsymbol{m})\boldsymbol{u}^1, \boldsymbol{u}^2\rangle - \langle \boldsymbol{v}, \nabla \widetilde{H}_{N,t}(\boldsymbol{m})\rangle \perp\!\!\!\perp \langle \boldsymbol{w}, \nabla \widetilde{H}_{N,t}(\boldsymbol{m})\rangle.$$

We calculate that

$$\begin{aligned}
\mathbb{E}\langle \nabla^2 \widetilde{H}_{N,t}(\boldsymbol{m})\boldsymbol{u}^1, \boldsymbol{u}^2\rangle\langle \boldsymbol{w}, \nabla \widetilde{H}_{N,t}(\boldsymbol{m})\rangle &= N^{-1}\left(\langle \boldsymbol{u}^1, \boldsymbol{w}\rangle\langle \boldsymbol{u}^2, \boldsymbol{m}\rangle + \langle \boldsymbol{u}^1, \boldsymbol{m}\rangle\langle \boldsymbol{u}^2, \boldsymbol{w}\rangle\right)\xi''(q_{\boldsymbol{m}}) \\
&\quad + N^{-2}\langle \boldsymbol{u}^1, \boldsymbol{m}\rangle\langle \boldsymbol{u}^2, \boldsymbol{m}\rangle\langle \boldsymbol{w}, \boldsymbol{m}\rangle\xi^{(3)}(q_{\boldsymbol{m}}), \\
\mathbb{E}\langle \boldsymbol{v}, \nabla \widetilde{H}_{N,t}(\boldsymbol{m})\rangle\langle \boldsymbol{w}, \nabla \widetilde{H}_{N,t}(\boldsymbol{m})\rangle &= \langle \boldsymbol{v}, \boldsymbol{w}\rangle\xi_t'(q_{\boldsymbol{m}}) + N^{-1}\langle \boldsymbol{v}, \boldsymbol{m}\rangle\langle \boldsymbol{w}, \boldsymbol{m}\rangle\xi''(q_{\boldsymbol{m}}).
\end{aligned}$$

Thus, $\boldsymbol{v}$ must satisfy

$$\begin{aligned}
&N^{-1}\left(\langle \boldsymbol{u}^2, \boldsymbol{m}\rangle\boldsymbol{u}^1 + \langle \boldsymbol{u}^1, \boldsymbol{m}\rangle\boldsymbol{u}^2\right)\xi''(q_{\boldsymbol{m}}) + N^{-2}\langle \boldsymbol{u}^1, \boldsymbol{m}\rangle\langle \boldsymbol{u}^2, \boldsymbol{m}\rangle\xi^{(3)}(q_{\boldsymbol{m}})\boldsymbol{m} \\
&= \xi_t'(q_{\boldsymbol{m}})\boldsymbol{v} + N^{-1}\langle \boldsymbol{v}, \boldsymbol{m}\rangle\xi''(q_{\boldsymbol{m}})\boldsymbol{m}.
\end{aligned}$$

This has solution $\boldsymbol{v} = a_1\boldsymbol{u}^1 + a_2\boldsymbol{u}^2 + a_3\boldsymbol{m}$, where

$$\begin{aligned}
a_1 &= \frac{\xi''(q_{\boldsymbol{m}})}{N\xi_t'(q_{\boldsymbol{m}})}\langle \boldsymbol{u}^2, \boldsymbol{m}\rangle, \qquad a_2 = \frac{\xi''(q_{\boldsymbol{m}})}{N\xi_t'(q_{\boldsymbol{m}})}\langle \boldsymbol{u}^1, \boldsymbol{m}\rangle, \\
a_3 &= \frac{\langle \boldsymbol{u}^1, \boldsymbol{m}\rangle\langle \boldsymbol{u}^2, \boldsymbol{m}\rangle}{N^2(\xi_t'(q_{\boldsymbol{m}}) + q_{\boldsymbol{m}}\xi''(q_{\boldsymbol{m}}))}\left(\xi^{(3)}(q_{\boldsymbol{m}}) - \frac{2\xi''(q_{\boldsymbol{m}})^2}{\xi_t'(q_{\boldsymbol{m}})}\right).
\end{aligned}$$

Thus

$$\mathbb{E}\left[\langle \nabla^2 \widetilde{H}_{N,t}(\boldsymbol{m})\boldsymbol{u}^1, \boldsymbol{u}^2\rangle\middle|\nabla \widetilde{H}_{N,t}(\boldsymbol{m})\right] = a_1\langle \boldsymbol{u}^1, \nabla \widetilde{H}_{N,t}(\boldsymbol{m})\rangle + a_2\langle \boldsymbol{u}^2, \nabla \widetilde{H}_{N,t}(\boldsymbol{m})\rangle + a_3\langle \boldsymbol{m}, \nabla \widetilde{H}_{N,t}(\boldsymbol{m})\rangle,$$

which implies

$$\begin{aligned}
\mathbb{E}\left[\nabla^2 \widetilde{H}_{N,t}(\boldsymbol{m})\middle|\nabla \widetilde{H}_{N,t}(\boldsymbol{m})\right] &= \frac{\xi''(q_{\boldsymbol{m}})}{N\xi_t'(q_{\boldsymbol{m}})}(\boldsymbol{m}\nabla \widetilde{H}_{N,t}(\boldsymbol{m})^\top + \nabla \widetilde{H}_{N,t}(\boldsymbol{m})\boldsymbol{m}^\top) \\
&\quad + \frac{\langle \boldsymbol{m}, \nabla \widetilde{H}_{N,t}(\boldsymbol{m})\rangle_N}{N(\xi_t'(q_{\boldsymbol{m}}) + q_{\boldsymbol{m}}\xi''(q_{\boldsymbol{m}}))}\left(\xi^{(3)}(q_{\boldsymbol{m}}) - \frac{2\xi''(q_{\boldsymbol{m}})^2}{\xi_t'(q_{\boldsymbol{m}})}\right)\boldsymbol{m}\boldsymbol{m}^\top,
\end{aligned}$$

as desired. The conditionally random part of $\nabla^2 \widetilde{H}_{N,t}$ is thus

$$\boldsymbol{M} = \nabla^2 \widetilde{H}_{N,t}(\boldsymbol{m}) - \mathbb{E}\left[\nabla^2 \widetilde{H}_{N,t}(\boldsymbol{m})\middle|\nabla \widetilde{H}_{N,t}(\boldsymbol{m})\right].$$

Direct evaluation of covariances $\mathbb{E}\,\boldsymbol{M}(i_1, j_1)\boldsymbol{M}(i_2, j_2)$ gives the covariance structure (6.80). The calculation is greatly simplified by the fact that $\langle \boldsymbol{e}_i, \boldsymbol{m}\rangle = 0$ for all $i \neq 1$, which implies e.g. that $\boldsymbol{M}(i,j) = \langle \nabla^2 \widetilde{H}_{N,t}(\boldsymbol{m})\boldsymbol{e}_i, \boldsymbol{e}_j\rangle$ for all $i, j \neq 1$. ☐

Corollary 6.6.14. *Let $\iota > 0$ be sufficiently small. Let $\boldsymbol{m}, \boldsymbol{z} \in \mathbb{R}^N$ with $\|\boldsymbol{m}\|_N^2 = q_{\boldsymbol{m}}$ and $\langle \boldsymbol{m}, \boldsymbol{x}\rangle_N = q_{\boldsymbol{x}}$, such that $|q_{\boldsymbol{m}} - q_*|, |q_{\boldsymbol{x}} - q_*|, \|\boldsymbol{z}\|_N \le \iota$. Conditional on $\nabla \mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m}) = \boldsymbol{z}$,*

$$\begin{aligned}
\nabla^2 \mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m}) &\stackrel{d}{=} \left(\frac{2+q_*}{q_*}\xi''(q_*) - (1-q_*)\xi^{(3)}(q_*) - \frac{2}{(1-q_*)^2}\right)\frac{\boldsymbol{m}\boldsymbol{m}^\top}{N} + \xi''(q_*)\frac{(\boldsymbol{x}-\boldsymbol{m})(\boldsymbol{x}-\boldsymbol{m})^\top}{N} \\
&\quad - \left((1-q_{\boldsymbol{m}})\xi''(q_{\boldsymbol{m}}) + \frac{1}{1-q_{\boldsymbol{m}}}\right)\boldsymbol{I} + \boldsymbol{M} + \boldsymbol{E}. \qquad (6.81)
\end{aligned}$$

Here, $\boldsymbol{M}$ is as in (6.80), and $\boldsymbol{E}$ is a $(\boldsymbol{x}, \boldsymbol{m}, \boldsymbol{z})$-measurable symmetric matrix satisfying $\|\boldsymbol{E}\|_{\mathsf{op}} \le o_\iota(1)$, whose kernel contains $\mathsf{span}(\boldsymbol{x}, \boldsymbol{m}, \boldsymbol{z})^\perp$.

Proof. In the below calculations, $\boldsymbol{E}$ is an error term satisfying the above, which may change from line to line. Conditioning on $\nabla \mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m}) = \boldsymbol{z}$ is equivalent to conditioning on

$$\nabla \widetilde{H}_{N,t}(\boldsymbol{m}) = \widetilde{\boldsymbol{z}} \equiv \boldsymbol{z} - \xi_t'(q_{\boldsymbol{x}})\boldsymbol{x} + \left((1-q_{\boldsymbol{m}})\xi''(q_{\boldsymbol{m}}) + \frac{1}{1-q_{\boldsymbol{m}}}\right)\boldsymbol{m}.$$

By Lemma 6.6.12,

$$
\begin{aligned}
\nabla^2 \widetilde{H}_{N,t}(\boldsymbol{m}) &\overset{d}{=} \frac{\xi''(q_{\boldsymbol{m}})}{\xi_t'(q_{\boldsymbol{m}})} \cdot \frac{\boldsymbol{m}\widetilde{\boldsymbol{z}}^\top + \widetilde{\boldsymbol{z}}\boldsymbol{m}^\top}{N} + \frac{\langle \boldsymbol{m}, \widetilde{\boldsymbol{z}} \rangle_N}{\xi_t'(q_{\boldsymbol{m}}) + q_{\boldsymbol{m}}\xi''(q_{\boldsymbol{m}})} \left(\xi^{(3)}(q_{\boldsymbol{m}}) - \frac{2\xi''(q_{\boldsymbol{m}})^2}{\xi_t'(q_{\boldsymbol{m}})} \right) \frac{\boldsymbol{m}\boldsymbol{m}^\top}{N} + \boldsymbol{M} \\
&= -\xi''(q_*) \frac{\boldsymbol{m}\boldsymbol{x}^\top + \boldsymbol{x}\boldsymbol{m}^\top}{N} + \frac{2\xi''(q_*)}{\xi_t'(q_*)} \left((1-q_*)\xi''(q_*) + \frac{1}{1-q_*} \right) \frac{\boldsymbol{m}\boldsymbol{m}^\top}{N} \\
&\quad + \frac{q_*(-\xi_t'(q_*) + (1-q_*)\xi''(q_*) + \frac{1}{1-q_*})}{\xi_t'(q_*) + q_*\xi''(q_*)} \left(\xi^{(3)}(q_*) - \frac{2\xi''(q_*)^2}{\xi_t'(q_*)} \right) \frac{\boldsymbol{m}\boldsymbol{m}^\top}{N} + \boldsymbol{M} + \boldsymbol{E} \\
&\overset{(6.39)}{=} -\xi''(q_*) \frac{\boldsymbol{m}\boldsymbol{x}^\top + \boldsymbol{x}\boldsymbol{m}^\top}{N} + \left(\frac{2\xi''(q_*)}{q_*} + (1-q_*)\xi^{(3)}(q_*) \right) \frac{\boldsymbol{m}\boldsymbol{m}^\top}{N} + \boldsymbol{M} + \boldsymbol{E}.
\end{aligned}
$$

Then

$$
\begin{aligned}
\nabla^2 \mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m}) &= \nabla^2 \widetilde{H}_{N,t}(\boldsymbol{m}) + \xi''(q_*) \frac{\boldsymbol{x}\boldsymbol{x}^\top}{N} - \left((1-q_*)\xi^{(3)}(q_*) - \xi''(q_*) + \frac{1}{(1-q_*)^2} \right) \frac{2\boldsymbol{m}\boldsymbol{m}^\top}{N} \\
&\quad - \left((1-q_{\boldsymbol{m}})\xi''(q_{\boldsymbol{m}}) + \frac{1}{1-q_{\boldsymbol{m}}} \right) \boldsymbol{I} + \boldsymbol{E}.
\end{aligned}
$$

Combining gives the conclusion. □

Lemma 6.6.15. *Let $\iota > 0$ be sufficiently small and $|q_{\boldsymbol{m}} - q_*| \le \iota$. Fix an orthonormal basis $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_N$ of $\mathbb{R}^N$ as discussed above (6.80). Let $\boldsymbol{M}$ be as in (6.80). Let $\boldsymbol{M}_*$ be sampled from the same law, except with $q_{\boldsymbol{m}}$ replaced by q_*, and with*

$$\boldsymbol{M}_*(i,j) = 0, \qquad \forall i,j \in \{1,2\}.$$

There is a coupling of $\boldsymbol{M}, \boldsymbol{M}_*$ such that with probability $1 - e^{-cN}$, $\|\boldsymbol{M} - \boldsymbol{M}_*\|_{\mathsf{op}} \le o_\iota(1)$.

Proof. Let $\boldsymbol{M}'$ be the matrix with $\boldsymbol{M}'(i,j) = 0$ for all $i,j \in \{1,2\}$, and otherwise $\boldsymbol{M}'(i,j) = \boldsymbol{M}(i,j)$. Since the $\boldsymbol{M}(i,j)$ have variance $O(N^{-1})$, with probability $1 - e^{-cN}$, $\|\boldsymbol{M} - \boldsymbol{M}'\|_{\mathsf{op}} \le \iota$.

For all $(i,j) \notin \{1,2\}^2$, $|\mathbb{E}\,\boldsymbol{M}(i,j)^2 - \mathbb{E}\,\boldsymbol{M}_*(i,j)^2| \le O(\iota)/N$. We couple $\boldsymbol{M}$ and $\boldsymbol{M}_*$ as follows. If $\mathbb{E}\,\boldsymbol{M}(i,j)^2 \le \mathbb{E}\,\boldsymbol{M}_*(i,j)^2$, we first sample $\boldsymbol{M}(i,j)$ from its law, and then sample

$$\boldsymbol{M}_*(i,j) = \boldsymbol{M}(i,j) + \upsilon_{i,j} g_{i,j},$$

for $g_{i,j} \sim \mathcal{N}(0, 1/N)$ and suitable $\upsilon_{i,j} = O(\iota^{1/2})$. If $\mathbb{E}\,\boldsymbol{M}(i,j)^2 \ge \mathbb{E}\,\boldsymbol{M}_*(i,j)^2$, we follow a similar procedure, sampling $\boldsymbol{M}_*(i,j)$ first. Let $\boldsymbol{E} = \boldsymbol{M}' - \boldsymbol{M}_*$. Then

$$\boldsymbol{E}(i,j) = (\varepsilon_{i,j}\upsilon_{i,j}g_{i,j})_{i,j\in[N]},$$

for some (deterministic) signs $\varepsilon_{i,j} \in \{\pm 1\}$. Let $\upsilon = \max_{i,j}(\upsilon_{i,j})$. There exists a random symmetric Gaussian matrix $\boldsymbol{E}'$, independent of $\boldsymbol{E}$, such that $\boldsymbol{E} + \boldsymbol{E}' =_d \upsilon \boldsymbol{G}$, where $\boldsymbol{G} \sim \mathsf{GOE}(N)$. Define

$$\mathcal{K} = \left\{ \boldsymbol{A} \in \mathbb{R}^{N\times N} \text{ symmetric} : \|\boldsymbol{A}\|_{\mathsf{op}} \le 3\upsilon \right\},$$

Note that

$$\mathbb{P}(\boldsymbol{E} + \boldsymbol{E}' \notin \mathcal{K}) = \mathbb{P}(\|\boldsymbol{G}\|_{\mathsf{op}} > 3) \le e^{-cN},$$

while by convexity of $\mathcal{K}$ and symmetry of $\boldsymbol{E}'$,

$$\mathbb{P}(\boldsymbol{E} + \boldsymbol{E}' \notin \mathcal{K} | \boldsymbol{E} \notin \mathcal{K}) \ge \frac{1}{2}.$$

It follows that $\mathbb{P}(\boldsymbol{E} \notin \mathcal{K}) \le 2e^{-cN}$, concluding the proof. □

Lemma 6.6.16. *Let $\boldsymbol{G} \sim \mathsf{GOE}(N)$ and $\boldsymbol{g} \sim \mathcal{N}(0, \boldsymbol{I}_N/N)$. For any $a, b \in \mathbb{R}$, $\iota > 0$,*

$$\left| \sup_{\boldsymbol{v}\in\mathbb{S}^{N-1}} \{a\langle \boldsymbol{G}\boldsymbol{v}, \boldsymbol{v}\rangle + 2b\langle \boldsymbol{g}, \boldsymbol{v}\rangle\} - 2\sqrt{a^2+b^2} \right| \le \iota$$

with probability $1 - e^{-cN}$.

Proof. By [CS17, Proposition 1.1],

$$\operatorname*{p-lim}_{N\to\infty} \sup_{\|\boldsymbol{v}\|_2=1} a\langle \boldsymbol{G}\boldsymbol{v},\boldsymbol{v}\rangle + 2b\langle \boldsymbol{g},\boldsymbol{v}\rangle = 2\sqrt{a^2+b^2}.$$

For each fixed $\boldsymbol{v}\in\mathbb{S}^{N-1}$, $a\langle \boldsymbol{G}\boldsymbol{v},\boldsymbol{v}\rangle + 2b\langle \boldsymbol{g},\boldsymbol{v}\rangle$ has variance $O(N^{-1})$. The result follows from Borell-TIS. $\square$

Proposition 6.6.17. *Let $\nabla^2\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m})$ be as in Eq. (6.81). There exist $C^{\mathsf{spec}}_{\max} > C^{\mathsf{spec}}_{\min} > 0$ such that for sufficiently small $\iota>0$,*

$$\mathsf{spec}(\nabla^2\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m})) \subseteq [-C^{\mathsf{spec}}_{\max}, -C^{\mathsf{spec}}_{\min}]$$

with probability $1-e^{-cN}$.

Proof. Let $\widetilde{\boldsymbol{m}} = \boldsymbol{m}/\|\boldsymbol{m}\|_2$, $\widetilde{\boldsymbol{x}} = P^{\perp}_{\boldsymbol{m}}\boldsymbol{x}/\|P^{\perp}_{\boldsymbol{m}}\boldsymbol{x}\|_2$. Throughout this proof, we will denote by $\boldsymbol{E}$, $\boldsymbol{E}_1$, $\boldsymbol{E}_2$, and so on error terms with the same meaning as in Corollary 6.6.14, namely $(\boldsymbol{x},\boldsymbol{m},\boldsymbol{z})$-measurable symmetric matrices satisfying $\|\boldsymbol{E}\|_{\mathsf{op}} \le o_\iota(1)$, whose kernel contains $\mathsf{span}(\boldsymbol{x},\boldsymbol{m},\boldsymbol{z})^\perp$. In particular

$$q_*\widetilde{\boldsymbol{m}}\widetilde{\boldsymbol{m}}^\top - \frac{\boldsymbol{m}\boldsymbol{m}^\top}{N} =: \boldsymbol{E}_1,\quad (1-q_*)\widetilde{\boldsymbol{x}}\widetilde{\boldsymbol{x}}^\top - \frac{(\boldsymbol{x}-\boldsymbol{m})(\boldsymbol{x}-\boldsymbol{m})^\top}{N} =: \boldsymbol{E}_2.$$

Let $\boldsymbol{e}_1,\dots,\boldsymbol{e}_N$ be an orthonormal basis of $\mathbb{R}^N$ with $\boldsymbol{e}_1 = \widetilde{\boldsymbol{m}}$, $\boldsymbol{e}_2 = \widetilde{\boldsymbol{x}}$. Let $\boldsymbol{M}_*$ be defined in Lemma 6.6.15, coupled to $\boldsymbol{M}$ so that $\|\boldsymbol{M}-\boldsymbol{M}_*\|_{\mathsf{op}} \le o_\iota(1)$ with probability $1-e^{-cN}$. Taking ι small, it suffices to show

$$-C^{\mathsf{spec}}_{\max}\boldsymbol{I} \preceq \boldsymbol{A} \preceq -C^{\mathsf{spec}}_{\min}\boldsymbol{I}, \tag{6.82}$$

for

$$\begin{aligned}\boldsymbol{A} = &\left((2+q_*)\xi''(q_*) - q_*(1-q_*)\xi^{(3)}(q_*) - \frac{2q_*}{(1-q_*)^2}\right)\widetilde{\boldsymbol{m}}\widetilde{\boldsymbol{m}}^\top + (1-q_*)\xi''(q_*)\widetilde{\boldsymbol{x}}\widetilde{\boldsymbol{x}}^\top \\ &- \left((1-q_*)\xi''(q_*) + \frac{1}{1-q_*}\right)\boldsymbol{I} + \boldsymbol{M}_*.\end{aligned}$$

By comparing $\boldsymbol{M}_*$ to a large constant multiple of a GOE, identically to the proof of Lemma 6.6.15, we can show $\|\boldsymbol{M}_*\|_{\mathsf{op}} = O(1)$ with probability $1-e^{-cN}$. On this event, all terms in $\boldsymbol{A}$ have bounded operator norm, and thus $-C^{\mathsf{spec}}_{\max}\boldsymbol{I} \preceq \boldsymbol{A}$. For the upper bound in (6.82), let

$$\psi = \xi''(q_*) + q_*\xi^{(3)}(q_*) - \frac{q_*\xi''(q_*)^2}{\xi'_t(q_*)}.$$

which (recall Remark 6.6.13) is nonnegative. Then

$$\begin{aligned}\boldsymbol{A} = &\left(-(1-q_*)\psi + q_*\left(\xi''(q_*) - \frac{1}{(1-q_*)^2}\right) - (1-q_*)^2\left(\xi''(q_*) - \frac{1}{(1-q_*)^2}\right)^2\right)\boldsymbol{e}_1\boldsymbol{e}_1^\top \\ &- \frac{1}{1-q_*}\boldsymbol{e}_2\boldsymbol{e}_2^\top - \left((1-q_*)\xi''(q_*) + \frac{1}{1-q_*}\right)\sum_{i=3}^N \boldsymbol{e}_i\boldsymbol{e}_i^\top + \boldsymbol{M}_*\end{aligned}$$

By (6.5), there exists $c_0>0$ depending only on ξ such that

$$\begin{aligned}&\boldsymbol{A} \preceq \boldsymbol{A}' - c_0(\boldsymbol{e}_1\boldsymbol{e}_1^\top + \boldsymbol{e}_2\boldsymbol{e}_2^\top), \qquad \text{where} \\ &\boldsymbol{A}' = -(1-q_*)\psi\boldsymbol{e}_1\boldsymbol{e}_1^\top - (1-q_*)\xi''(q_*)\boldsymbol{e}_2\boldsymbol{e}_2^\top - \left((1-q_*)\xi''(q_*) + \frac{1}{1-q_*}\right)\sum_{i=3}^N \boldsymbol{e}_i\boldsymbol{e}_i^\top + \boldsymbol{M}_* \\ &\quad = -(1-q_*)\psi\boldsymbol{e}_1\boldsymbol{e}_1^\top - (1-q_*)\xi''(q_*)\sum_{i=2}^N \boldsymbol{e}_i\boldsymbol{e}_i^\top - \frac{1}{1-q_*}\sum_{i=3}^N \boldsymbol{e}_i\boldsymbol{e}_i^\top + \boldsymbol{M}_*.\end{aligned}$$

By (6.80), (with $\boldsymbol{M}_*(i,j)$ having the same meaning as above)

$$\begin{aligned}
(\boldsymbol{M}_*(1,i) : 3 \le i \le N) &\overset{d}{=} \sqrt{\psi}\boldsymbol{g}^1, & \boldsymbol{g}^1 &\sim \mathcal{N}(0, \boldsymbol{I}_{N-2}/N),\\
(\boldsymbol{M}_*(2,i) : 3 \le i \le N) &\overset{d}{=} \sqrt{\xi''(q_*)}\boldsymbol{g}^2, & \boldsymbol{g}^2 &\sim \mathcal{N}(0, \boldsymbol{I}_{N-2}/N),\\
(\boldsymbol{M}_*(i,j) : 3 \le i,j \le N) &\overset{d}{=} \sqrt{\xi''(q_*)\cdot\frac{N-2}{N}}\boldsymbol{G} & \boldsymbol{G} &\sim \mathsf{GOE}(N-2),
\end{aligned}$$

and $\boldsymbol{g}^1, \boldsymbol{g}^2, \boldsymbol{G}$ are independent. Fix a, b with $a^2+b^2 \le 1$ and consider temporarily the restricted set

$$\mathbb{S}^{N-1}_{a,b} = \left\{\boldsymbol{v} \in \mathbb{S}^{N-1} : \langle \boldsymbol{v}, \boldsymbol{e}_1\rangle = a, \langle \boldsymbol{v}, \boldsymbol{e}_2\rangle = b\right\}.$$

For any $\boldsymbol{v} \in \mathbb{S}^{N-1}_{a,b}$ we can write

$$\boldsymbol{v} = a\boldsymbol{e}_1 + b\boldsymbol{e}_2 + \sqrt{1-a^2-b^2}\boldsymbol{w},$$

where $\boldsymbol{w} \in \mathbb{S}^{N-1}_{0,0}$. Because we defined $\boldsymbol{M}_*(i,j) = 0$ for all $i,j \in \{1,2\}$,

$$\begin{aligned}
\langle \boldsymbol{M}_*\boldsymbol{v}, \boldsymbol{v}\rangle &= 2a\sqrt{(1-a^2-b^2)\psi}\langle \boldsymbol{g}^1, \boldsymbol{w}\rangle + 2b\sqrt{(1-a^2-b^2)\xi''(q_*)}\langle \boldsymbol{g}^2, \boldsymbol{w}\rangle\\
&\quad + (1-a^2-b^2)\sqrt{\xi''(q_*)\cdot\frac{N-2}{N}}\langle \boldsymbol{G}\boldsymbol{w}, \boldsymbol{w}\rangle.
\end{aligned}$$

By Lemma 6.6.16, with probability $1-e^{-cN}$,

$$\left|\sup_{\boldsymbol{v}\in\mathbb{S}^{N-1}_{a,b}} \langle \boldsymbol{M}_*\boldsymbol{v}, \boldsymbol{v}\rangle - 2\sqrt{f(a,b)}\right| \le \iota,$$

where

$$\begin{aligned}
f(a,b) &= (1-a^2-b^2)^2\xi''(q_*) + a^2(1-a^2-b^2)\psi + b^2(1-a^2-b^2)\xi''(q_*)\\
&= (1-a^2-b^2)\left((1-a^2)\xi''(q_*) + a^2\psi\right).
\end{aligned}$$

On this event, for all $\boldsymbol{v} \in \mathbb{S}^{N-1}_{a,b}$,

$$\begin{aligned}
\langle \boldsymbol{A}\boldsymbol{v}, \boldsymbol{v}\rangle &\le -(1-q_*)\psi a^2 - (1-q_*)\xi''(q_*)(1-a^2) - \frac{1-a^2-b^2}{1-q_*} + 2\sqrt{f(a,b)} - c_0(a^2+b^2) + \iota\\
&= -\left(\sqrt{(1-q_*)\left((1-a^2)\xi''(q_*) + a^2\psi\right)} - \sqrt{\frac{1-a^2-b^2}{1-q_*}}\right)^2 - c_0(a^2+b^2) + \iota.
\end{aligned}$$

At $a = b = 0$, the first term is strictly negative by (6.5). So, there exists $c_1 > 0$, depending only on ξ, such that for all $a^2+b^2 \le 1$,

$$-\left(\sqrt{(1-q_*)\left((1-a^2)\xi''(q_*) + a^2\psi\right)} - \sqrt{\frac{1-a^2-b^2}{1-q_*}}\right)^2 - c_0(a^2+b^2) \le -c_1.$$

We have thus shown that, for fixed a, b, with probability $1-e^{-cN}$,

$$\sup_{\boldsymbol{v}\in\mathbb{S}^{N-1}_{a,b}} \langle \boldsymbol{A}\boldsymbol{v}, \boldsymbol{v}\rangle \le -c_1 + \iota. \tag{6.83}$$

Recall that $\|\boldsymbol{A}\|_{\mathsf{op}} = O(1)$ with probability $1-e^{-cN}$. So, the map

$$(a,b) \mapsto \sup_{\boldsymbol{v}\in\mathbb{S}^{N-1}_{a,b}} \langle \boldsymbol{A}\boldsymbol{v}, \boldsymbol{v}\rangle$$

is $O(1)$-Lipschitz. By a union bound, with proability $1-e^{-cN}$ (6.83) holds for all (a,b) in a ι-net of $a^2+b^2 \le 1$. On this event,

$$\sup_{\boldsymbol{v}\in\mathbb{S}^{N-1}} \langle \boldsymbol{A}\boldsymbol{v}, \boldsymbol{v}\rangle \le -c_1 + O(\iota).$$

Taking $C^{\mathsf{spec}}_{\min} = c_1/2$ and ι small enough completes the proof. $\square$

Proof of Proposition 6.6.9. Let ι' be given by Proposition 6.6.11. By this proposition, for any $\boldsymbol{m} \in \mathcal{S}_0$, $r \in I_\iota$,

$$\mathbb{P}\left(E_1(\boldsymbol{m},\iota')^c \middle| \nabla_{\mathsf{sp}}\widetilde{H}_{N,t}(\boldsymbol{m}) = \boldsymbol{0}, \widetilde{\partial}_{\mathsf{rad}}\widetilde{H}_{N,t}(\boldsymbol{m}) = r\right) \le e^{-cN}.$$

Since $\|\nabla\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m})\|_N \le \iota'$ on $E_1(\boldsymbol{m},\iota')$, and $\nabla_{\mathsf{sp}}\widetilde{H}_{N,t}(\boldsymbol{m}), \widetilde{\partial}_{\mathsf{rad}}\widetilde{H}_{N,t}(\boldsymbol{m})$ are $\nabla\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m})$-measurable,

$$\mathbb{P}\left(E_1(\boldsymbol{m},\iota') \cap E_2(\boldsymbol{m})^c \middle| \nabla_{\mathsf{sp}}\widetilde{H}_{N,t}(\boldsymbol{m}) = \boldsymbol{0}, \widetilde{\partial}_{\mathsf{rad}}\widetilde{H}_{N,t}(\boldsymbol{m}) = r\right) \le \sup_{\|\boldsymbol{z}\|_N \le \iota'} \mathbb{P}\left(E_2(\boldsymbol{m})^c \middle| \nabla\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m}) = \boldsymbol{z}\right).$$

By Corollary 6.6.14 and Proposition 6.6.17, this last probability is $\le e^{-cN}$. This completes the proof. $\square$

6.6.4 Proof of conditioning bound

Propositions 6.3.6 and 6.6.1 directly imply parts (a) and (b) of Proposition 6.4.4. We now prove the remainder of this proposition.

Proof of Proposition 6.4.4(c). Set $\upsilon > 0$ small enough that $\max(\upsilon, \iota'(\upsilon)) \le \iota/2$, for the function ι' from Proposition 6.6.1. Suppose K_N holds and the events in Propositions 6.4.3 and 6.6.1 hold with tolerance υ. This occurs with probability $1 - e^{-cN}$.

By (6.42) and (6.43), for suitably large K_{AMP}, $\boldsymbol{m}^{\mathsf{AMP}} \in \mathcal{S}_\upsilon \subseteq \mathcal{S}_{\iota/2}$ and $\boldsymbol{m}^{\mathsf{AMP}}$ is an υ-approximate critical point of $\mathcal{F}_{\mathsf{TAP}}$. By Proposition 6.6.1(b), this implies $\|\boldsymbol{m}^{\mathsf{AMP}} - \boldsymbol{m}^{\mathsf{TAP}}\|_N \le \iota/2$. $\square$

We now turn to the proof of part (d). Define

$$K(\boldsymbol{m}) = P^{\perp}_{\mathsf{span}(\boldsymbol{x},\boldsymbol{m})}\nabla^2\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m})P^{\perp}_{\mathsf{span}(\boldsymbol{x},\boldsymbol{m})}.$$

We will treat this as a $(N-2)\times(N-2)$ matrix, after a suitable change of coordinates. The following fact is a consequence of Corollary 6.6.14.

Fact 6.6.18. *Let $\|\boldsymbol{m}\|_N^2 = q_{\boldsymbol{m}} < 1$. Conditional on $\nabla\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m}) = \boldsymbol{0}$,*

$$K(\boldsymbol{m}) \overset{d}{=} -\left((1-q_{\boldsymbol{m}})\xi''(q_{\boldsymbol{m}}) + \frac{1}{1-q_{\boldsymbol{m}}}\right)\boldsymbol{I} + \sqrt{\xi''(q_{\boldsymbol{m}})\frac{N-2}{N}}\boldsymbol{G}, \qquad \boldsymbol{G} \sim \mathsf{GOE}(N-2).$$

The next fact is verified by direct calculation.

Fact 6.6.19. *For any $\boldsymbol{m}$, $\nabla\widetilde{H}_{N,t}(\boldsymbol{m})$ is Gaussian, with variance $\xi_t'(q_{\boldsymbol{m}}) + q_{\boldsymbol{m}}\xi''(q_{\boldsymbol{m}})$ in the direction of $\boldsymbol{m}$ and $\xi_t'(q_{\boldsymbol{m}})$ in all directions orthogonal to $\boldsymbol{m}$.*

We will need the following technical lemma, which we prove in Subsection 6.6.5.

Lemma 6.6.20. *For all $\iota > 0$ sufficiently small, there exists a constant $C > 0$ such that*

$$\int_{\mathcal{S}_\iota} \mathbb{E}\left[|\det K(\boldsymbol{m})| \middle| \nabla\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m}) = \boldsymbol{0}\right] \varphi_{\nabla\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m})}(\boldsymbol{0})\ \mathsf{d}^N(\boldsymbol{m}) \le C.$$

Proof of Proposition 6.4.4(d). We will apply Lemma 6.3.7 with $\mathcal{F}_{\mathsf{TAP}}$ for $\mathcal{F}$, $\boldsymbol{m}^{\mathsf{AMP}}$ for $\boldsymbol{m}_0$, the interior of $\mathcal{S}_\iota$ for D, and $C^{\mathsf{spec}}_{\min}/2$ for c_{spec}. Note that (6.47) implies $\varepsilon \le c^2_{\mathsf{spec}}/10c_{\mathsf{op}}$. We next verify that the event $\mathcal{E}$ is contained in the event $\mathcal{E}_{\mathsf{cond}}$ defined in Lemma 6.3.7. Suppose $\mathcal{E}$ holds. Then event $\mathcal{H}(c_{\mathsf{op}})$ holds by part (a). $\|\nabla\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m}^{\mathsf{AMP}})\|_N \le \varepsilon$ by definition, and by parts (a), (b), and (c),

$$\lambda_{\max}(\nabla^2\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m}^{\mathsf{AMP}})) \le \lambda_{\max}(\nabla^2\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m}^{\mathsf{TAP}})) + c_{\mathsf{op}}\|\boldsymbol{m}^{\mathsf{AMP}} - \boldsymbol{m}^{\mathsf{TAP}}\|_{\mathsf{op},N} \le -C^{\mathsf{spec}}_{\min} + c_{\mathsf{op}}\iota/2 \le -C^{\mathsf{spec}}_{\min}/2$$

for small enough ι. Thus $\mathcal{G}(\varepsilon, c_{\mathsf{spec}})$ holds. We have $\|\boldsymbol{m}^{\mathsf{AMP}}\|_N \le 1$ because $\boldsymbol{m}^{\mathsf{AMP}} \in \mathcal{S}_{\iota/2}$, by part (c). Also, (6.47) implies $\frac{5\varepsilon}{c_{\mathsf{spec}}} \le \frac{\iota}{2}$, so $U = \mathsf{B}^N(\boldsymbol{m}^{\mathsf{AMP}}, 5\varepsilon/c_{\mathsf{spec}}) \subseteq \mathcal{S}_\iota$.

By (6.56), $\mathbb{E}\,\nabla\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m})$ is continuous in $\boldsymbol{m}$, and by Fact 6.6.19, $\mathsf{Cov}(\nabla\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m}))$ is uniformly lower bounded for all $\boldsymbol{m} \in \mathcal{S}_\iota$. This verifies the regularity condition in Lemma 6.3.7. By this lemma,

$$\mathbb{E}[X\mathbf{1}\{\mathcal{E}\}] \le \int_{S_\iota} \mathbb{E}\left[|\det\nabla^2\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m})|X\mathbf{1}\{\mathcal{E}\} \middle| \nabla\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m}) = \boldsymbol{0}\right] \varphi_{\nabla\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m})}(\boldsymbol{0})\ \mathsf{d}^N(\boldsymbol{m}).$$

By Hölder's inequality,

$$\begin{aligned}&\mathbb{E}\left[|\det\nabla^2\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m})|X\mathbf{1}\{\mathcal{E}\}\middle|\nabla\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m})=\mathbf{0}\right]\\&\le\mathbb{E}\left[|\det\nabla^2\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m})|^{1+\delta^{-1}}\mathbf{1}\{\mathcal{E}\}\middle|\nabla\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m})=\mathbf{0}\right]^{\delta/(1+\delta)}\mathbb{E}\left[X^{1+\delta}\mathbf{1}\{\mathcal{E}\}\middle|\nabla\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m})=\mathbf{0}\right]^{1/(1+\delta)}.\end{aligned}$$

On event $\mathcal{E}$, the eigenvalues of $\nabla^2\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m})$ lie in $[-C^{\mathsf{spec}}_{\max},-C^{\mathsf{spec}}_{\min}]$ and interlace those of $K(\boldsymbol{m})$. So,

$$|\det\nabla^2\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m})|\le(C^{\mathsf{spec}}_{\max})^2|\det K(\boldsymbol{m})|.$$

Thus,

$$\begin{aligned}&\mathbb{E}\left[|\det\nabla^2\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m})|^{1+\delta^{-1}}\mathbf{1}\{\mathcal{E}\}\middle|\nabla\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m})=\mathbf{0}\right]^{\delta/(1+\delta)}\\&\le(C^{\mathsf{spec}}_{\max})^2\,\mathbb{E}\left[|\det K(\boldsymbol{m})|^{1+\delta^{-1}}\middle|\nabla\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m})=\mathbf{0}\right]^{\delta/(1+\delta)}\\&\le(C^{\mathsf{spec}}_{\max})^2C'_\delta\,\mathbb{E}\left[|\det K(\boldsymbol{m})|\middle|\nabla\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m})=\mathbf{0}\right].\end{aligned}$$

for some $C'_\delta>0$. The last estimate is by Fact 6.6.18, (6.5), and Lemma 6.6.6, similarly to (6.79). Combining,

$$\begin{aligned}\mathbb{E}[X\mathbf{1}\{\mathcal{E}\}]\le(C^{\mathsf{spec}}_{\max})^2C'_\delta\int_{\mathcal{S}_\iota}&\mathbb{E}\left[|\det K(\boldsymbol{m})|\middle|\nabla\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m})=\mathbf{0}\right]\varphi_{\nabla\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m})}(\mathbf{0})\,\mathsf{d}^N(\boldsymbol{m})\\&\times\sup_{\boldsymbol{m}\in\mathcal{S}_\iota}\mathbb{E}\left[X^{1+\delta}\mathbf{1}\{\mathcal{E}\}\middle|\nabla\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m})=\mathbf{0}\right]^{1/(1+\delta)}.\end{aligned}$$

Finally, by Lemma 6.6.20, this integral is bounded by a constant $C>0$. Thus the result holds with $C_\delta=(C^{\mathsf{spec}}_{\max})^2C'_\delta C$. □

6.6.5 Determinant concentration and estimate of Kac–Rice integral

In this subsection, we provide the deferred proofs of Lemmas 6.6.6 and 6.6.20. These are the final ingredients to the proof of Proposition 6.4.4.

Proof of Lemma 6.6.6. For any compact $K\subseteq(2,+\infty)$, we may pick $\varepsilon>0$ such small enough that $r\ge 2+2\varepsilon$ for all $r\in K$. Let $\mathcal{E}_\varepsilon$ be the event that $\|\boldsymbol{G}\|_{\mathsf{op}}\le 2+\varepsilon$. It is classical that $\mathbb{P}(\mathcal{E}_\varepsilon)\ge 1-e^{-cN}$. For $r\in K$, let

$$f(x)=\log\max(|r-x|,\varepsilon),$$

which is ε^{-1}-Lipschitz. Let $\lambda_1,\ldots,\lambda_N$ be the eigenvalues of $\boldsymbol{G}$ and define

$$\mathsf{Tr}f(\boldsymbol{G})=\sum_{i=1}^N f(\lambda_i).$$

By [GZ00, Theorem 1.1(b)], for all $s\ge 0$,

$$\mathbb{P}\left(|\mathsf{Tr}f(\boldsymbol{G})-\mathbb{E}\,\mathsf{Tr}f(\boldsymbol{G})|\ge s\right)\le 2\exp(-\varepsilon^2s^2/8).\tag{6.84}$$

Note that $|\det(r\boldsymbol{I}-\boldsymbol{G})|\le\exp(\mathsf{Tr}f(\boldsymbol{G}))$, and equality holds if $\boldsymbol{G}\in\mathcal{E}_\varepsilon$. Thus,

$$\mathbb{E}\left[|\det(r\boldsymbol{I}-\boldsymbol{G})|^t\right]\le\mathbb{E}\left[\exp(t\mathsf{Tr}f(\boldsymbol{G}))\right],\qquad\mathbb{E}\left[|\det(r\boldsymbol{I}-\boldsymbol{G})|\right]\ge\mathbb{E}\left[\exp(\mathsf{Tr}f(\boldsymbol{G}))\mathbf{1}\{\mathcal{E}_\varepsilon\}\right].$$

By (6.84), there exists $C_{\varepsilon,t}$ depending on ε,t such that

$$\mathbb{E}\left[\exp(t\mathsf{Tr}f(\boldsymbol{G}))\right]\le C_{\varepsilon,t}\exp(t\,\mathbb{E}\,\mathsf{Tr}f(\boldsymbol{G})).$$

By Cauchy–Schwarz,

$$\mathbb{E}\left[\exp(\mathsf{Tr}f(\boldsymbol{G}))\mathbf{1}\{\mathcal{E}_\varepsilon^c\}\right]\le\mathbb{E}\left[\exp(2\mathsf{Tr}f(\boldsymbol{G}))\right]^{1/2}\mathbb{P}(\mathcal{E}_\varepsilon^c)^{1/2}\le C_{\varepsilon,2}^{1/2}e^{-cN/2}\exp(\mathbb{E}\,\mathsf{Tr}f(\boldsymbol{G})),$$

which implies

$$\begin{aligned}\mathbb{E}\left[\exp(\mathsf{Tr}f(\boldsymbol{G}))\mathbf{1}\{\mathcal{E}_\varepsilon\}\right] &\ge \mathbb{E}\left[\exp(\mathsf{Tr}f(\boldsymbol{G}))\right] - C_{\varepsilon,2}^{1/2}e^{-cN/2}\exp(\mathbb{E}\,\mathsf{Tr}f(\boldsymbol{G})) \\ &\ge (1-C_{\varepsilon,2}^{1/2}e^{-cN/2})\exp(\mathbb{E}\,\mathsf{Tr}f(\boldsymbol{G})).\end{aligned}$$

Thus,

$$\frac{\mathbb{E}\left[|\det(r\boldsymbol{I}-\boldsymbol{G})|^t\right]^{1/t}}{\mathbb{E}\left[|\det(r\boldsymbol{I}-\boldsymbol{G})|\right]} \le \frac{C_{\varepsilon,t}^{1/t}\exp(\mathbb{E}\,\mathsf{Tr}f(\boldsymbol{G}))}{(1-C_{\varepsilon,2}^{1/2}e^{-cN/2})\exp(\mathbb{E}\,\mathsf{Tr}f(\boldsymbol{G}))}$$

is bounded by a constant depending only on ε, t. □

Lemma 6.6.21. *Let $\boldsymbol{G} \sim \mathsf{GOE}(N)$. For all r in any compact subset of $(2,+\infty)$, there exists $C>0$ such that*

$$\mathbb{E}[|\det(r\boldsymbol{I}-\boldsymbol{G})|] \le C\exp(N\Phi(r)),$$

where

$$\Phi(r) = \frac{1}{4}r^2 - \frac{1}{2} - \frac{1}{4}r\sqrt{r^2-4} + \log\frac{r+\sqrt{r^2-4}}{2}.$$

Proof. Follows from [BČNS22, Lemma 2.1 and 2.2(i)] with $N+1$ for N and $\sqrt{\frac{N}{2(N+1)}}r$ for x. Note that the matrix $\mathsf{GOE}_{N-1}(N^{-1})$ therein is defined with typical spectral radius $\sqrt{2}$, while our $\mathsf{GOE}(N)$ has spectral radius 2. □

Proof of Lemma 6.6.20. Throughout this proof, $C>0$ is a constant, uniform over $\boldsymbol{m}\in\mathcal{S}_\iota$, which may change from line to line. Let $\|\boldsymbol{m}\|_N^2 = q_{\boldsymbol{m}}$ and $\langle \boldsymbol{x},\boldsymbol{m}\rangle_N = q_{\boldsymbol{x}}$, so that $q_{\boldsymbol{m}}, q_{\boldsymbol{x}} \in [q_*-\iota, q_*+\iota]$. By Fact 6.6.18, for $\boldsymbol{G}\sim\mathsf{GOE}(N-2)$,

$$\mathbb{E}\left[|\det K(\boldsymbol{m})|\big|\nabla\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m})=\boldsymbol{0}\right] = \left(\xi''(q_{\boldsymbol{m}})\cdot\frac{N-2}{N}\right)^{(N-2)/2}\mathbb{E}\left[\left|\det\left(\sqrt{\frac{N}{N-2}}r_{\boldsymbol{m}}\boldsymbol{I}-\boldsymbol{G}\right)\right|\right]$$
$$\text{where}\quad r_{\boldsymbol{m}} = (1-q_{\boldsymbol{m}})\xi''(q_{\boldsymbol{m}})^{1/2} + \frac{1}{(1-q_{\boldsymbol{m}})\xi''(q_{\boldsymbol{m}})^{1/2}}.$$

By (6.5), for $q_{\boldsymbol{m}}\in[q_*-\iota,q_*+\iota]$, $r_{\boldsymbol{m}}$ takes values in a compact subset of $(2,+\infty)$. By Lemma 6.6.21,

$$\mathbb{E}\left[|\det K(\boldsymbol{m})|\big|\nabla\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m})=\boldsymbol{0}\right] \le C\exp(Nf_1(q_{\boldsymbol{m}})), \qquad f_1(q_{\boldsymbol{m}}) := \frac{1}{2}\log\xi''(q_{\boldsymbol{m}}) + \Phi(r_{\boldsymbol{m}}).$$

By (6.5),

$$\sqrt{r_{\boldsymbol{m}}^2-4} = \frac{1}{(1-q_{\boldsymbol{m}})\xi''(q_{\boldsymbol{m}})^{1/2}} - (1-q_{\boldsymbol{m}})\xi''(q_{\boldsymbol{m}})^{1/2}.$$

So, f_1 simplifies to

$$f_1(q_{\boldsymbol{m}}) = \frac{1}{2}(1-q_{\boldsymbol{m}})^2\xi''(q_{\boldsymbol{m}}) - \log(1-q_{\boldsymbol{m}}).$$

On the other hand, by (6.56),

$$\nabla\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m}) = \nabla\widetilde{H}_{N,t}(\boldsymbol{m}) + \xi_t'(q_{\boldsymbol{x}})\left(\boldsymbol{x}-\frac{q_{\boldsymbol{x}}}{q_{\boldsymbol{m}}}\boldsymbol{m}\right) - \left((1-q_{\boldsymbol{m}})\xi''(q_{\boldsymbol{m}}) + \frac{1}{1-q_{\boldsymbol{m}}} - \frac{q_{\boldsymbol{x}}\xi_t'(q_{\boldsymbol{x}})}{q_{\boldsymbol{m}}}\right)\boldsymbol{m}.$$

Since $\boldsymbol{x}-\frac{q_{\boldsymbol{x}}}{q_{\boldsymbol{m}}}\boldsymbol{m}$ is orthogonal to $\boldsymbol{m}$, Fact 6.6.19 yields

$$\varphi_{\nabla\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m})}(\boldsymbol{0}) \le C\exp(Nf_2(q_{\boldsymbol{m}},q_{\boldsymbol{x}}))$$

where

$$\begin{aligned}f_2(q_{\boldsymbol{m}},q_{\boldsymbol{x}}) = &-\frac{1}{2}\log(2\pi\xi_t'(q_{\boldsymbol{m}})) - \frac{\xi_t'(q_{\boldsymbol{x}})^2}{2\xi_t'(q_{\boldsymbol{m}})}\left(1-\frac{q_{\boldsymbol{x}}^2}{q_{\boldsymbol{m}}}\right) \\ &-\frac{q_{\boldsymbol{m}}}{2(\xi_t'(q_{\boldsymbol{m}})+q_{\boldsymbol{m}}\xi''(q_{\boldsymbol{m}}))}\left((1-q_{\boldsymbol{m}})\xi''(q_{\boldsymbol{m}}) + \frac{1}{1-q_{\boldsymbol{m}}} - \frac{q_{\boldsymbol{x}}\xi_t'(q_{\boldsymbol{x}})}{q_{\boldsymbol{m}}}\right)^2.\end{aligned}$$

Combining the above,

$$\begin{aligned}&\int_{\mathcal{S}_\iota} \mathbb{E}\left[|\det K(\boldsymbol{m})|\middle|\nabla\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m})=\boldsymbol{0}\right]\varphi_{\nabla\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m})}(\boldsymbol{0})\ \mathsf{d}^N(\boldsymbol{m})\\ &\le CN\int_{q_*-\iota}^{q_*+\iota}\int_{q_*-\iota}^{q_*+\iota}\exp\left(N(f_1(q_{\boldsymbol{m}})+f_2(q_{\boldsymbol{m}},q_{\boldsymbol{x}})+f_3(q_{\boldsymbol{m}},q_{\boldsymbol{x}}))\right)\ \mathsf{d}q_{\boldsymbol{x}}\mathsf{d}q_{\boldsymbol{m}}. \end{aligned}\tag{6.85}$$

Here $CN\exp(Nf_3(q_{\boldsymbol{m}},q_{\boldsymbol{x}}))$ is a volumetric factor and

$$f_3(q_{\boldsymbol{m}},q_{\boldsymbol{x}})=\frac{1}{2}+\frac{1}{2}\log(2\pi(q_{\boldsymbol{m}}-q_{\boldsymbol{x}}^2)).$$

Let

$$\begin{aligned}F(q_{\boldsymbol{m}},q_{\boldsymbol{x}})&=f_1(q_{\boldsymbol{m}})+f_2(q_{\boldsymbol{m}},q_{\boldsymbol{x}})+f_3(q_{\boldsymbol{m}},q_{\boldsymbol{x}})\\ &=\frac{1}{2}+\frac{1}{2}\log\frac{q_{\boldsymbol{m}}-q_{\boldsymbol{x}}^2}{\xi_t'(q_{\boldsymbol{m}})(1-q_{\boldsymbol{m}})^2}+\frac{1}{2}(1-q_{\boldsymbol{m}})^2\xi''(q_{\boldsymbol{m}})-\frac{\xi_t'(q_{\boldsymbol{x}})^2}{2\xi_t'(q_{\boldsymbol{m}})}\left(1-\frac{q_{\boldsymbol{x}}^2}{q_{\boldsymbol{m}}}\right)\\ &\quad-\frac{q_{\boldsymbol{m}}}{2(\xi_t'(q_{\boldsymbol{m}})+q_{\boldsymbol{m}}\xi''(q_{\boldsymbol{m}}))}\left((1-q_{\boldsymbol{m}})\xi''(q_{\boldsymbol{m}})+\frac{1}{1-q_{\boldsymbol{m}}}-\frac{q_{\boldsymbol{x}}\xi_t'(q_{\boldsymbol{x}})}{q_{\boldsymbol{m}}}\right)^2.\end{aligned}$$

To conclude, we will verify that $F(q_*,q_*)=0$, $\nabla F(q_*,q_*)=0$, and F is $\Omega(1)$-strongly concave over $q_{\boldsymbol{m}},q_{\boldsymbol{x}}\in[q_*-\iota,q_*+\iota]$. This will imply that the integral in (6.85) is $O(N^{-1})$ and finish the proof.

Recall from (6.39) that $\xi_t'(q_*)=\frac{q_*}{1-q_*}$. The following identity will be used repeatedly in the calculations below to simplify the final term in F and its derivatives:

$$\frac{1}{\xi_t'(q_*)+q_*\xi''(q_*)}\left((1-q_*)\xi''(q_*)+\frac{1}{1-q_*}-\xi_t'(q_*)\right)=\frac{1-q_*}{q_*}.$$

Using this, we verify that

$$\begin{aligned}F(q_*,q_*)&=\frac{1}{2}+\frac{1}{2}(1-q_*)^2\xi''(q_*)-\frac{q_*}{2}-\frac{q_*}{2(\frac{q_*}{1-q_*}+q_*\xi''(q_*))}\left((1-q_*)\xi''(q_*)+1\right)^2\\ &=\frac{1}{2}(1-q_*)^2\xi''(q_*)+\frac{1-q_*}{2}-\frac{1-q_*}{2}\left((1-q_*)\xi''(q_*)+1\right)=0.\end{aligned}$$

We also calculate

$$\begin{aligned}\frac{\partial F}{\partial q_{\boldsymbol{m}}}(q_{\boldsymbol{m}},q_{\boldsymbol{x}})&=\frac{1}{2(q_{\boldsymbol{m}}-q_{\boldsymbol{x}}^2)}-\frac{\xi''(q_{\boldsymbol{m}})}{2\xi_t'(q_{\boldsymbol{m}})}+\frac{1}{1-q_{\boldsymbol{m}}}-(1-q_{\boldsymbol{m}})\xi''(q_{\boldsymbol{m}})+\frac{1}{2}(1-q_{\boldsymbol{m}})^2\xi^{(3)}(q_{\boldsymbol{m}})\\ &\quad+\frac{\xi_t'(q_{\boldsymbol{x}})^2\xi''(q_{\boldsymbol{m}})}{2\xi_t'(q_{\boldsymbol{m}})^2}\left(1-\frac{q_{\boldsymbol{x}}^2}{q_{\boldsymbol{m}}}\right)-\frac{\xi_t'(q_{\boldsymbol{x}})^2}{2\xi_t'(q_{\boldsymbol{m}})}\frac{q_{\boldsymbol{x}}^2}{q_{\boldsymbol{m}}^2}\\ &\quad-\frac{\xi_t'(q_{\boldsymbol{m}})-q_{\boldsymbol{m}}\xi''(q_{\boldsymbol{m}})-q_{\boldsymbol{m}}^2\xi^{(3)}(q_{\boldsymbol{m}})}{2(\xi_t'(q_{\boldsymbol{m}})+q_{\boldsymbol{m}}\xi''(q_{\boldsymbol{m}}))^2}\left((1-q_{\boldsymbol{m}})\xi''(q_{\boldsymbol{m}})+\frac{1}{1-q_{\boldsymbol{m}}}-\frac{q_{\boldsymbol{x}}\xi_t'(q_{\boldsymbol{x}})}{q_{\boldsymbol{m}}}\right)^2\\ &\quad-\frac{q_{\boldsymbol{m}}}{\xi_t'(q_{\boldsymbol{m}})+q_{\boldsymbol{m}}\xi''(q_{\boldsymbol{m}})}\left((1-q_{\boldsymbol{m}})\xi''(q_{\boldsymbol{m}})+\frac{1}{1-q_{\boldsymbol{m}}}-\frac{q_{\boldsymbol{x}}\xi_t'(q_{\boldsymbol{x}})}{q_{\boldsymbol{m}}}\right)\\ &\qquad\times\left(-\xi''(q_{\boldsymbol{m}})+(1-q_{\boldsymbol{m}})\xi^{(3)}(q_{\boldsymbol{m}})+\frac{1}{(1-q_{\boldsymbol{m}})^2}+\frac{q_{\boldsymbol{x}}\xi_t'(q_{\boldsymbol{x}})}{q_{\boldsymbol{m}}^2}\right),\\ \frac{\partial F}{\partial q_{\boldsymbol{x}}}(q_{\boldsymbol{m}},q_{\boldsymbol{x}})&=-\frac{q_{\boldsymbol{x}}}{q_{\boldsymbol{m}}-q_{\boldsymbol{x}}^2}-\frac{\xi_t'(q_{\boldsymbol{x}})\xi''(q_{\boldsymbol{x}})}{\xi_t'(q_{\boldsymbol{m}})}\left(1-\frac{q_{\boldsymbol{x}}^2}{q_{\boldsymbol{m}}}\right)+\frac{q_{\boldsymbol{x}}\xi_t'(q_{\boldsymbol{x}})^2}{q_{\boldsymbol{m}}\xi_t'(q_{\boldsymbol{m}})}\\ &\quad+\frac{\xi_t'(q_{\boldsymbol{x}})+q_{\boldsymbol{x}}\xi''(q_{\boldsymbol{x}})}{\xi_t'(q_{\boldsymbol{m}})+q_{\boldsymbol{m}}\xi''(q_{\boldsymbol{m}})}\left((1-q_{\boldsymbol{m}})\xi''(q_{\boldsymbol{m}})+\frac{1}{1-q_{\boldsymbol{m}}}-\frac{q_{\boldsymbol{x}}\xi_t'(q_{\boldsymbol{x}})}{q_{\boldsymbol{m}}}\right).\end{aligned}$$

Thus

$$\begin{aligned}\frac{\partial F}{\partial q_{\boldsymbol{m}}}(q_*,q_*) &= \frac{1}{2q_*(1-q_*)} - \frac{(1-q_*)\xi''(q_*)}{2q_*} + \frac{1}{1-q_*} - (1-q_*)\xi''(q_*) + \frac{1}{2}(1-q_*)^2\xi^{(3)}(q_*)\\ &\quad + \frac{(1-q_*)\xi''(q_*)}{2} - \frac{q_*}{2(1-q_*)} - \frac{(1-q_*)^2}{2q_*^2}\left(\frac{q_*}{1-q_*} - q_*\xi''(q_*) - q_*^2\xi^{(3)}(q_*)\right)\\ &\quad - (1-q_*)\left(-\xi''(q_*) + (1-q_*)\xi^{(3)}(q_*) + \frac{1}{(1-q_*)^2} + \frac{1}{1-q_*}\right) = 0,\end{aligned}$$

and

$$\frac{\partial F}{\partial q_{\boldsymbol{x}}}(q_*,q_*) = -\frac{1}{1-q_*} - (1-q_*)\xi''(q_*) + \xi_t'(q_*) + \left((1-q_*)\xi''(q_*) + \frac{1}{1-q_*} - \xi_t'(q_*)\right) = 0.$$

By similar calculations, we find the following formulas for the second derivative. Let

$$\Delta_0 = \xi''(q_*) - \frac{1}{(1-q_*)^2} \overset{(6.5)}{<} 0$$

and

$$\Delta_1 = \frac{(1-q_*)^3\Delta_0^3 - q_*(1-q_*)\Delta_0^2}{2q_*^2(1+(1-q_*)\xi''(q_*))}, \qquad \Delta_2 = -\frac{(1-q_*)^2}{q_*}\Delta_0^2 + \Delta_0.$$

Then

$$\frac{\partial^2 F}{\partial q_{\boldsymbol{m}}^2}(q_*,q_*) = \Delta_1 + \Delta_2, \qquad \frac{\partial^2 F}{\partial q_{\boldsymbol{m}}\partial q_{\boldsymbol{x}}}(q_*,q_*) = -\Delta_2, \qquad \frac{\partial^2 F}{\partial q_{\boldsymbol{x}}^2}(q_*,q_*) = \Delta_2.$$

It follows that

$$\nabla^2 F(q_*,q_*) = \Delta_1(1,0)^{\otimes 2} + \Delta_2(1,-1)^{\otimes 2} \preceq -CI_2$$

for some $C>0$ depending only on ξ. Since $\nabla^2 F$ is clearly locally Lipschitz around (q_*,q_*), $\nabla^2 F(q_{\boldsymbol{m}},q_{\boldsymbol{x}}) \preceq -CI_2/2$ for all $q_{\boldsymbol{m}},q_{\boldsymbol{x}} \in [q_*-\iota, q_*+\iota]$ for suitably small ι. This concludes the proof. □

6.6.6 Algorithmic guarantees and Lipschitz continuity of correction

In this subsection, we prove Proposition 6.4.5.

Proof of Proposition 6.4.5(a). By (6.44),

$$\mu_t(\mathsf{Band}(\boldsymbol{m}^{\text{AMP}}, [q_*-\iota/2, q_*+\iota/2])) \ge 1 - e^{-cN}.$$

Since $\|\boldsymbol{m}^{\text{AMP}} - \boldsymbol{m}^{\text{TAP}}\|_N \le \iota/2$, we have

$$\mathsf{Band}(\boldsymbol{m}^{\text{TAP}}, [q_*-\iota, q_*+\iota]) \supseteq \mathsf{Band}(\boldsymbol{m}^{\text{AMP}}, [q_*-\iota/2, q_*+\iota/2]).$$

□

Proof of Proposition 6.4.5(b). On K_N, the maps $\boldsymbol{m} \mapsto \lambda_{\max}(\nabla^2 \mathcal{F}_{\text{TAP}}(\boldsymbol{m}))$ and $\boldsymbol{m} \mapsto \lambda_{\min}(\nabla^2 \mathcal{F}_{\text{TAP}}(\boldsymbol{m}_+))$ are $O(1)$-Lipschitz (over $\|\boldsymbol{m}\|_N \le 1-\varepsilon$, for any $\varepsilon > 0$). Combined with (6.46), this implies

$$\mathsf{spec}(\nabla^2 \mathcal{F}_{\text{TAP}}(\boldsymbol{m})) \subseteq \left[-2C^{\mathsf{spec}}_{\max}, -\frac{1}{2}C^{\mathsf{spec}}_{\min}\right], \qquad \forall \|\boldsymbol{m} - \boldsymbol{m}^{\text{TAP}}\|_N \le \iota.$$

Thus $\nabla^2 \mathcal{F}_{\text{TAP}}$ is strongly concave and well-conditioned in the convex region $\|\boldsymbol{m} - \boldsymbol{m}^{\text{TAP}}\|_N \le \iota$. It is classical (see e.g. [Nes03]) that for suitable $\eta > 0$, gradient descent

$$\boldsymbol{u}^{k+1} = \boldsymbol{u}^k - \eta \nabla \mathcal{F}_{\text{TAP}}(\boldsymbol{u}^k)$$

initialized from $\boldsymbol{u}^0$ in this region satisfies

$$\|\boldsymbol{u}^k - \boldsymbol{m}^{\text{TAP}}\|_N \le (1-\varepsilon)^k \|\boldsymbol{u}^0 - \boldsymbol{m}^{\text{TAP}}\|_N \le \iota(1-\varepsilon)^k.$$

for some $\varepsilon > 0$. In particular $\boldsymbol{u}^0 = \boldsymbol{m}^{\text{AMP}}$ is in this region. Recalling $\boldsymbol{m}^{\mathsf{GD}} = \boldsymbol{u}^{K_{\mathsf{GD}}(N)}$ and $K_{\mathsf{GD}}(N) = \lfloor K^*_{\mathsf{GD}} \log N \rfloor$, we conclude

$$\|\boldsymbol{m}^{\mathsf{GD}} - \boldsymbol{m}^{\text{TAP}}\|_N \le \iota(1-\varepsilon)^{K_{\mathsf{GD}}(N)} \le N^{-10}$$

for suitably large K^*_{GD}. This implies part (b). □

We now turn to the proof of part (c). Recall from below (6.22) that $\boldsymbol{I}_{N-1}$ denotes the identity operator on $\mathsf{T}_{\boldsymbol{m}}$; we sometimes write this as $\boldsymbol{I}_{N-1}^{\boldsymbol{m}}$ to emphasize the dependence on $\boldsymbol{m}$.

Lemma 6.6.22. *Let $\gamma_* = (1-q_*)^{-1} + (1-q_*)\xi''(q_*)$. Let $\iota, \boldsymbol{m}^{\mathsf{TAP}} \in \mathcal{S}_\iota$ be as in Proposition 6.4.4. There exists $\iota' = o_\iota(1)$ such that with probability $1-e^{-cN}$, ($\boldsymbol{m}^{\mathsf{TAP}}$ is defined and)*

$$\mathsf{spec}\ \boldsymbol{A}^{(2)}(\boldsymbol{m}^{\mathsf{TAP}}) \subseteq [-(2+\iota')\sqrt{\xi''(q_*)}, (2+\iota')\sqrt{\xi''(q_*)}] \tag{6.86}$$

and

$$\left|\frac{1}{N}\mathsf{Tr}\left((\gamma_* \boldsymbol{I}_{N-1} - \boldsymbol{A}^{(2)}(\boldsymbol{m}^{\mathsf{TAP}}))^{-1}\right) - (1-q_*)\right| \le \iota'. \tag{6.87}$$

Proof. Let $\mathcal{E}_{\mathsf{spec}}$ be the event that (6.86), (6.87) both hold, and let $\mathcal{E}$ be as in Proposition 6.4.4. By Proposition 6.4.4(d) with $\delta = 1/2$,

$$\begin{aligned}
\mathbb{P}(\mathcal{E}_{\mathsf{spec}}^c) &\le \mathbb{P}(\mathcal{E}^c) + \mathbb{P}(\mathcal{E}_{\mathsf{spec}}^c \cap \mathcal{E}) \\
&\le e^{-cN} + C_{1/2} \sup_{\boldsymbol{m}\in\mathcal{S}_\iota} \mathbb{P}\left[\mathcal{E}_{\mathsf{spec}}^c \cap \mathcal{E} \middle| \nabla \mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m}) = \boldsymbol{0}\right]^{1/2}.
\end{aligned}$$

We will show that this probability is e^{-cN}, uniformly in $\boldsymbol{m} \in \mathcal{S}_\iota$. Note that on $\mathcal{E}$, we have deterministically $\boldsymbol{m}^{\mathsf{TAP}} = \boldsymbol{m}$.

Let $q_{\boldsymbol{m}} = \|\boldsymbol{m}\|_N^2 = \in [q_*-\iota, q_*+\iota]$. One checks analogously to Fact 6.6.5 that conditional on $\nabla\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m}) = \boldsymbol{0}$, we have $\boldsymbol{A}^{(2)}(\boldsymbol{m}) =_d \sqrt{\xi''(q_{\boldsymbol{m}})\frac{N-1}{N}}\boldsymbol{G}$, $\boldsymbol{G} \sim \mathsf{GOE}(N-1)$. It is classical that $\mathsf{spec}(\boldsymbol{G}) \subseteq [-2-\iota, 2+\iota]$ with probability $1-e^{-cN}$, so (6.86) holds with conditional probability $1-e^{-cN}$. Note that by (6.5),

$$\gamma_* - 2\sqrt{\xi''(q_*)} = \frac{1}{1-q_*}\left(1-(1-q_*)\xi''(q_*)^{1/2}\right)^2 > 0. \tag{6.88}$$

So, for small enough ι, when (6.86) holds the matrix $\gamma_*\boldsymbol{I}_{N-1} - \boldsymbol{A}^{(2)}(\boldsymbol{m})$ is positive semidefinite with smallest eigenvalue bounded away from 0. Recall the semicircle measure

$$\rho_{\mathsf{sc}}(\lambda) = \frac{1}{2\pi}\sqrt{4-\lambda^2}\ \mathsf{d}\lambda. \tag{6.89}$$

Applying [GZ00, Theorem 1.1(b)] as in the proof of Lemma 6.6.6 shows that with probability $1-e^{-cN}$,

$$\left|\frac{1}{N}\mathsf{Tr}\left((\gamma_*\boldsymbol{I}_{N-1} - \boldsymbol{A}^{(2)}(\boldsymbol{m}))^{-1}\right) - \int \frac{\rho_{\mathsf{sc}}(\mathsf{d}\lambda)}{\gamma_* - \sqrt{\xi''(q_{\boldsymbol{m}})}\lambda}\right| \le \iota.$$

This integral evaluates as

$$\int \frac{\rho_{\mathsf{sc}}(\mathsf{d}\lambda)}{\gamma_* - \sqrt{\xi''(q_*)}\lambda} + o_\iota(1) = 1 - q_* + o_\iota(1).$$

Thus, for suitable ι', (6.87) holds with conditional probability $1-e^{-cN}$, as desired. □

Lemma 6.6.23. *Suppose the event K_N in Proposition 6.3.6 holds. For any $\delta > 0$, there exists L such that for all $\delta \le \|\boldsymbol{m}_1\|_N, \|\boldsymbol{m}_2\|_N \le 1$, (treating $\boldsymbol{A}^{(2)}(\boldsymbol{m}_i)$ as a matrix in $\mathbb{R}^{N\times N}$, and $\boldsymbol{A}^{(3)}(\boldsymbol{m}_i)$ as a tensor in $(\mathbb{R}^N)^{\otimes 3}$)*

$$\begin{aligned}
\|\boldsymbol{A}^{(2)}(\boldsymbol{m}_1) - \boldsymbol{A}^{(2)}(\boldsymbol{m}_2)\|_{\mathsf{op},N} &\le L\|\boldsymbol{m}_1 - \boldsymbol{m}_2\|_N, \\
\|\boldsymbol{A}^{(3)}(\boldsymbol{m}_1) - \boldsymbol{A}^{(3)}(\boldsymbol{m}_2)\|_{\mathsf{op},N} &\le L\|\boldsymbol{m}_1 - \boldsymbol{m}_2\|_N.
\end{aligned}$$

Proof. Let $\mathsf{proj}_{\boldsymbol{m}}^\perp$ denote the projection operator to the orthogonal complement of $\boldsymbol{m}$. Then $\boldsymbol{A}^{(2)}(\boldsymbol{m}) = P_{\boldsymbol{m}}^\perp \nabla^2 H_N(\boldsymbol{m}) P_{\boldsymbol{m}}^\perp$. So,

$$\begin{aligned}
\|\boldsymbol{A}^{(2)}(\boldsymbol{m}_1) - \boldsymbol{A}^{(2)}(\boldsymbol{m}_2)\|_{\mathsf{op},N} &\le \|P_{\boldsymbol{m}_1}^\perp \nabla^2 H_N(\boldsymbol{m}_1) P_{\boldsymbol{m}_1}^\perp - P_{\boldsymbol{m}_1}^\perp \nabla^2 H_N(\boldsymbol{m}_1) P_{\boldsymbol{m}_2}^\perp\|_{\mathsf{op},N} \\
&\quad + \|P_{\boldsymbol{m}_1}^\perp \nabla^2 H_N(\boldsymbol{m}_1) P_{\boldsymbol{m}_2}^\perp - P_{\boldsymbol{m}_2}^\perp \nabla^2 H_N(\boldsymbol{m}_1) P_{\boldsymbol{m}_2}^\perp\|_{\mathsf{op},N} \\
&\quad + \|P_{\boldsymbol{m}_2}^\perp \nabla^2 H_N(\boldsymbol{m}_1) P_{\boldsymbol{m}_2}^\perp - P_{\boldsymbol{m}_2}^\perp \nabla^2 H_N(\boldsymbol{m}_2) P_{\boldsymbol{m}_2}^\perp\|_{\mathsf{op},N} \\
&\le 2\|P_{\boldsymbol{m}_1}^\perp - P_{\boldsymbol{m}_2}^\perp\|_{\mathsf{op},N} \max(\|\nabla^2 H_N(\boldsymbol{m}_1)\|_{\mathsf{op},N}, \|\nabla^2 H_N(\boldsymbol{m}_2)\|_{\mathsf{op},N}) \\
&\quad + \|\nabla^2 H_N(\boldsymbol{m}_1) - \nabla^2 H_N(\boldsymbol{m}_2)\|_{\mathsf{op},N}
\end{aligned}$$

On event K_N,

$$\|\nabla^2 H_N(\boldsymbol{m}_1)\|_{\mathsf{op},N}, \|\nabla^2 H_N(\boldsymbol{m}_2)\|_{\mathsf{op},N} \le C_2,$$
$$\|\nabla^2 H_N(\boldsymbol{m}_1) - \nabla^2 H_N(\boldsymbol{m}_2)\|_{\mathsf{op},N} \le C_3\|\boldsymbol{m}_1 - \boldsymbol{m}_2\|_N.$$

Finally, for a constant C_δ depending on δ,

$$\begin{aligned}\|P^\perp_{\boldsymbol{m}_1} - P^\perp_{\boldsymbol{m}_2}\|_{\mathsf{op},N} &= \left\|\frac{\boldsymbol{m}_1\boldsymbol{m}_1^\top}{\|\boldsymbol{m}_1\|^2} - \frac{\boldsymbol{m}_2\boldsymbol{m}_2^\top}{\|\boldsymbol{m}_2\|^2}\right\|_{\mathsf{op},N} \\ &\le \left\|\frac{\boldsymbol{m}_1\boldsymbol{m}_1^\top}{\|\boldsymbol{m}_1\|^2} - \frac{\boldsymbol{m}_1\boldsymbol{m}_1^\top}{\|\boldsymbol{m}_2\|^2}\right\|_{\mathsf{op},N} + \left\|\frac{\boldsymbol{m}_1\boldsymbol{m}_1^\top}{\|\boldsymbol{m}_2\|^2} - \frac{\boldsymbol{m}_2\boldsymbol{m}_2^\top}{\|\boldsymbol{m}_2\|^2}\right\|_{\mathsf{op},N} \le C_\delta\|\boldsymbol{m}_1 - \boldsymbol{m}_2\|_N.\end{aligned}$$

This proves the inequality for $\boldsymbol{A}^{(2)}$. The proof for $\boldsymbol{A}^{(3)}$ is analogous. □

Lemma 6.6.24. *There exists $L > 0$ such that with probability $1 - e^{-cN}$, for all $\boldsymbol{m}_1, \boldsymbol{m}_2 \in \mathsf{B}_N(\boldsymbol{m}^{\mathsf{TAP}}, \iota)$ (treating $\boldsymbol{Q}(\boldsymbol{m}_i)$ as a matrix in $\mathbb{R}^{N\times N}$)*

$$\|\boldsymbol{Q}(\boldsymbol{m}_1)\|_{\mathsf{op},N} \le L\,,$$
$$\|\boldsymbol{Q}(\boldsymbol{m}_1) - \boldsymbol{Q}(\boldsymbol{m}_2)\|_{\mathsf{op},N} \le L\|\boldsymbol{m}_1 - \boldsymbol{m}_2\|_N\,.$$

Proof. Suppose K_N holds and (6.86), (6.87) from Lemma 6.6.22 hold. Then, for some $\iota'' = o_\iota(1)$ and all $\boldsymbol{m} \in \mathsf{B}_N(\boldsymbol{m}^{\mathsf{TAP}}, \iota)$,

$$\mathsf{spec}\ \boldsymbol{A}^{(2)}(\boldsymbol{m}) \subseteq [-(2+\iota'')\sqrt{\xi''(q_*)}, (2+\iota'')\sqrt{\xi''(q_*)}] \tag{6.90}$$

and

$$\left|\frac{1}{N}\mathsf{Tr}\left((\gamma_*\boldsymbol{I}_{N-1} - \boldsymbol{A}^{(2)}(\boldsymbol{m}))^{-1}\right) - (1-q_*)\right| \le \iota''. \tag{6.91}$$

When (6.90) holds, the calculation (6.88) shows γ_* is bounded away from $\mathsf{spec}\ \boldsymbol{A}^{(2)}(\boldsymbol{m})$. Thus,

$$\gamma \mapsto \frac{1}{N}\mathsf{Tr}\left((\gamma\boldsymbol{I}_{N-1} - \boldsymbol{A}^{(2)}(\boldsymbol{m}))^{-1}\right)$$

has derivative $\Omega(1)$ in a neighborhood of γ_*. It follows from (6.91) that $\gamma_{*,N}(\boldsymbol{m}) = \gamma_* + o_\iota(1)$ uniformly for all $\boldsymbol{m} \in \mathsf{B}_N(\boldsymbol{m}^{\mathsf{TAP}}, \iota)$. This is also bounded away from $\mathsf{spec}\ \boldsymbol{A}^{(2)}(\boldsymbol{m})$, so

$$\|\boldsymbol{Q}(\boldsymbol{m})\|_{\mathsf{op},N} = \|(\gamma_{*,N}(\boldsymbol{m})\boldsymbol{I}_N - \boldsymbol{A}^{(2)}(\boldsymbol{m}))^{-1}\|_{\mathsf{op},N}$$

is bounded. Let $\boldsymbol{m}_1, \boldsymbol{m}_2 \in \mathsf{B}_N(\boldsymbol{m}^{\mathsf{TAP}}, \iota)$. There exists a rotation operator $\boldsymbol{R}$ from $\mathsf{T}_{\boldsymbol{m}_1}$ to $\mathsf{T}_{\boldsymbol{m}_2}$ such that $\left\|\boldsymbol{R} - \boldsymbol{I}^{\boldsymbol{m}_1}_{N-1}\right\|_{\mathsf{op},N} \le \|\boldsymbol{m}_1 - \boldsymbol{m}_2\|_N$. Recall $q_{\boldsymbol{m}} = \|\boldsymbol{m}\|_N^2$. The definition of $\gamma_{*,N}(\boldsymbol{m})$ implies

$$\begin{aligned}q_{\boldsymbol{m}_2} - q_{\boldsymbol{m}_1} &= \frac{1}{N}\mathsf{Tr}\left((\gamma_{*,N}(\boldsymbol{m}_1)\boldsymbol{I}^{\boldsymbol{m}_1}_{N-1} - \boldsymbol{A}^{(2)}(\boldsymbol{m}_1))^{-1} - (\gamma_{*,N}(\boldsymbol{m}_2)\boldsymbol{I}^{\boldsymbol{m}_1}_{N-1} - \boldsymbol{A}^{(2)}(\boldsymbol{m}_2))^{-1}\right) \\ &= \frac{1}{N}\mathsf{Tr}\Bigg(\boldsymbol{Q}(\boldsymbol{m}_1)\left((\gamma_{*,N}(\boldsymbol{m}_2) - \gamma_{*,N}(\boldsymbol{m}_1))\boldsymbol{I}_N - (\boldsymbol{A}^{(2)}(\boldsymbol{m}_1) - \boldsymbol{R}^{-1}\boldsymbol{A}^{(2)}(\boldsymbol{m}_2)\boldsymbol{R})\right) \\ &\qquad \boldsymbol{R}^{-1}\boldsymbol{Q}(\boldsymbol{m}_2)\boldsymbol{R}\Bigg).\end{aligned}$$

Thus,

$$\begin{aligned}&(\gamma_{*,N}(\boldsymbol{m}_1) - \gamma_{*,N}(\boldsymbol{m}_2))\frac{1}{N}\mathsf{Tr}(\boldsymbol{Q}(\boldsymbol{m}_1)\boldsymbol{R}^{-1}\boldsymbol{Q}(\boldsymbol{m}_2)\boldsymbol{R}) \\ &= q_{\boldsymbol{m}_2} - q_{\boldsymbol{m}_1} + \frac{1}{N}\mathsf{Tr}\Bigg(\boldsymbol{Q}(\boldsymbol{m}_1)(\boldsymbol{A}^{(2)}(\boldsymbol{m}_1) - \boldsymbol{R}^{-1}\boldsymbol{A}^{(2)}(\boldsymbol{m}_2)\boldsymbol{R})\boldsymbol{R}^{-1}\boldsymbol{Q}(\boldsymbol{m}_2)\boldsymbol{R}\Bigg).\end{aligned} \tag{6.92}$$

Note that

$$\|\boldsymbol{A}^{(2)}(\boldsymbol{m}_1) - \boldsymbol{R}^{-1}\boldsymbol{A}^{(2)\boldsymbol{R}}(\boldsymbol{m}_2)\|_{\mathsf{op},N} \le \|\boldsymbol{A}^{(2)}(\boldsymbol{m}_1)\|_{\mathsf{op},N}\|\boldsymbol{I}_{N-1}^{\boldsymbol{m}_1} - \boldsymbol{R}\|_{\mathsf{op},N} + \|\boldsymbol{A}^{(2)}(\boldsymbol{m}_1) - \boldsymbol{A}^{(2)}(\boldsymbol{m}_2)\|_{\mathsf{op},N},$$

and thus the absolute value of the right-hand side of (6.92) is upper bounded by

$$|q_{\boldsymbol{m}_2} - q_{\boldsymbol{m}_1}| + \|\boldsymbol{Q}(\boldsymbol{m}_1)\|_{\mathsf{op},N}\|\boldsymbol{Q}(\boldsymbol{m}_2)\|_{\mathsf{op},N}\|\boldsymbol{A}^{(2)}(\boldsymbol{m}_1) - \boldsymbol{R}^{-1}\boldsymbol{A}^{(2)}\boldsymbol{R}(\boldsymbol{m}_2)\|_{\mathsf{op},N} \le L\|\boldsymbol{m}_1 - \boldsymbol{m}_2\|_N,$$

by Lemma 6.6.23. As discussed above, $\boldsymbol{Q}(\boldsymbol{m}_1), \boldsymbol{R}^{-1}\boldsymbol{Q}(\boldsymbol{m}_2)\boldsymbol{R} \succeq c\boldsymbol{I}_{N-1}^{\boldsymbol{m}_1}$ for some constant $c > 0$, so

$$\frac{1}{N}\mathsf{Tr}(\boldsymbol{Q}(\boldsymbol{m}_1)\boldsymbol{R}^{-1}\boldsymbol{Q}(\boldsymbol{m}_2)\boldsymbol{R}) \ge \frac{1}{N}\mathsf{Tr}((c\boldsymbol{I}_{N-1}^{\boldsymbol{m}_1})^2) \ge c^2/2$$

is bounded away from 0. It follows that, after adjusting L,

$$|\gamma_{*,N}(\boldsymbol{m}_1) - \gamma_{*,N}(\boldsymbol{m}_2)| \le L\|\boldsymbol{m}_1 - \boldsymbol{m}_2\|_N.$$

Finally, (adjusting L again)

$$\begin{aligned}
&\|\boldsymbol{Q}(\boldsymbol{m}_1) - \boldsymbol{Q}(\boldsymbol{m}_2)\|_{\mathsf{op},N} \\
&= \left\|\boldsymbol{Q}(\boldsymbol{m}_1)\left((\gamma_{*,N}(\boldsymbol{m}_1) - \gamma_{*,N}(\boldsymbol{m}_2))\boldsymbol{I}_N - (\boldsymbol{A}^{(2)}(\boldsymbol{m}_1) - \boldsymbol{A}^{(2)}(\boldsymbol{m}_2))\right)\boldsymbol{Q}(\boldsymbol{m}_2)\right\|_{\mathsf{op},N} \\
&\le \|\boldsymbol{Q}(\boldsymbol{m}_1)\|_{\mathsf{op},N}\|\boldsymbol{Q}(\boldsymbol{m}_2)\|_{\mathsf{op},N}\left(|\gamma_{*,N}(\boldsymbol{m}_1) - \gamma_{*,N}(\boldsymbol{m}_2)| + \|\boldsymbol{A}^{(2)}(\boldsymbol{m}_1) - \boldsymbol{A}^{(2)}(\boldsymbol{m}_2)\|_{\mathsf{op},N}\right) \\
&\le L\|\boldsymbol{m}_1 - \boldsymbol{m}_2\|_N.
\end{aligned}$$

□

Proof of Proposition 6.4.5(c). For any $\|\boldsymbol{v}\|_2 = 1$,

$$\begin{aligned}
2|\langle \boldsymbol{\Delta}(\boldsymbol{m}_1) - \boldsymbol{\Delta}(\boldsymbol{m}_2), \boldsymbol{v}\rangle| &= |\langle \boldsymbol{A}^{(3)}(\boldsymbol{m}_1) \otimes \boldsymbol{v}, \boldsymbol{Q}(\boldsymbol{m}_1) \otimes \boldsymbol{Q}(\boldsymbol{m}_1)\rangle - \langle \boldsymbol{A}^{(3)}(\boldsymbol{m}_2) \otimes \boldsymbol{v}, \boldsymbol{Q}(\boldsymbol{m}_2) \otimes \boldsymbol{Q}(\boldsymbol{m}_2)\rangle| \\
&\le |\langle (\boldsymbol{A}^{(3)}(\boldsymbol{m}_1) - \boldsymbol{A}^{(3)}(\boldsymbol{m}_2)) \otimes \boldsymbol{v}, \boldsymbol{Q}(\boldsymbol{m}_1) \otimes \boldsymbol{Q}(\boldsymbol{m}_1)\rangle| \\
&+ |\langle \boldsymbol{A}^{(3)}(\boldsymbol{m}_2) \otimes \boldsymbol{v}, (\boldsymbol{Q}(\boldsymbol{m}_1) - \boldsymbol{Q}(\boldsymbol{m}_2)) \otimes \boldsymbol{Q}(\boldsymbol{m}_1)\rangle| \\
&+ |\langle \boldsymbol{A}^{(3)}(\boldsymbol{m}_2) \otimes \boldsymbol{v}, \boldsymbol{Q}(\boldsymbol{m}_2) \otimes (\boldsymbol{Q}(\boldsymbol{m}_1) - \boldsymbol{Q}(\boldsymbol{m}_2))\rangle|.
\end{aligned}$$

By the previous two lemmas, this is bounded by $\frac{L}{\sqrt{N}}\|\boldsymbol{m}_1 - \boldsymbol{m}_2\|_N$, for some $L > 0$. Since this holds for all $\boldsymbol{v}$, we have $\|\boldsymbol{\Delta}(\boldsymbol{m}_1) - \boldsymbol{\Delta}(\boldsymbol{m}_2)\|_2 \le \frac{L}{2\sqrt{N}}$, and thus $\|\boldsymbol{\Delta}(\boldsymbol{m}_1) - \boldsymbol{\Delta}(\boldsymbol{m}_2)\|_N \le \frac{L}{2N}$. □

6.7 Local computation of magnetization: proof of Proposition 6.4.6

Recall that $H_{N,t}(\boldsymbol{\sigma}) = H_N(\boldsymbol{\sigma}) + \langle \boldsymbol{y}_t, \boldsymbol{\sigma}\rangle$ with $\boldsymbol{y}_t = t\boldsymbol{x} + \boldsymbol{B}_t$, and define

$$\boldsymbol{x}^\perp = \boldsymbol{x} - \frac{\langle \boldsymbol{x}, \boldsymbol{m}\rangle_N}{\|\boldsymbol{m}\|_N^2}\boldsymbol{m}\,.$$

as well as the bands (for $\|\boldsymbol{m}\|_N^2 = q$)

$$\mathsf{Band}_*(\iota) := \mathsf{Band}(\boldsymbol{m}, I(\iota)) \cap \mathsf{Band}(\boldsymbol{x}, I(\iota))\,, \quad I(\iota) := [q_* - \iota, q_* + \iota]\,, \tag{6.93}$$

$$D_N(a,b) = \Big\{\boldsymbol{\sigma} \in S_N : \langle \boldsymbol{\sigma}, \boldsymbol{m}\rangle_N = aq,\, \langle \boldsymbol{\sigma}, \boldsymbol{x}\rangle_N = b\Big\}. \tag{6.94}$$

We recall the definition of truncated magnetization from Proposition 6.4.6:

$$\widetilde{\boldsymbol{m}}_{2\iota}(\boldsymbol{m}) = \frac{\int_{\mathsf{Band}_*(2\iota)} \boldsymbol{\sigma} \exp(H_{N,t}(\boldsymbol{\sigma}))\ \mu_0(\mathsf{d}\boldsymbol{\sigma})}{\int_{\mathsf{Band}_*(2\iota)} \exp(H_{N,t}(\boldsymbol{\sigma}))\ \mu_0(\mathsf{d}\boldsymbol{\sigma})}. \tag{6.95}$$

In Sections 6.7.1 and 6.7.2 we will prove Proposition 6.4.6. For the readers' convenience, we reproduce the statement below.

Proposition 6.4.6. *Define $\boldsymbol{A}_2 := \boldsymbol{A}^{(2)}(\boldsymbol{m})$, $\boldsymbol{A}_3 := \boldsymbol{A}^{(3)}(\boldsymbol{m})$ as per Eq. (6.21) and $\gamma_* = \gamma_{*,N}$ as per Eq. (6.22). Further recall the definition of $\mathcal{S}_\iota$ on Eq. (6.45), namely $\mathcal{S}_\iota := \big\{\boldsymbol{m} \in \mathbb{R}^N : |\langle \boldsymbol{m}, \boldsymbol{x}\rangle_N - q_*|, |\langle \boldsymbol{m}, \boldsymbol{m}\rangle_N - q_*| \le \iota\big\}$. Then we have, for appropriate constant $\delta, \iota > 0$,*

$$\begin{aligned}
&\sup_{\boldsymbol{m}\in\mathcal{S}_\iota} \mathbb{E}\left[\|\boldsymbol{m} + \boldsymbol{\Delta}(\boldsymbol{m}) - \widetilde{\boldsymbol{m}}_{2\iota}(\boldsymbol{m})\|_N^{2+\delta} \middle| \nabla\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m}) = \boldsymbol{0}\right] \\
&= \sup_{\boldsymbol{m}\in\mathcal{S}_\iota} \mathbb{E}\left[N^{-1-\delta/2} \left(\sum_{i=1}^N \left(\big[\widetilde{\boldsymbol{m}}_{2\iota}(\boldsymbol{m}) - \boldsymbol{m}\big]_i - \left(\frac{1}{2}\langle \boldsymbol{A}_3, \boldsymbol{Q}_{i,\cdot} \otimes \boldsymbol{Q}\rangle\right)\right)^2\right)^{1+\delta/2} \middle| \nabla\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m}) = \boldsymbol{0}\right] \\
&\le N^{-1-\delta}\,, \qquad (6.96)
\end{aligned}$$

$$\boldsymbol{Q} := (\gamma_* \boldsymbol{I} - \boldsymbol{A}_2)^{-1}\,. \qquad (6.97)$$

Our approach to proving Proposition 6.4.6 is based on decomposing

$$\mathsf{Band}_*(2\iota) = \cup_{r,s\in I(2\iota)} \mathsf{Band}(\boldsymbol{m}, \{r\}) \cap \mathsf{Band}(\boldsymbol{x}, \{s\}) = \cup_{a,b\in L(2\iota)} D_N(a,b)\,, \qquad (6.98)$$

where, for $q = \|\boldsymbol{m}\|_N^2$, $c = \langle \boldsymbol{x}, \boldsymbol{m}\rangle_N$,

$$L(2\iota) = \{(a,b) : \, qa \in I(2\iota), b \in I(2\iota)\}\,.$$

Note that for $\boldsymbol{m} \in \mathcal{S}_\iota$, we have $q, c \in I(\iota)$, and thus $L(2\iota)$ is a neighborhood of $(0,0)$ of radius of order ι. For any $r, s \in I(2\iota)$, we will see that the Hamiltonian restricted to $\mathsf{Band}(\boldsymbol{m}, \{r\}) \cap \mathsf{Band}(\boldsymbol{x}, \{s\})$ (conditional on $\nabla\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m}) = \boldsymbol{0}$) is equivalent to that of a mixed p-spin model in its replica symmetric phase with a small magnetic field. We will therefore devote Section 6.7.1 to study this problem. In Section 6.7.2 we will use this result, and integrate it over a, b to prove Proposition 6.4.6.

6.7.1 Conditional magnetization per band

As anticipated, in this section we will compute a good approximation to the magnetization for general spherical models with small external field. While we will apply this result to the effective Hamiltonian in the band, hence in dimension $N-2$, throughout this section, we adopt general notations for such a model, cf. Eq. (6.1) and recast $N-2$ as N. We write

$$H(\boldsymbol{\sigma}) = \langle \boldsymbol{u}, \boldsymbol{\sigma}\rangle + H_{\ge 2}(\boldsymbol{\sigma}) = \langle \boldsymbol{u}, \boldsymbol{\sigma}\rangle + \sum_{p\ge 2} H_p(\boldsymbol{\sigma})\,, \qquad (6.99)$$

$$H_p(\boldsymbol{\sigma}) = \frac{1}{N^{(p-1)/2}} \sum_{i_1,\dots,i_p=1}^N g_{i_1,\dots,i_p} \sigma_{i_1} \dots \sigma_{i_p}\,, \qquad g_{i_1,\dots,i_p} \overset{i.i.d.}{\sim} \mathcal{N}(0, \beta_p^2)\,. \qquad (6.100)$$

We will write throughout $H_{\ge i}(\boldsymbol{\sigma}) = \sum_{p\ge i} H_p(\boldsymbol{\sigma})$. We recast the mixture of $H_{\ge 2}$ as $\xi(s) = \sum_{p\ge 2} \beta_p^2 s^p$.

The results of this subsection hold for all models satisfying the replica symmetry condition (6.31), which we reproduce for convenience:

$$\xi''(0) < 1, \qquad \xi(q) + q + \log(1-q) < 0 \quad \forall q \in (0,1). \qquad (6.101)$$

This holds under the main condition (6.5) by integrating twice and, as we will see in (6.203), will hold for the effective model on the band $\mathsf{Band}(\boldsymbol{m}, \{r\}) \cap \mathsf{Band}(\boldsymbol{x}, \{s\})$, for all $r, s \in I(2\iota)$. Note that the first inequality in (6.101) implies $\beta_2^2 < 1/2$.

We will always assume $\|\boldsymbol{u}\|_2 \le c_0\sqrt{N}$, with c_0 a small constant, and in some lemmas $\|\boldsymbol{u}\|_2 \le N^{c_0}$. We will denote by $\mu(\mathsf{d}\boldsymbol{\sigma}) \propto \exp(H(\boldsymbol{\sigma}))\,\mu_0(\mathsf{d}\boldsymbol{\sigma})$ the corresponding Gibbs measure.

Note that we can view H_2 as a quadratic form with $H_2(\boldsymbol{\sigma}) = \langle \boldsymbol{\sigma}, \boldsymbol{W}^{(2)}\boldsymbol{\sigma}\rangle$ (with entries $W_{ij}^{(2)} = (g_{ij} + g_{ji})/2\sqrt{N}$). Hence $\boldsymbol{W}^{(2)}$ is a GOE matrix scaled by $\beta_2/\sqrt{2}$. We will work in the orthonormal basis diagonalizing $\boldsymbol{W}^{(2)}$ and its the spectrum of be $\boldsymbol{\Lambda} = (\Lambda_i)_{i\le N}$, with $\Lambda_1 \ge \Lambda_2 \ge \dots \ge \Lambda_N$. We will occasionally identify $\boldsymbol{\Lambda}$ with the diagonal matrix with diagonal entries Λ_i. We also write $\boldsymbol{W}^{(3)}$ for the symmetric 3rd-order tensor such that $H_3(\boldsymbol{\sigma}) = \langle \boldsymbol{W}^{(3)}, \boldsymbol{\sigma}^{\otimes 3}\rangle$, written in the basis of eigenvectors of $\boldsymbol{W}^{(2)}$. That is, $\boldsymbol{W}^{(3)}$ is obtained by rotating $\widetilde{\boldsymbol{w}}^{(3)}$ with entries $\widetilde{W}_{ijk}^{(3)} = (g_{ijk} + \text{permutations})/6N$.

Given a symmetric matrix $\boldsymbol{A} \in \mathbb{R}^{N\times N}$, we define

$$G(\gamma) = G(\gamma; \boldsymbol{A}, \boldsymbol{u}) := \gamma - \frac{1}{2N} \log\det(\gamma \boldsymbol{I} - \boldsymbol{A}) + \frac{1}{4N}\langle \boldsymbol{u}, (\gamma\boldsymbol{I} - \boldsymbol{A})^{-1}\boldsymbol{u}\rangle, \tag{6.102}$$

$$\gamma_* = \gamma_*(\boldsymbol{A}, \boldsymbol{u}) = \underset{\gamma > \lambda_{\max}(\boldsymbol{A})}{\arg\min}\; G(\gamma; \boldsymbol{A}, \boldsymbol{u})\,. \tag{6.103}$$

Note that G is convex with $\lim_{\gamma\downarrow\lambda_{\max}(\boldsymbol{A})} G'(\gamma) = -\infty$, $\lim_{\gamma\uparrow+\infty} G'(\gamma) = +\infty$, so γ_* is also the unique solution to $G'(\gamma) = 0$. We will omit the argument $\boldsymbol{A}$ or $\boldsymbol{u}$ whenever clear from the context (in particular, we typically omit $\boldsymbol{u}$ and omit $\boldsymbol{A}$ when $\boldsymbol{A} = \boldsymbol{\Lambda}$ is the diagonal matrix containing the eigenvalues of $\boldsymbol{W}$).

Lemma 6.7.1. *There exits $c_0 > 0$ such that, for $\|\boldsymbol{u}\| \le N^{c_0}$, and under the additional assumptions above, the following holds. Let $\gamma_* = \gamma_*(\boldsymbol{\Lambda})$ and define, for $j \le N$*

$$\widehat{m}_j := \frac{u_j}{2(\gamma_* - \Lambda_j)} + \frac{1}{2(\gamma_* - \Lambda_j)} \sum_{i=1}^{N} \frac{W^{(3)}_{jii}}{\gamma_* - \Lambda_i}\,. \tag{6.104}$$

Then, for some $c > 0$, with probability $1 - N^{-c}$, the following holds for all $i \in [N]$:

$$\int \sigma_i\, \mu(\mathsf{d}\boldsymbol{\sigma}) = (1 + O(N^{-c}))\Big(\widehat{m}_i + O(N^{-1/2-c})\Big).$$

Together with further estimates, we will use Lemma 6.7.1 to prove the following lemma, which is the main result of the section.

Lemma 6.7.2. *Let $\alpha \ge 2$. There exists $c_0 > 0$ such that, for $\|\boldsymbol{u}\| \le N^{c_0}$, and under the additional assumptions above, we have for some $c > 0$ that*

$$\mathbb{E}\left[\left\|\int \boldsymbol{\sigma}\, \mu(\mathsf{d}\boldsymbol{\sigma}) - \widehat{\boldsymbol{m}}\right\|_N^{\alpha}\right] = O(N^{-\alpha/2-c}). \tag{6.105}$$

The rest of this section is devoted to the proof of Lemma 6.7.1 and Lemma 6.7.2.

Quadratic Hamiltonians

We begin by proving several supporting lemmas about quadratic models. The Laplace transform allows us to compute accurately various statistics of quadratic models. We note that the use of Laplace transforms in studying spherical quadratic models has been utilized before, for example for analyzing the fluctuation of the free energy in [BL16]. We will however need accurate control over a number of statistics beyond the free energy.

Lemma 6.7.3. *For $\boldsymbol{A} \in \mathbb{R}^{N\times N}$ a GOE matrix scaled by $\alpha < 1/2$, and $\boldsymbol{u} \in \mathbb{R}^N$ such that $\|\boldsymbol{u}\|^2 \le \varepsilon N$ for $\varepsilon > 0$ depending only on $1/2 - \alpha$, there exists a constant $c > 0$ such that, defining and $G(\gamma) = G(\gamma; \boldsymbol{A})$, $\gamma_* = \gamma_*(\boldsymbol{A})$, we have that the following claim holds with probability at least $1 - \exp(-cN)$:*

$$\int e^{\langle \boldsymbol{\sigma}, \boldsymbol{A}\boldsymbol{\sigma}\rangle + \langle \boldsymbol{u}, \boldsymbol{\sigma}\rangle} \mu_0(\mathsf{d}\boldsymbol{\sigma}) = (1 + O(N^{-c}))\sqrt{\frac{2}{G''(\gamma_*)}}\,(2e)^{-N/2}\, e^{NG(\gamma_*)}, \tag{6.106}$$

and, for $\boldsymbol{v}_k$ any eigenvector of $\boldsymbol{A}$ (uniformly over k)

$$\frac{\int \langle \boldsymbol{v}_k, \boldsymbol{\sigma}\rangle e^{\langle \boldsymbol{\sigma}, \boldsymbol{A}\boldsymbol{\sigma}\rangle + \langle \boldsymbol{u}, \boldsymbol{\sigma}\rangle}\, \mu_0(\mathsf{d}\boldsymbol{\sigma})}{\int e^{\langle \boldsymbol{\sigma}, \boldsymbol{A}\boldsymbol{\sigma}\rangle + \langle \boldsymbol{u}, \boldsymbol{\sigma}\rangle}\, \mu_0(\mathsf{d}\boldsymbol{\sigma})} = (1 + O(N^{-c}))\frac{\langle \boldsymbol{v}_k, \boldsymbol{u}\rangle}{2(\gamma_* - \lambda_k(\boldsymbol{A}))}\,. \tag{6.107}$$

Proof. By a change of basis, we can assume that $\boldsymbol{A} = \boldsymbol{\Lambda}$ is diagonal (and its entries ordered). Let

$$E(\ell) := \ell^{N/2-1} \int \exp\Big(\langle \boldsymbol{u}, \boldsymbol{\sigma}\rangle \sqrt{\ell} + \langle \boldsymbol{\Lambda}, \boldsymbol{\sigma}^{\otimes 2}\rangle \ell\Big)\, \mu_0(\mathsf{d}\boldsymbol{\sigma})\,. \tag{6.108}$$

Then the Laplace transform of E is given by

$$F(t) = \int_0^{\infty} e^{-t\ell}\, E(\ell)\, \mathsf{d}\ell,$$

and one has (for $\Re(\gamma) > \Lambda_1 = \max_{i\le N} \Lambda_i$)

$$E(\ell) = \frac{1}{2\pi i}\int_{N\gamma - i\infty}^{N\gamma + i\infty} e^{t\ell}\, F(t)\, \mathsf{d}t.$$

We evaluate, for $\Re(t) > \Lambda_1$,

$$\begin{aligned}
F(Nt) &= \int_0^{\infty}\int e^{-Nt\ell} \exp\Big(\langle \boldsymbol{u}, \boldsymbol{\sigma}\rangle\sqrt{\ell} + \langle \boldsymbol{\Lambda}, \boldsymbol{\sigma}^{\otimes 2}\rangle \ell\Big)\ell^{N/2-1}\mu_0(\mathsf{d}\boldsymbol{\sigma})\,\mathsf{d}\ell \\
&= \frac{\Gamma(N/2)}{(N\pi)^{N/2}}\int_{\mathbb{R}^N} \exp\Big(-t\|\boldsymbol{y}\|^2 + \langle \boldsymbol{u}, \boldsymbol{y}\rangle + \langle \boldsymbol{\Lambda}, \boldsymbol{y}^{\otimes 2}\rangle\Big)\,\mathsf{d}\boldsymbol{y} \\
&= \frac{\Gamma(N/2)}{(N\pi)^{N/2}}\int_{\mathbb{R}^N} \exp\Big\{-\sum_{i=1}^N (t-\Lambda_i)y_i^2 + \sum_{i=1}^N u_i y_i\Big\}\,\mathsf{d}\boldsymbol{y} \\
&= \frac{\Gamma(N/2)}{N^{N/2}}\exp\left\{-\frac{1}{2}\sum_{i=1}^N \log(t-\Lambda_i) + \sum_{i=1}^N \frac{u_i^2}{4(t-\Lambda_i)}\right\}.
\end{aligned}$$

Hence, by the inverse Laplace transform, for all $\gamma \in \mathbb{R}$, $\gamma > \max_{i\le N}\Lambda_i$,

$$\begin{aligned}
E(1) &= \frac{N}{2\pi i}\int_{\gamma - i\infty}^{\gamma + i\infty} e^{Nt}\, F(Nt)\,\mathsf{d}t \\
&= \frac{\Gamma(N/2)}{2\pi N^{N/2-1}}\int_{-\infty}^{\infty} \exp\left\{N(\gamma + iz) - \frac{1}{2}\sum_{i=1}^N \log(\gamma + iz - \Lambda_i) + \sum_{i=1}^N \frac{u_i^2}{4(\gamma + iz - \Lambda_i)}\right\}\mathsf{d}z \\
&= \frac{\Gamma(N/2)}{2\pi N^{N/2-1}}\int_{-\infty}^{\infty} \exp\big(NG(\gamma + iz)\big)\mathsf{d}z, \qquad (6.109)
\end{aligned}$$

where $G(x) = G(x; \boldsymbol{\Lambda})$ is defined as per Eq. (6.102).

Let γ_* be defined as per Eq. (6.103). Per the discussion below (6.103), γ_* is the unique solution to $G'(\gamma_*) = 0$. Explicitly,

$$N - \frac{1}{2}\sum_{i=1}^N \frac{1}{\gamma_* - \Lambda_i} - \sum_{i=1}^N \frac{u_i^2}{4(\gamma_* - \Lambda_i)^2} = 0.$$

Our assumption on α and $\|\boldsymbol{u}\|$ implies that $\gamma_* - \max_i \boldsymbol{\Lambda}_i > \delta$ for an appropriate δ depending only on $1/2 - \alpha$. We will set $\gamma = \gamma_*$ in Eq. (6.109). Note that

$$\Re(G(\gamma_* + iz) - G(\gamma_*)) = -\frac{1}{4N}\sum_{i=1}^N \log(1 + z^2/(\gamma_* - \Lambda_i)^2) - \frac{1}{4N}\sum_{i=1}^N \frac{u_i^2 z^2}{(\gamma_* - \Lambda_i)(z^2 + (\gamma_* - \Lambda_i)^2)}.$$

For $|z| \in ((\log N)/\sqrt{N}, 1)$, we have $\Re(G(\gamma_* + iz) - G(\gamma_*)) < -cz^2$, and for $|z| \ge 1$, we have $\Re(G(\gamma_* + iz) - G(\gamma_*)) < -c\log(1 + cz^2)$. This implies that

$$\int_{|z| > \frac{\log N}{\sqrt{N}}} \exp\Big(NG(\gamma_* + iz) - NG(\gamma_*)\Big)\,\mathsf{d}z < e^{-c(\log N)^2}.$$

On the other hand, for $|z| \le (\log N)/\sqrt{N}$ we use the Taylor expansion:

$$G(\gamma_* + iz) = G(\gamma_*) + \sum_{j=1}^k \frac{(iz)^j}{j!}G^{(j)}(\gamma_*) + \mathsf{Err}_{N,k+1},$$

$$\mathsf{Err}_{N,k+1} \le \frac{C}{k!}\left(\frac{\log N}{\sqrt{N}}\right)^k \sup_{|z| \le (\log N)/\sqrt{N}} \left|G^{(k)}(\gamma_* + iz)\right|.$$

We have that

$$G^{(1)}(z) = 1 - \frac{1}{2N}\sum_{i=1}^N \frac{1}{z-\Lambda_i} - \frac{1}{4N}\sum_{i=1}^N \frac{u_i^2}{(z-\Lambda_i)^2},$$

$$G^{(j)}(z) = \frac{(-1)^j (j-1)!}{2N}\sum_{i=1}^N \frac{1}{(z-\Lambda_i)^j} + \frac{(-1)^j j!}{4N}\sum_{i=1}^N \frac{u_i^2}{(z-\Lambda_i)^{j+1}}.$$

In particular, with probability $1-\exp(-cN)$ over $\boldsymbol{A}$, $\sup_{|z|\le(\log N)/\sqrt{N}} |G^{(j)}(\gamma_*+iz)| \le j!C^j$ for a finite constant $C>0$ as long as $\|\boldsymbol{u}\|^2 \le N$. Hence, we have (for $\ell_N := (\log N)/\sqrt{N}$ and $J_N := [-\ell_N, \ell_N]$)

$$\begin{aligned}\int_{J_N} \exp\Big(NG(\gamma_*+iz) - NG(\gamma_*)\Big)\,\mathsf{d}z &= \int_{J_N} e^{-NG^{(2)}(\gamma_*)z^2/2} \exp\big(O(N\ell_N^3)\big)\,\mathsf{d}z\\ &= \sqrt{\frac{2\pi}{NG^{(2)}(\gamma_*)}}\Big(1+O(N^{-1/2+\varepsilon})\Big)\,.\end{aligned}$$

Together with Eq. (6.109), we get

$$E(1) = \frac{\Gamma(N/2)}{2\pi N^{N/2-1}} e^{NG(\gamma_*)}\left\{\sqrt{\frac{2\pi}{NG^{(2)}(\gamma_*)}} + O(N^{-1+\varepsilon})\right\}\,,$$

which yields (6.106) by Stirling's formula.

By a similar argument, we obtain the integral of the spin $\sigma_k = \langle \boldsymbol{v}_k, \boldsymbol{\sigma}\rangle$ (recall we are working in the basis in which $\boldsymbol{A}$ is diagonal). Let

$$E_k(\ell) := \ell^{N/2-1}\int \sqrt{\ell}\sigma_k \exp\Big(\langle \boldsymbol{u},\boldsymbol{\sigma}\rangle\sqrt{\ell} + \langle \boldsymbol{\Lambda}, \boldsymbol{\sigma}^{\otimes 2}\rangle \ell\Big)\,\mu_0(\mathsf{d}\boldsymbol{\sigma}).$$

Then the Laplace transform can be evaluated as

$$\begin{aligned}F_k(Nt) &= \int_0^\infty\int e^{-Nt\ell}\sqrt{\ell}\sigma_k \exp\Big(\langle \boldsymbol{u},\boldsymbol{\sigma}\rangle\sqrt{\ell} + \langle \boldsymbol{\Lambda}, \boldsymbol{\sigma}^{\otimes 2}\rangle \ell\Big)\ell^{N/2-1}\mu_0(\mathsf{d}\boldsymbol{\sigma})\,\mathsf{d}\ell\\ &= \frac{\Gamma(N/2)}{(N\pi)^{N/2}}\int_{\mathbb{R}^N} y_k \exp\big(-t\|\boldsymbol{y}\|^2 + \langle \boldsymbol{u},\boldsymbol{y}\rangle + \langle \boldsymbol{\Lambda}, \boldsymbol{y}^{\otimes 2}\rangle\big)\,\mathsf{d}\boldsymbol{y}\\ &= \frac{\Gamma(N/2)}{N^{N/2}}\frac{u_k}{2(t-\Lambda_k)}\exp\left\{-\frac{1}{2}\sum_{i=1}^N \log(t-\Lambda_i) + \sum_{i=1}^N \frac{u_i^2}{4(t-\Lambda_i)}\right\}.\end{aligned}$$

Then we apply the same strategy as for computing $E(1)$. By inverse Laplace transform:

$$E_k(1) = \frac{\Gamma(N/2)}{2\pi N^{N/2-1}} u_k \int_{-\infty}^\infty \frac{1}{2(\gamma_*+iz-\Lambda_k)}\exp\big(NG(\gamma_*+iz)\big)\mathsf{d}z, \tag{6.110}$$

We make a negligible error in restricting to $J_N := [-\ell_N, \ell_N]$ (for $\ell_N := (\log N)/\sqrt{N}$)

$$\begin{aligned}E_k(1) &= \frac{\Gamma(N/2)}{2\pi N^{N/2-1}} u_k \left\{\int_{J_N} \frac{1}{2(\gamma_*+iz-\Lambda_k)}\exp\big(NG(\gamma_*+iz)\big)\mathsf{d}z + O(e^{-c(\log N)^2})\right\}\\ &= \frac{\Gamma(N/2)}{2\pi N^{N/2-1}} u_k \left\{\sqrt{\frac{2\pi}{NG^{(2)}(\gamma_*)}}\frac{1}{2(\gamma_*-\Lambda_k)} + O(N^{-1+\varepsilon})\right\}\,.\end{aligned}$$

Comparing with the above, we get

$$E_k(1) = (1+O(N^{-c}))\frac{u_k}{2(\gamma_*-\Lambda_k)}\int e^{\langle \boldsymbol{\sigma}, \boldsymbol{A}\boldsymbol{\sigma}\rangle + \langle \boldsymbol{u},\boldsymbol{\sigma}\rangle}\mu_0(\mathsf{d}\boldsymbol{\sigma}).$$

This gives (6.107). □

Lemma 6.7.4. *Let $\boldsymbol{A} \in \mathbb{R}^{N\times N}$ be a GOE matrix scaled by $\alpha < 1/2$. Assume that $\boldsymbol{u} \in \mathbb{R}^N$ is such that $\|\boldsymbol{u}\| \le N^{c_0}$. For $\ell \in [L]$, consider a collection of pairs of indices (i_ℓ, j_ℓ) with $i_\ell \neq j_\ell \in [2k]$ and integers $r_\ell \ge 3$. Let $R = \sum_{\ell=1}^L r_\ell$. We have that the following claim holds with probability at least $1 - \exp(-cN)$:*

Uniformly in $h \in [N]$,

$$\frac{\int \prod_{i=1}^{2k} \boldsymbol{\sigma}_h^i \prod_{\ell=1}^L \langle \boldsymbol{\sigma}^{i_\ell}, \boldsymbol{\sigma}^{j_\ell}\rangle^{r_\ell} \exp\left(\sum_{i=1}^{2k}\langle \boldsymbol{u}, \boldsymbol{\sigma}^i\rangle + \langle \boldsymbol{\sigma}^i, \boldsymbol{\Lambda}\boldsymbol{\sigma}^i\rangle\right)\mu_0^{\otimes 2k}(\mathsf{d}\boldsymbol{\sigma})}{\int \exp\left(\sum_{i=1}^{2k}\langle \boldsymbol{u}, \boldsymbol{\sigma}^i\rangle + \langle \boldsymbol{\sigma}^i, \boldsymbol{\Lambda}\boldsymbol{\sigma}^i\rangle\right)\mu_0^{\otimes 2k}(\mathsf{d}\boldsymbol{\sigma})}$$
$$= O_{k,L}\left(|u_h|^{2k-2\min(k,L)} N^{(R-\min(k,L))/2}(1+\|\boldsymbol{u}\|)^{2R}\right).$$

Proof. As before, we perform a change of basis and assume that $\boldsymbol{A} = \boldsymbol{\Lambda}$ is diagonal. Consider

$$\begin{aligned}&E(z_1,\dots,z_{2k})\\ &:= (\prod_{i=1}^{2k} z_i)^{N/2-1}\int \prod_{i=1}^{2k}(\boldsymbol{\sigma}_h^i\sqrt{z_i})\prod_{\ell=1}^L\langle \boldsymbol{\sigma}^{i_\ell}\sqrt{z_{i_\ell}}, \boldsymbol{\sigma}^{j_\ell}\sqrt{z_{j_\ell}}\rangle^{r_\ell}\exp\left(\sum_{i=1}^{2k}\langle \boldsymbol{u}, \boldsymbol{\sigma}^i\rangle\sqrt{z_i} + \langle \boldsymbol{\sigma}^i, \boldsymbol{\Lambda}\boldsymbol{\sigma}^i\rangle z_i\right)\mu_0^{\otimes 2k}(\mathsf{d}\boldsymbol{\sigma}).\end{aligned}$$

Then the multivariate Laplace transform of E is

$$\begin{aligned}F(N(t_1,\dots,t_{2k})) &= \frac{\Gamma(N/2)^{2k}}{(N\pi)^{kN}}\int_{\mathbb{R}^N}\prod_{i=1}^{2k} y_h^i \prod_{\ell=1}^L \langle \boldsymbol{y}^{i_\ell}, \boldsymbol{y}^{j_\ell}\rangle^{r_\ell}\exp\left(-\sum_{i=1}^{2k}\left(t_i\|\boldsymbol{y}^i\|^2 - \langle \boldsymbol{u}, \boldsymbol{y}^i\rangle - \langle \boldsymbol{\Lambda}, (\boldsymbol{y}^i)^{\otimes 2}\rangle\right)\right)\mathsf{d}\boldsymbol{y}\\ &= \frac{\Gamma(N/2)^{2k}}{N^{kN}}\exp\left\{-\sum_{i=1}^{2k}\left(\frac{1}{2}\sum_{h'=1}^N \log(t_i - \Lambda_{h'}) + \sum_{h'=1}^N \frac{u_{h'}^2}{4(t_i-\Lambda_{h'})}\right)\right\}\cdot \mathfrak{G},\end{aligned}$$

where, for $\boldsymbol{y}^i = \frac{1}{2}(t_i - \boldsymbol{\Lambda})^{-1}\boldsymbol{u} + \boldsymbol{w}^i$ and $\boldsymbol{w}^i$ independently distributed according $\mathcal{N}\left(0, \frac{1}{2}(t_i - \boldsymbol{\Lambda})^{-1}\right)$,

$$\begin{aligned}\mathfrak{G} &:= \mathbb{E}\left[\prod_{i=1}^{2k} y_h^i \prod_{\ell=1}^L \langle \boldsymbol{y}^{i_\ell}, \boldsymbol{y}^{j_\ell}\rangle^{r_\ell}\right]\\ &= \mathbb{E}\Bigg[\prod_{i=1}^{2k}\left(\frac{u_h}{2(t_i - \Lambda_h)} + w_h^i\right)\\ &\prod_{\ell=1}^L\left(\langle \boldsymbol{w}^{i_\ell}, \boldsymbol{w}^{j_\ell}\rangle + \frac{1}{2}\langle \boldsymbol{w}^{i_\ell}, (t_{j_\ell} - \boldsymbol{\Lambda})^{-1}\boldsymbol{u}\rangle + \frac{1}{2}\langle \boldsymbol{w}^{j_\ell}, (t_{i_\ell} - \boldsymbol{\Lambda})^{-1}\boldsymbol{u}\rangle + \frac{1}{4}\langle (t_{j_\ell} - \boldsymbol{\Lambda})^{-1}\boldsymbol{u}, (t_{i_\ell} - \boldsymbol{\Lambda})^{-1}\boldsymbol{u}\rangle\right)^{r_\ell}\Bigg].\end{aligned} \tag{6.111}$$

Assume that $\min_{i\in[2k]}\Re(t_i - \max_{h'}\Lambda_{h'}) > c > 0$, and recall that $\|\boldsymbol{u}\| \le N^{c_0}$. Define $\mathcal{R}$ a tuple of sets of length $2k+R$, where, for $1 \le a \le 2k$, $\mathcal{R}_a$ is a subset of $\{a\}$, and for $a > 2k$, $\mathcal{R}_a$ is a subset of the pair of indices $\{i_\ell, j_\ell\}$ in the corresponding term. As such, each tuple $\mathcal{R}$ represents a term in the expansion of (6.111). If there exists an index in $[2k]$ that appears an odd number of times among the sets in $\mathcal{R}$, then the contribution of the corresponding term to (6.111) is 0. Consider the tuples $\mathcal{R}$ where each index appears an even number of times. Let $B(\mathcal{R})$ be the collection of indices $a \le 2k$ where $\mathcal{R}_a = \{a\}$, and let $b(\mathcal{R}) = |B(\mathcal{R})|$. The indices $B(\mathcal{R})$ must appear an odd number of times among the remaining sets $(\mathcal{R}_j)_{j=2k+1}^{2k+R}$. In each possible way to pick out terms among $(\mathcal{R}_j)_{j=2k+1}^{2k+R}$ so that each index in $B(\mathcal{R})$ appears at least once, let $d(\mathcal{R})$ denote the number of sets $|\mathcal{R}_j| = 1$ among these terms. Among the remaining terms, each of the index in $B(\mathcal{R})$ not covered by the $d(\mathcal{R})$ sets can be matched to terms among $(\mathcal{R}_j)_{j=2k+1}^{2k+R}$. Consider an arbitrary way to pair up all remaining indices appearing in the terms into pairs; let $f(\mathcal{R}) \le R - d(\mathcal{R}) - (b(\mathcal{R}) - d(\mathcal{R}))/2$ be the number of such pairs. For each such term and fixed pairing, we can upper bound its contribution to (6.111) by

$$O\left(|u_h|^{2k-b(\mathcal{R})+d(\mathcal{R})}N^{f(\mathcal{R})/2}(1+\|\boldsymbol{u}\|^2)^R\right) \le O\left(\max_{a\le\min(k,L)}|u_h|^{2k-2a}N^{(R-a)/2}(1+\|\boldsymbol{u}\|)^{2R}\right),$$

noting the constraints $0 \le d(\mathcal{R}) \le b(\mathcal{R})$, $f(\mathcal{R}) \le R - d(\mathcal{R}) - (b(\mathcal{R}) - d(\mathcal{R}))/2$. Thus, we have

$$\mathfrak{G} = O_{k,L}\left(\max_{a \le \min(k,L)} |u_h|^{2k-2a} N^{(R-a)/2}(1+\|\boldsymbol{u}\|)^{2R}\right).$$

Hence, for $R = \sum_\ell r_\ell \ge 3L$,

$$\mathfrak{G} = O_{k,L}\left(|u_h|^{2k-2\min(k,L)} N^{(R-\min(k,L))/2}(1+\|\boldsymbol{u}\|)^{2R}\right).$$

Taking the inverse Laplace transform and integrating on $t_i = \gamma_* + ix_i$, for γ_* defined in Eq. (6.103), noting that, similar to Lemma 6.7.3, we can restrict the integration to the range $x_i \in [-\ell_N, \ell_N]$ for $\ell_N = (\log N)/\sqrt{N}$, we obtain that

$$\begin{aligned}
&\frac{\int \prod_{i=1}^{2k} \boldsymbol{\sigma}_h^i \prod_{\ell=1}^{L} \langle \boldsymbol{\sigma}^{i_\ell}, \boldsymbol{\sigma}^{j_\ell}\rangle^{r_\ell} \exp\left(\sum_{i=1}^{2k}\langle \boldsymbol{u}, \boldsymbol{\sigma}^i\rangle + \langle \boldsymbol{\sigma}^i, \boldsymbol{\Lambda}\boldsymbol{\sigma}^i\rangle\right)\mu_0^{\otimes 2k}(\mathsf{d}\boldsymbol{\sigma})}{\int \exp\left(\sum_{i=1}^{2k}\langle \boldsymbol{u}, \boldsymbol{\sigma}^i\rangle + \langle \boldsymbol{\sigma}^i, \boldsymbol{\Lambda}\boldsymbol{\sigma}^i\rangle\right)\mu_0^{\otimes 2k}(\mathsf{d}\boldsymbol{\sigma})} \\
&= O_{k,L}\left(|u_h|^{2k-2\min(k,L)} N^{(R-\min(k,L))/2}(1+\|\boldsymbol{u}\|)^{2R}\right).
\end{aligned}$$

□

The next lemma states that, under a purely quadratic Hamiltonian, and for small field, the overlap concentrates near zero.

Lemma 6.7.5 (Overlap concentration in quadratic models)**.** *Define*

$$A_2(t)^c := \{(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2) \in S_N \times S_N : |\langle \boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2\rangle_N| \ge t\}.$$

Assuming that $\|\boldsymbol{u}\|^2 \le \delta N$ *for* δ *sufficiently small, we have for some constant* $c > 0$ *that, with probability* $1 - e^{-cN}$,

$$\frac{\int_{A_2(t)^c} \exp(H_{\le 2}(\boldsymbol{\sigma}^1) + H_{\le 2}(\boldsymbol{\sigma}^2))\mu_0^{\otimes 2}(\mathsf{d}\boldsymbol{\sigma})}{\int \exp(H_{\le 2}(\boldsymbol{\sigma}^1) + H_{\le 2}(\boldsymbol{\sigma}^2))\mu_0^{\otimes 2}(\mathsf{d}\boldsymbol{\sigma})} \le \exp\Big\{-cN\big(t - \|\boldsymbol{u}\|_N\big)_+^2\Big\}. \tag{6.112}$$

Proof. Consider the Hamiltonian $H(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2) = H_{\le 2}(\boldsymbol{\sigma}^1) + H_{\le 2}(\boldsymbol{\sigma}^2) + 2\theta\langle \boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2\rangle$. Let Λ_i be the eigenvalues of the quadratic component $\boldsymbol{A}$ of H. Using the Laplace transform as in Lemma 6.7.3,

$$\begin{aligned}
&\int \exp\left(\langle \boldsymbol{u}, \boldsymbol{\sigma}^1 + \boldsymbol{\sigma}^2\rangle + \langle \boldsymbol{A}, (\boldsymbol{\sigma}^1)^{\otimes 2} + (\boldsymbol{\sigma}^2)^{\otimes 2}\rangle\right)\exp\left(2\theta\langle \boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2\rangle\right)\mu_0^{\otimes 2}(\mathsf{d}\boldsymbol{\sigma}) \\
&= \left(\frac{\Gamma(N/2)}{(2\pi)N^{N/2-1}}\right)^2 \int_{-\infty}^{\infty} \exp\left(2N\tilde{G}_\theta(\gamma_1 + iz_1, \gamma_2 + iz_2)\right)\mathsf{d}z_1\mathsf{d}z_2,
\end{aligned}$$

where

$$\tilde{G}_\theta(z_1, z_2) = \frac{z_1 + z_2}{2} - \frac{1}{4N}\sum_{i=1}^{N}\log((z_1 - \Lambda_i)(z_2 - \Lambda_i) - \theta^2) + \frac{1}{8N}\sum_{i=1}^{N}\frac{u_i^2(z_1 + z_2 - 2\Lambda_i + 2\theta)}{(z_1 - \Lambda_i)(z_2 - \Lambda_i) - \theta^2}.$$

We also denote

$$G_\theta(z) = G_\theta(z, z) = z - \frac{1}{4N}\sum_{i=1}^{N}\log((z - \Lambda_i)^2 - \theta^2) + \frac{1}{4N}\sum_{i=1}^{N}\frac{u_i^2(z - \Lambda_i + \theta)}{(z - \Lambda_i)^2 - \theta^2}.$$

Let $\gamma_*(\theta)$ be a stationary point of G_θ on $\mathbb{R}$ so that $\gamma_*(\theta) > \max\Lambda_i + \theta$ (there exists a unique such point), and $\gamma_* = \gamma_*(0)$.

As in Lemma 6.7.3, noting that

$$\begin{aligned}
&\exp\left(2\Re\left(\log((\gamma_*(\theta)-\Lambda_i+iz_1)(\gamma_*(\theta)-\Lambda_i+iz_2)-\theta^2)-\log((\gamma_*(\theta)-\Lambda_i)^2-\theta^2)\right)\right)\\
&=\left(\frac{(\gamma_*(\theta)-\Lambda_i)^2-\theta^2-z_1z_2}{(\gamma_*(\theta)-\Lambda_i)^2-\theta^2}\right)^2+\left(\frac{(\gamma_*(\theta)-\Lambda_i)(z_1+z_2)}{(\gamma_*(\theta)-\Lambda_i)^2-\theta^2}\right)^2\\
&=\frac{((\gamma_*(\theta)-\Lambda_i)^2-\theta^2)^2+(z_1z_2)^2+2\theta^2z_1z_2+(z_1^2+z_2^2)(\gamma_*(\theta)-\Lambda_i)^2}{((\gamma_*(\theta)-\Lambda_i)^2-\theta^2)^2}\\
&\geq 1+\frac{z_1^2+z_2^2}{(\gamma_*(\theta)-\Lambda_i)^2-\theta^2}.
\end{aligned} \tag{6.113}$$

Furthermore,

$$\begin{aligned}
&\Re\left(\frac{(\gamma_*(\theta)-\Lambda_i+\theta)+i(z_1+z_2)/2}{N((\gamma_*(\theta)-\Lambda_i)^2-\theta^2-z_1z_2+i(\gamma_*(\theta)-\Lambda_i)(z_1+z_2))}-\frac{1}{N(\gamma_*(\theta)-\Lambda_i-\theta)}\right)\\
&=\frac{-(z_1z_2)^2-\theta(\gamma_*(\theta)-\Lambda_i+\theta)(z_1+z_2)^2/2-((\gamma_*(\theta)-\Lambda_i)^2-\theta^2)(z_1^2+z_2^2)/2}{N(\gamma_*(\theta)-\Lambda_i-\theta)(((\gamma_*(\theta)-\Lambda_i)^2-\theta^2-z_1z_2)^2+((\gamma_*(\theta)-\Lambda_i)(z_1+z_2))^2)}\\
&\leq 0.
\end{aligned} \tag{6.114}$$

Given (6.113) and (6.114), we can proceed as in Lemma 6.7.3 to restrict the integral over z_1 and z_2 to the range $|z_1|,|z_2|<(\log N)/\sqrt{N}$, incurring an error $e^{-c(\log N)^2}$. Then by similarly expanding around $(\gamma_*(\theta),\gamma_*(\theta))$, we obtain

$$\begin{aligned}
&\int\exp\left(\langle\boldsymbol{u},\boldsymbol{\sigma}^1+\boldsymbol{\sigma}^2\rangle+\langle\boldsymbol{A},(\boldsymbol{\sigma}^1)^{\otimes 2}+(\boldsymbol{\sigma}^2)^{\otimes 2}\rangle\right)\exp\left(2\theta\langle\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2\rangle\right)\mu_0^{\otimes 2}(\mathsf{d}\boldsymbol{\sigma})\\
&=\left(\frac{\Gamma(N/2)}{(2\pi)N^{N/2-1}}\right)^2e^{2NG_\theta(\gamma_*(\theta))}\left\{\frac{2\pi}{N\det(\nabla^2\tilde{G}_\theta(\gamma_*(\theta),\gamma_*(\theta)))^{1/2}}+O(N^{-3/2+\epsilon})\right\}
\end{aligned}$$

When $\|\boldsymbol{u}\|^2\leq\delta N$, we have for G as in (6.102) that

$$G_\theta(\gamma_*(\theta))-G(\gamma_*)=O(\theta^2)+O(\theta\|\boldsymbol{u}\|^2/N).$$

On the other hand, by Lemma 6.7.3,

$$\begin{aligned}
&\int\exp(\langle\boldsymbol{u},\boldsymbol{\sigma}^1+\boldsymbol{\sigma}^2\rangle+\langle\boldsymbol{A},(\boldsymbol{\sigma}^1)^{\otimes 2}+(\boldsymbol{\sigma}^2)^{\otimes 2}\rangle)\mu_0^{\otimes 2}(\mathsf{d}\boldsymbol{\sigma})\\
&=(1+O(N^{-c}))\left(\frac{\Gamma(N/2)}{(2\pi)N^{N/2-1}}\right)^2\left(\sqrt{\frac{2\pi}{NG''(\gamma_*)}}+O(N^{-3/2+\varepsilon})\right)^2.
\end{aligned}$$

In particular,

$$\begin{aligned}
&\frac{\int_{t\leq|\langle\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2\rangle_N|}\exp(H_{\leq 2}(\boldsymbol{\sigma}^1)+H_{\leq 2}(\boldsymbol{\sigma}^2))\mu_0^{\otimes 2}(\mathsf{d}\boldsymbol{\sigma})}{\int_{\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2}\exp(H_{\leq 2}(\boldsymbol{\sigma}^1)+H_{\leq 2}(\boldsymbol{\sigma}^2))\mu_0^{\otimes 2}(\mathsf{d}\boldsymbol{\sigma})}\\
&\leq\exp(-2N\theta t)\frac{\int_{\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2}\exp(\langle\boldsymbol{u},\boldsymbol{\sigma}^1+\boldsymbol{\sigma}^2\rangle+N\langle\boldsymbol{A},(\boldsymbol{\sigma}^1)^{\otimes 2}+(\boldsymbol{\sigma}^2)^{\otimes 2}\rangle)\exp\left(2\theta\langle\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2\rangle\right)\mu_0^{\otimes 2}(\mathsf{d}\boldsymbol{\sigma})}{\int_{\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2}\exp(\langle\boldsymbol{u},\boldsymbol{\sigma}^1+\boldsymbol{\sigma}^2\rangle+N\langle\boldsymbol{A},(\boldsymbol{\sigma}^1)^{\otimes 2}+(\boldsymbol{\sigma}^2)^{\otimes 2}\rangle)\mu_0^{\otimes 2}(\mathsf{d}\boldsymbol{\sigma})}\\
&\leq\exp\Big(-2N\theta t+O(N\theta^2)+O(\theta\|\boldsymbol{u}\|^2)\Big).
\end{aligned}$$

Optimizing over θ, we obtain Eq. (6.112). □

We will also need the following lemma, giving an accurate expansion of moments of overlaps in perturbations of quadratic Hamiltonians.

Lemma 6.7.6. *Let $\boldsymbol{A} \in \mathbb{R}^{N\times N}$ be a GOE matrix scaled by $\alpha < 1/2$ with eigenvalues given by $\boldsymbol{\Lambda}$. Let $\boldsymbol{\Delta} \in \mathbb{R}^{N\times N}$ be an independent GOE matrix scaled by $\beta > 0$ and $|\zeta_1|, |\zeta_2| \le C(\log N)/\sqrt{N}$. For $i = 1, 2$, let $\tilde{\boldsymbol{\Lambda}}_i = \boldsymbol{\Lambda} + \zeta_i \boldsymbol{\Delta}$. Assume that $\boldsymbol{u} \in \mathbb{R}^N$ is such that $\|\boldsymbol{u}\| \le N^{c_0}$. Let $r \ge 0$ and $L > 0$. We have that the following claim holds with probability at least $1 - \exp(-cN)$: There exist $C_{i,j} = O_{r,L}(\|\boldsymbol{u}\|^{2r} + N^{\lfloor r/2 \rfloor})$ for $i, j \le L$ such that*

$$\frac{\int \langle \boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2\rangle^r \exp\left(\sum_{i=1}^2 \langle \boldsymbol{u}, \boldsymbol{\sigma}^i\rangle + \langle \boldsymbol{\sigma}^i, \tilde{\boldsymbol{\Lambda}}_i \boldsymbol{\sigma}^i\rangle\right) \mu_0^{\otimes 2}(\mathsf{d}\boldsymbol{\sigma})}{\int \exp\left(\sum_{i=1}^2 \langle \boldsymbol{u}, \boldsymbol{\sigma}^i\rangle + \langle \boldsymbol{\sigma}^i, \boldsymbol{\Lambda} \boldsymbol{\sigma}^i\rangle\right) \mu_0^{\otimes 2}(\mathsf{d}\boldsymbol{\sigma})}$$
$$= C_{0,0} + \sum_{i,j=0,(i,j)\neq(0,0)}^{L} C_{i,j}\zeta_1^i \zeta_2^j + O_L(N^{-L/2} + e^{-N^c}).$$

Proof. Consider

$$E(z_1, z_2) := (\prod_{i=1}^2 z_i)^{N/2-1} \int \langle \boldsymbol{\sigma}^1 \sqrt{z_1}, \boldsymbol{\sigma}^2 \sqrt{z_2}\rangle^r \exp\left(\sum_{i=1}^2 \langle \boldsymbol{u}, \boldsymbol{\sigma}^i\rangle \sqrt{z_i} + \langle \boldsymbol{\sigma}^i, \tilde{\boldsymbol{\Lambda}}_i \boldsymbol{\sigma}^i\rangle z_i\right) \mu_0^{\otimes 2}(\mathsf{d}\boldsymbol{\sigma}).$$

Then the multivariate Laplace transform of E is

$$F(N(t_1, t_2)) = \frac{\Gamma(N/2)^2}{(N\pi)^N} \int_{\mathbb{R}^N} \langle \boldsymbol{y}^1, \boldsymbol{y}^2\rangle^r \exp\left(-\sum_{i=1}^2 \left(t_i \|\boldsymbol{y}^i\|^2 - \langle \boldsymbol{u}, \boldsymbol{y}^i\rangle - \langle \tilde{\boldsymbol{\Lambda}}_i, (\boldsymbol{y}^i)^{\otimes 2}\rangle\right)\right) \mathsf{d}\boldsymbol{y}$$
$$= \frac{\Gamma(N/2)^2}{N^N} \exp\left\{-\sum_{i=1}^2 \left(\frac{1}{2} \log\det(t_i \boldsymbol{I}_N - \tilde{\boldsymbol{\Lambda}}_i) + \sum_{h'=1}^N \frac{1}{4N} \langle (t_i \boldsymbol{I}_N - \tilde{\boldsymbol{\Lambda}}_i)^{-1}, \boldsymbol{u}\boldsymbol{u}^T\rangle\right)\right\} \cdot \mathfrak{G}(\sigma),$$

where, for $\boldsymbol{Z}^i = (t_i \boldsymbol{I}_N - \tilde{\boldsymbol{\Lambda}}_i)^{-1}$, $\boldsymbol{y}^i = \frac{1}{2}\boldsymbol{Z}^i \boldsymbol{u} + \boldsymbol{w}^i$ and $\boldsymbol{w}^i$ independently distributed according $\mathcal{N}\left(0, \frac{1}{2}\boldsymbol{Z}^i\right)$,

$$\mathfrak{G}(t_1, t_2; \sigma_1, \sigma_2) := \mathbb{E}\left[\langle \boldsymbol{y}^1, \boldsymbol{y}^2\rangle^r\right]$$
$$= \mathbb{E}\left[\left(\langle \boldsymbol{w}^1, \boldsymbol{w}^2\rangle + \frac{1}{2}\langle \boldsymbol{w}^1, \boldsymbol{Z}^2 \boldsymbol{u}\rangle + \frac{1}{2}\langle \boldsymbol{w}^2, \boldsymbol{Z}^1 \boldsymbol{u}\rangle + \frac{1}{4}\langle \boldsymbol{Z}^1 \boldsymbol{u}, \boldsymbol{Z}^2 \boldsymbol{u}\rangle\right)^r\right].$$

Let $\mathfrak{G}_0(t_1, t_2) = \mathfrak{G}(t_1, t_2; 0, 0)$. Note that $\mathfrak{G}$ is a rational function of t_i, and hence extends to complex values of t_i. We next consider the Taylor expansion in ζ_1, ζ_2 of $\mathfrak{G}$. Write $\boldsymbol{w}^i = (\frac{1}{2}\boldsymbol{Z}^i)^{1/2} \widetilde{\boldsymbol{W}}^i$ for $\widetilde{\boldsymbol{W}}^i \sim \mathcal{N}(0, \boldsymbol{I}_N)$. Note that

$$\|\partial_{\zeta_1}^i \boldsymbol{Z}^1\|_{\mathsf{op}} = O_i(\beta).$$

We can thus bound the derivatives of $\mathfrak{G}(t_1, t_2; \zeta_1, \zeta_2)$ for $|t_1 - \gamma_*|, |t_2 - \gamma_*| \le C(\log N)/\sqrt{N}$ as

$$|\partial_{\zeta_1}^i \partial_{\zeta_2}^j \mathfrak{G}(t_1, t_2; \zeta_1, \zeta_2)| \le O_{r,i+j}\left(\|\boldsymbol{u}\|^{2r} + N^{r/2}\right) \tag{6.115}$$

for r even, and

$$|\partial_{\zeta_1}^i \partial_{\zeta_2}^j \mathfrak{G}(t_1, t_2; \zeta_1, \zeta_2)| \le O_{r,i+j}\left(\|\boldsymbol{u}\|^{2r} + \|\boldsymbol{u}\|^2 N^{(r-1)/2}\right) \tag{6.116}$$

for r odd. We can thus write

$$\mathfrak{G}(t_1, t_2; \zeta_1, \zeta_2) = \mathfrak{G}_0(t_1, t_2) + \sum_{i,j\le L,(i,j)\neq(0,0)} C_{i,j}\zeta_1^i \zeta_2^j + O(\max(\zeta_1, \zeta_2)^{L+1}),$$

where $|C_{i,j}| = O_{r,i+j}\left(\|\boldsymbol{u}\|^{2r} + N^{\lfloor r/2 \rfloor}\right)$.

Let

$$F(\zeta_1, \zeta_2) := \frac{\int \langle \boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2\rangle^r \exp\left(\sum_{i=1}^2 \langle \boldsymbol{u}, \boldsymbol{\sigma}^i\rangle + \langle \boldsymbol{\sigma}^i, \tilde{\boldsymbol{\Lambda}}_i \boldsymbol{\sigma}^i\rangle\right) \mu_0^{\otimes 2}(\mathsf{d}\boldsymbol{\sigma})}{\int \exp\left(\sum_{i=1}^2 \langle \boldsymbol{u}, \boldsymbol{\sigma}^i\rangle + \langle \boldsymbol{\sigma}^i, \boldsymbol{\Lambda} \boldsymbol{\sigma}^i\rangle\right) \mu_0^{\otimes 2}(\mathsf{d}\boldsymbol{\sigma})}. \tag{6.117}$$

Next, we take the inverse Laplace transform and integrate on $t_i = \gamma_* + ix_i$, for γ_* defined in Eq. (6.103). We note that, for $G(\gamma) = G(\gamma; \boldsymbol{\Lambda}, \boldsymbol{u})$ and $\tilde{G}_i(\gamma) = G(\gamma; \tilde{\boldsymbol{\Lambda}}_i, \boldsymbol{u})$,

$$\begin{aligned}\tilde{G}_i'(z) = G'(z) &+ \frac{1}{2N}\mathsf{Tr}((z-\boldsymbol{\Lambda})^{-1}(\boldsymbol{I} - (z-\boldsymbol{\Lambda})(z-\tilde{\boldsymbol{\Lambda}}_i)^{-1})) \\ &+ \frac{1}{4N}\langle \boldsymbol{u}, (z-\boldsymbol{\Lambda})^{-1}(\boldsymbol{I} - (z-\boldsymbol{\Lambda})(z-\tilde{\boldsymbol{\Lambda}}_i)^{-2}(z-\boldsymbol{\Lambda}))(z-\boldsymbol{\Lambda})^{-1}\boldsymbol{u}\rangle.\end{aligned}$$

Moreover, $(z-\boldsymbol{\Lambda})(z-\tilde{\boldsymbol{\Lambda}})^{-1} = (\boldsymbol{I} - \zeta_i\boldsymbol{\Delta}(z-\boldsymbol{\Lambda})^{-1})^{-1}$, and $(z-\boldsymbol{\Lambda})(z-\tilde{\boldsymbol{\Lambda}}_i)^{-2}(z-\boldsymbol{\Lambda}) = (\boldsymbol{I} - \zeta_i\boldsymbol{\Delta}(z-\boldsymbol{\Lambda})^{-1})^{-1}(\boldsymbol{I} - (z-\boldsymbol{\Lambda})^{-1}\zeta_i\boldsymbol{\Delta})^{-1}$. Expanding in $\zeta_i\boldsymbol{\Delta}$, we can show that for $|\zeta_i| \le C(\log N)/\sqrt{N}$, $|\tilde{G}_i'(\gamma_*)| \le N^{-1+o(1)}$. Hence, by an argument similar to Lemma 6.7.3, we can restrict the integration on $t_i = \gamma_* + ix_i$ to the range $x_i \in [-\ell_N, \ell_N]$ for $\ell_N = (\log N)/\sqrt{N}$, and obtain that

$$F(\zeta_1, \zeta_2) = F(0,0) + \sum_{i,j\le L, (i,j)\ne(0,0)} C_{i,j}\zeta_1^i\zeta_2^j + O_L(N^{-L/2} + e^{-N^c}).$$

□

Estimates of restricted partition functions

In this section we estimate modified partition functions that are obtained by suitable restrictions of the integral over $\boldsymbol{\sigma}$, always under the assumption (6.101). Namely, for any Borel set $U \subseteq (S_N)^{\otimes m}$,

$$Z_m(U) := \int_U e^{\sum_{i=1}^m H(\boldsymbol{\sigma}^i)}\mu_0^{\otimes m}(\mathsf{d}\boldsymbol{\sigma}), \tag{6.118}$$

with subscript omitted if $m = 1$. If $U = S_N$, we write simply $Z = Z(S_N)$. We also denote by $Z_{\le 2,m}(U)$ the same integral whereby $H(\boldsymbol{\sigma})$ is replaced by $H_{\le 2}(\boldsymbol{\sigma})$:

$$Z_m(U) := \int_U e^{\sum_{i=1}^m H_{\le 2}(\boldsymbol{\sigma}^i)}\mu_0^{\otimes m}(\mathsf{d}\boldsymbol{\sigma}), \tag{6.119}$$

We will occasionally omit the subscript m when the dimension of U is clear from the context.

Throughout this section, we follow the notations $\langle \boldsymbol{x}, \boldsymbol{y}\rangle_N = \langle \boldsymbol{x}, \boldsymbol{y}\rangle/N$, so $\langle \boldsymbol{x}, \boldsymbol{x}\rangle_N = \|\boldsymbol{x}\|_N^2$.

As for the restrictions, an important role is played by the typical set:

$$T(\delta) = \left\{\boldsymbol{\sigma} \in S_N : \int_{\boldsymbol{\sigma}': |\langle \boldsymbol{\sigma}', \boldsymbol{\sigma}\rangle_N| > \delta} e^{H(\boldsymbol{\sigma}')}\mu_0(\mathsf{d}\boldsymbol{\sigma}') < e^{-c_1(\delta)N}\min\left(\int e^{H(\boldsymbol{\sigma})}\mu_0(\mathsf{d}\boldsymbol{\sigma}); e^{N\xi(1)/2}\right)\right\}. \tag{6.120}$$

We further define $A_m(\delta) \subseteq (S_N)^m$ to be the set of m-uples of vectors which are nearly orthogonal. Namely:

$$A_m(\delta) := \left\{(\boldsymbol{\sigma}^i)_{i\le m} : \boldsymbol{\sigma}^i \in S_N, |\langle \boldsymbol{\sigma}^i, \boldsymbol{\sigma}^j\rangle_N| \le \delta \ \forall i \ne j\right\}. \tag{6.121}$$

Finally, we consider the set of m-uples in $T = T(\delta)$ that are nearly orthogonal:

$$A_m(T, \delta) := \left\{(\boldsymbol{\sigma}^i)_{i\le m} : \boldsymbol{\sigma}^i \in T, |\langle \boldsymbol{\sigma}^i, \boldsymbol{\sigma}^j\rangle_N| \le \delta \ \forall i \ne j\right\}. \tag{6.122}$$

In particular $A_m(T, \delta) = T^m \cap A_m(\delta)$.

Our first lemma establishes that, under the Gibbs measure, non-typical points are exponentially rare.

Lemma 6.7.7 (Most points are typical)**.** *For any $\delta > 0$, there exists $u(\delta), c_1(\delta), c_2(\delta) > 0$ such that the following holds. Let $H(\boldsymbol{\sigma})$ be defined as per Eq. (6.99) and suppose $\|\boldsymbol{u}\| \le u(\delta)\sqrt{N}$. Let $T(\delta)$ be defined as per Eq. (6.120).*

Then, with probability at least $1 - \exp(-c_2(\delta)N)$,

$$Z(T(\delta)) \ge (1 - e^{-Nc_2(\delta)}) \cdot Z. \tag{6.123}$$

Furthermore, there is $c_3(\delta) > 0$ such that, with probability at least $1 - \exp(-c_3(\delta)N)$,

$$Z_{\le 2}(T(\delta)^c) \le e^{-c_3(\delta)N} Z_{\le 2}. \tag{6.124}$$

Finally

$$\mathbb{E}\left[\int_{T(\delta)^c} e^{H_{\ge 2}(\boldsymbol{\sigma})}\ \mu_0(\mathsf{d}\boldsymbol{\sigma})\right] \le e^{-c_1(\delta)N}\, \mathbb{E}\, Z_{\ge 2}. \tag{6.125}$$

Proof. The second inequality in (6.101) is termed "strictly RS" in [HS23b], see Eq. (2.7) therein. By Proposition 3.1 of that paper,

$$\mathbb{E}\int_{T(\delta)} e^{H_{\ge 2}(\boldsymbol{\sigma})}\ \mu_0(\mathsf{d}\boldsymbol{\sigma}) \ge (1 - e^{-c_1(\delta)N})e^{N\xi(1)/2}.$$

(While this proposition states a bound of $(1 - o(1))\exp(N\xi(1)/2)$, its proof shows the $1 - o(1)$ is in fact $1 - e^{-c_1(\delta)N}$.) As $\mathbb{E}\, Z_{\ge 2} = \exp(N\xi(1)/2)$, for $Z_{\ge 2} := \int_{S_N} \exp H_{\ge 2}(\boldsymbol{\sigma})\ \mu_0(\mathsf{d}\boldsymbol{\sigma})$, this implies Eq. (6.125).

By Markov's inequality, with probability $1 - e^{-c_1(\delta)N/5}$,

$$\int_{T(\delta)^c} e^{H_{\ge 2}(\boldsymbol{\sigma})}\ \mu_0(\mathsf{d}\boldsymbol{\sigma}) \le e^{-4c_1(\delta)N/5}\, \mathbb{E}\, Z_{\ge 2}.$$

By [Tal06a, Proposition 2.3], (6.101) implies that $\frac{1}{N}\log Z_{\ge 2} \to_p \xi(1)/2$. By standard concentration properties of $\frac{1}{N}\log Z_{\ge 2}$, with probability $1 - e^{-c_2(\delta)N}$,

$$Z_{\ge 2} \ge e^{-c_1(\delta)N/5}\, \mathbb{E}\, Z_{\ge 2}.$$

On the intersection of these events,

$$\int_{T(\delta)^c} e^{H_{\ge 2}(\boldsymbol{\sigma})}\ \mu_0(\mathsf{d}\boldsymbol{\sigma}) \le e^{-3c_1(\delta)N/5} Z_{\ge 2}.$$

Finally, set $u(\delta) = c_1(\delta)/5$, so that for all $\boldsymbol{\sigma} \in S_N$,

$$|H(\boldsymbol{\sigma}) - H_{\ge 2}(\boldsymbol{\sigma})| = |\langle \boldsymbol{u}, \boldsymbol{\sigma}\rangle| \le c_1(\delta)N/5.$$

Thus

$$\int_{T(\delta)^c} e^{H(\boldsymbol{\sigma})}\ \mu_0(\mathsf{d}\boldsymbol{\sigma}) \le e^{-c_1(\delta)N/5}\int_{S_N} e^{H(\boldsymbol{\sigma})}\ \mu_0(\mathsf{d}\boldsymbol{\sigma}).$$

The conclusion (6.123) follows with $c(\delta) = \min(c_1(\delta)/6, c_2(\delta)/2)$.

Finally, from Markov's inequality, we have with probability $1 - e^{-c_3(\delta)N}$ that

$$Z_{\le 2}(T(\delta)^c) \le e^{-c_3(\delta)N} e^{N\xi_{\le 2}(1)/2}.$$

Then (6.124) follows from standard concentration properties. $\square$

The next lemma states that we can anneal over terms of degree higher than 2 in the Hamiltonian. This will be the most important technical result of the section.

Lemma 6.7.8. *Let $H(\boldsymbol{\sigma})$ be defined as per Eq. (6.99) and define $T = T(\delta)$ as in Eq. (6.120). Assume that $\|\boldsymbol{u}\| \le N^{c_0}$ for c_0 sufficiently small given ξ. Under assumption (6.101), for all $L, k > 0$ and $\varepsilon > 0$, there exist $C = C(L,k) > 0$ such that the following holds with probability at least $1 - \exp(-N/C)$*

$$\mathbb{E}_{\ge 3}\Big\{\big(Z(T) - \mathbb{E}_{\ge 3}Z(T)\big)^{2k}\Big\} \le C\, N^{-L/2}\left(\mathbb{E}_{\ge 3}Z(T)\right)^{2k}, \tag{6.126}$$

and further

$$\mathbb{P}\left\{\big|Z - \mathbb{E}_{\ge 3}Z(T)\big| > \varepsilon\, \mathbb{E}_{\ge 3}Z(T)\right\} \le C\varepsilon^{-2L}N^{-L/2} + e^{-N/C}. \tag{6.127}$$

We also have, with probability at least $1-\exp(-N/C)$

$$\mathbb{E}_{\geq 3} Z(T) = (1+O(e^{-N/C}))\mathbb{E}_{\geq 3} Z. \tag{6.128}$$

Further, letting $(\boldsymbol{v}_k)_{k\leq N}$ *be the basis of eigenvectors of* $\boldsymbol{W}_2$*, for each* $i \in [N]$*,*

$$\mathbb{P}\left(\int_T \langle \boldsymbol{v}_i, \boldsymbol{\sigma}\rangle \, e^{H(\boldsymbol{\sigma})}\mu_0(\mathsf{d}\boldsymbol{\sigma}) \geq N^{\varepsilon}\|\boldsymbol{u}\|^{Ck}(|\langle \boldsymbol{v}_i, \boldsymbol{u}\rangle| + C\, N^{-1/2})\mathbb{E}_{\geq 3}\int e^{H(\boldsymbol{\sigma})}\mu_0(\mathsf{d}\boldsymbol{\sigma})\right) \leq C\left(N^{-2\varepsilon k} + e^{-N/C}\right). \tag{6.129}$$

Before proving this lemma, we state and prove a number of key estimates.

Our first lemma establishes that (in expectation) the partition function in $A_{2k}(\delta)$ is dominated by the subset $A_{2k}(\delta, T)$.

Lemma 6.7.9 (Orthogonal frames are mostly typical)**.** *Define* $T = T(\delta)$ *as in Eq. (6.120). We have for* $\delta > 0$ *sufficiently small and appropriate* $c, c' > 0$ *that, if* $\|\boldsymbol{u}\| \leq c'\sqrt{N}$*,*

$$\mathbb{E}\, Z_{2k}\big(\{(\boldsymbol{\sigma}^i)_{i\leq 2k} \in A_{2k}(\delta) : \boldsymbol{\sigma}^1 \in T^c\}\big) \leq e^{-cN}\mathbb{E}\, Z_{2k}\big(A_{2k}(\delta)\big). \tag{6.130}$$

As a consequence,

$$\mathbb{E}\, Z_{2k}\big(A_{2k}(\delta, T)\big) \geq (1-e^{-cN})\mathbb{E}\, Z_{2k}\big(A_{2k}(\delta)\big).$$

Proof. We have

$$\begin{aligned}\mathbb{E}\Big\{H_{\geq 2}(\boldsymbol{\rho})\Big|\sum_{i=1}^{2k} H_{\geq 2}(\boldsymbol{\sigma}^i)\Big\} &= \frac{\mathbb{E} H_{\geq 2}(\boldsymbol{\rho})\sum_{i=1}^{2k} H_{\geq 2}(\boldsymbol{\sigma}^i)}{\mathbb{E}(\sum_{i=1}^{2k} H_{\geq 2}(\boldsymbol{\sigma}^i))^2}\sum_{i=1}^{2k} H_{\geq 2}(\boldsymbol{\sigma}^i)\\ &= \frac{\sum_{i=1}^{2k}\xi(\langle \boldsymbol{\rho}, \boldsymbol{\sigma}^i\rangle_N)}{\sum_{i,j\in[2k]}\xi(\langle \boldsymbol{\sigma}^i, \boldsymbol{\sigma}^j\rangle_N)}\sum_{i=1}^{2k} H_{\geq 2}(\boldsymbol{\sigma}^i),\end{aligned}$$

and for $\widehat{H}(\boldsymbol{\rho}) = H_{\geq 2}(\rho) - \mathbb{E}[H_{\geq 2}(\boldsymbol{\rho})|\sum_{i=1}^{2k} H_{\geq 2}(\boldsymbol{\sigma}^i)]$,

$$\mathbb{E}\big[\widehat{H}(\boldsymbol{\rho}^1)\widehat{H}(\boldsymbol{\rho}^2)\big] = \xi(\langle \boldsymbol{\rho}^1, \boldsymbol{\rho}^2\rangle_N) - \frac{(\sum_{i=1}^{2k}\xi(\langle \boldsymbol{\rho}^1, \boldsymbol{\sigma}^i\rangle_N)(\sum_{i=1}^{2k}\xi(\langle \boldsymbol{\rho}^2, \boldsymbol{\sigma}^i\rangle_N)}{\sum_{i,j\in[2k]}\xi(\langle \boldsymbol{\sigma}^i, \boldsymbol{\sigma}^j\rangle_N)}.$$

For each $|q_1| \geq \delta$, and $q_2, \ldots, q_{2k} \in [0,1]$, consider the band $\mathsf{Band}_*(\{\boldsymbol{\sigma}^i\})$ of vectors $\boldsymbol{\rho}$ with $\langle \boldsymbol{\rho}, \boldsymbol{\sigma}^i\rangle = q_i$ for all $i \in [2k]$. Write $\boldsymbol{\rho} = \boldsymbol{x} + \sqrt{1-\tilde{q}^2}\boldsymbol{\tau}$ where $\boldsymbol{x} \in \mathrm{span}(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^{2k})$ and $\|\boldsymbol{\tau}\|^2 = N$, $\boldsymbol{\tau} \perp \mathrm{span}(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^{2k})$. Define the process $\overline{H}(\boldsymbol{\tau}) = \widehat{H}(\boldsymbol{\rho})$, which is a p-spin model with corresponding mixture $\widetilde{\xi}(t) = \widetilde{\xi}(t; \boldsymbol{q}, (\boldsymbol{\sigma}^i)_{i=1}^{2k})$ given by

$$\widetilde{\xi}(t; \boldsymbol{q}, (\boldsymbol{\sigma}^i)_{i=1}^{2k}) = \xi\big(\tilde{q}^2 + (1-\tilde{q}^2)t\big) - \frac{\big(\sum_{i=1}^{2k}\xi(q_i)\big)^2}{\sum_{i,j\in[2k]}\xi(\langle \boldsymbol{\sigma}^i, \boldsymbol{\sigma}^j\rangle_N)}.$$

We define the free energy

$$\Phi(\boldsymbol{q}; (\boldsymbol{\sigma}^i)_{i=1}^{2k}) := \frac{1}{N}\log\int_{\mathsf{Band}_*(\{\boldsymbol{\sigma}^i\})} e^{H_{\geq 2}(\boldsymbol{\rho})}\mu_0(\mathsf{d}\boldsymbol{\rho}).$$

Following the proof of Lemma 3.3 of [HS23b], the replica-symmetric bound implies that the following holds with high probability:

$$\Phi(\boldsymbol{q}; (\boldsymbol{\sigma}^i)_{i=1}^{2k}) \leq \frac{\sum_{i=1}^{2k}\xi(q_i)}{\sum_{i,j\in[2k]}\xi(\langle \boldsymbol{\sigma}^i, \boldsymbol{\sigma}^j\rangle_N)}\sum_{i=1}^{2k} H_{\geq 2}(\boldsymbol{\sigma}^i) + \frac{1}{2}\xi(1) - \frac{1}{2}\xi(\tilde{q}) + \frac{1}{2}\tilde{q} + \frac{1}{2}\log(1-\tilde{q}) + o_N(1). \tag{6.131}$$

By the generalized Bessel inequality, we have

$$\sum_{i=1}^{2k}\langle \boldsymbol{x},\boldsymbol{\sigma}^i\rangle_N^2 \le \|\boldsymbol{x}\|_N^2(2k)^{-1}\sum_{i,j\in[2k]}\langle \boldsymbol{\sigma}^i,\boldsymbol{\sigma}^j\rangle_N^2 = \|\boldsymbol{x}\|_N^2(2k)^{-1}(2k+(2k)^2\delta^2).$$

Hence,

$$\tilde{q}^2 = \|\boldsymbol{x}\|_N^2 \ge \frac{1}{1+2k\delta^2}\sum_{i=1}^{2k} q_i^2,$$

and since $\xi(0)=\xi'(0)=0$, this implies

$$\sum_{i=1}^{2k}\xi(q_i) \le \xi\big((1+2k\delta^2)^{1/2}\tilde{q}\big)\,.$$

We pick δ sufficiently small in c and k, and η small in δ. Given $\sum_{i=1}^{2k} H_{\ge 2}(\boldsymbol{\sigma}^i) = EN$ where $E \le \sum_{i,j\in[2k]}\xi(\langle \boldsymbol{\sigma}^i,\boldsymbol{\sigma}^j\rangle_N)+\eta$, whenever $q_1\ge\delta$, we have by assumption (6.101), with high probability

$$\Phi(\boldsymbol{q};(\boldsymbol{\sigma}^i)_{i=1}^{2k}) \le \frac{1}{2}\xi(1) - 10\eta.$$

Integrating over the $(q_i)_{i\le 2k}$ and using Gaussian concentration, we deduce that for $E\le \sum_{i,j\in[2k]}\xi(\langle \boldsymbol{\sigma}^i,\boldsymbol{\sigma}^j\rangle_N)+\eta$, we have

$$\mathbb{P}\left\{\int_{\boldsymbol{\rho}:\langle\boldsymbol{\rho},\boldsymbol{\sigma}_1\rangle_N>\delta} e^{H_{\ge 2}(\boldsymbol{\rho})}\mu_0(\mathsf{d}\boldsymbol{\rho}) \le e^{N(\xi(1)/2-9\eta)}\,\middle|\, \sum_{i=1}^{2k} H_{\ge 2}(\boldsymbol{\sigma}^i)=EN\right\} \ge 1-e^{-c(\eta)N}.$$

Up until now we worked with the Hamiltonian $H_{\ge 2}(\boldsymbol{\sigma})$, which does not include the term linear in $\boldsymbol{\sigma}$. Recall that $H(\boldsymbol{\sigma}) = \langle \boldsymbol{u},\boldsymbol{\sigma}\rangle + H_{\ge 2}(\boldsymbol{\sigma})$ and $\|\boldsymbol{u}\|\le c'\sqrt{N}$ so $|H(\boldsymbol{\sigma})-H_{\ge 2}(\boldsymbol{\sigma})|\le |\langle \boldsymbol{u},\boldsymbol{\sigma}\rangle| \le c'N$, assuming that $c'<\eta$, we have

$$\mathbb{P}\left\{\int_{\boldsymbol{\rho}:\langle\boldsymbol{\rho},\boldsymbol{\sigma}_1\rangle_N>\delta} e^{H(\boldsymbol{\rho})}\mu_0(\mathsf{d}\boldsymbol{\rho}) \le e^{N(\xi(1)/2-8\eta)}\,\middle|\, \sum_{i=1}^{2k} H_{\ge 2}(\boldsymbol{\sigma}^i)=EN\right\} \ge 1-e^{-c(\eta)N}.$$

Hence, under the same conditions

$$\mathbb{P}\left\{\boldsymbol{\sigma}^1\in T^c\,\middle|\,\sum_{i=1}^{2k} H_{\ge 2}(\boldsymbol{\sigma}^i)=EN\right\}\le e^{-c(\eta)N}\,.$$

Define the event

$$\mathcal{E}(\{\boldsymbol{\sigma}^i\}) := \left\{\sum_{i=1}^{2k} H_{\ge 2}(\boldsymbol{\sigma}^i) \ge N\left(\sum_{i,j\in[2k]}\xi(\langle \boldsymbol{\sigma}^i,\boldsymbol{\sigma}^j\rangle_N)+\eta\right)\right\}\,.$$

Thus, since $|H(\boldsymbol{\sigma})-H_{\ge 2}(\boldsymbol{\sigma})|\le c'N$, we can then conclude that

$$\begin{aligned}
&\mathbb{E}\left\{\int_{A_{2k}(\delta):\boldsymbol{\sigma}^1\in T^c} e^{\sum_{i=1}^{2k}H(\boldsymbol{\sigma}^i)}\mu_0^{\otimes 2k}(\mathsf{d}\boldsymbol{\sigma})\right\} = \mathbb{E}\left\{\int_{A_{2k}(\delta)} \mathbf{1}_{\boldsymbol{\sigma}^1\in T^c} e^{\sum_{i=1}^{2k}H(\boldsymbol{\sigma}^i)}\mu_0^{\otimes 2k}(\mathsf{d}\boldsymbol{\sigma})\right\}\\
&= \mathbb{E}\left\{\int_{A_{2k}(\delta)} \mathbb{P}\Big\{\boldsymbol{\sigma}^1\in T^c\Big|\sum_{i=1}^{2k}H_{\ge 2}(\boldsymbol{\sigma}^i)\Big\} e^{\sum_{i=1}^{2k}H(\boldsymbol{\sigma}^i)}\mu_0^{\otimes 2k}(\mathsf{d}\boldsymbol{\sigma})\right\}\\
&\le e^{-c(\eta)N+c'N}\mathbb{E}\int_{A_{2k}(\delta)} e^{\sum_{i=1}^{2k}H_{\ge 2}(\boldsymbol{\sigma}^i)}\mu_0^{\otimes 2k}(\mathsf{d}\boldsymbol{\sigma}) + e^{c'N}\mathbb{E}\int_{A_{2k}(\delta)} e^{\sum_{i=1}^{2k}H_{\ge 2}(\boldsymbol{\sigma}^i)}\mathbf{1}_{\mathcal{E}(\{\boldsymbol{\sigma}^i\})}\mu_0^{\otimes 2k}(\mathsf{d}\boldsymbol{\sigma})\\
&\le e^{-cN}\mathbb{E}\int_{A_{2k}(\delta)} e^{\sum_{i=1}^{2k}H(\boldsymbol{\sigma}^i)}\mu_0^{\otimes 2k}(\mathsf{d}\boldsymbol{\sigma}).
\end{aligned}$$

Here we assume $c' < c(\eta)/4$ and $c = c(\eta)/4$, and in the last step we used, for $U(\{\boldsymbol{\sigma}^i\}) := \sum_{i,j\in[2k]} \xi(\langle \boldsymbol{\sigma}^i, \boldsymbol{\sigma}^j \rangle_N)$,

$$\mathbb{E}\int_{A_{2k}(\delta)} e^{\sum_{i=1}^{2k} H_{\geq 2}(\boldsymbol{\sigma}^i)} \mathbf{1}_{\mathcal{E}(\{\boldsymbol{\sigma}^i\})} \mu_0^{\otimes 2k}(\mathsf{d}\boldsymbol{\sigma}) \leq \int_{A_{2k}(\delta)} \exp\Big\{N(1-s+s^2)U(\{\boldsymbol{\sigma}^i\}) - Ns\eta\Big\} \mu_0^{\otimes 2k}(\mathsf{d}\boldsymbol{\sigma}),$$

and chose δ, s suitably small. $\square$

The next lemma shows that integrals of S_N^{2k} with the product Gibbs measure are very precisely approximated by integral over tuples that are very close to orthogonal.

Lemma 6.7.10 (Near-orthogonal tuples dominate). *For $\delta > 0$ sufficiently small and appropriate $c, c_0 > 0$, if $\|\boldsymbol{u}\| \leq N^{c_0/2}$, then with probability $1 - e^{-cN}$ over $\boldsymbol{W}^{(2)}$, the following holds:*

1. *For quadratic Hamiltonians, the unrestricted partition function of $2k$ replicas is dominated by its restriction to $A_{2k}(N^{-1/2+c})$:*

$$Z_{\leq 2,2k}\big(A_{2k}(N^{-1/2+c})\big) \geq (1 - e^{-N^c}) \cdot \big(Z_{\leq 2}\big)^{2k}. \tag{6.132}$$

2. *The contribution of $A_{2k}(\delta) \setminus A_{2k}(N^{-1/2+c}) = \{(\boldsymbol{\sigma}^i)_{i\leq 2k} : \max_{i\neq j} |\langle \boldsymbol{\sigma}^i, \boldsymbol{\sigma}^j \rangle_N| \in [N^{-1/2+c}, \delta]\}$ is small:*

$$\mathbb{E}_{\geq 3} Z_{2k}\big(A_{2k}(\delta) \setminus A_{2k}(N^{-1/2+c})\big) \leq e^{-N^c + Nk\xi_{\geq 3}(1)} \big(Z_{\leq 2}\big)^{2k}. \tag{6.133}$$

3. *Annealing the restricted partition function over $H_{\geq 3}$ is roughly equivalent to complete annealing:*

$$\mathbb{E}_{\geq 3} Z_{2k}\big(A_{2k}(\delta)\big) \geq e^{-4k(\|\boldsymbol{u}\|+1)\sqrt{N}} \mathbb{E} Z_{2k}\big(A_{2k}(\delta)\big). \tag{6.134}$$

Proof. **Proof of 1.** By Lemma 6.7.5, for some constants $c_1, C_1 > 0$ that, with probability $1 - e^{-cN}$ over $\boldsymbol{W}^{(2)}$,

$$\begin{aligned} Z_{\leq 2,2k}\big(A_{2k}(N^{-1/2+c})\big) &\leq \sum_{i\neq j} \int_{S_N^{2k}} \mathbf{1}_{|\langle \boldsymbol{\sigma}^i, \boldsymbol{\sigma}^j \rangle_N| > N^{-1/2+c}} e^{\sum_{i=1}^{2k} H_{\leq 2}(\boldsymbol{\sigma}^i)} \mu_0^{\otimes 2k}(\mathsf{d}\boldsymbol{\sigma}) \\ &\leq e^{-N^c} \int_{S_N^{2k}} e^{\sum_{i=1}^{2k} H_{\leq 2}(\boldsymbol{\sigma}^i)} \mu_0^{\otimes 2k}(\mathsf{d}\boldsymbol{\sigma}), \end{aligned}$$

yielding (6.132).
Proof of 2. By a direct calculation, for any set $U \subseteq (S_N)^{2k}$:

$$\mathbb{E}_{\geq 3} Z_{2k}\big(U\big) = e^{Nk\xi_{\geq 3}(1)} \int_U e^{\sum_{i=1}^{2k} H_{\leq 2}(\boldsymbol{\sigma}^i)} \exp\left(N \sum_{i<j<2k} \xi_{\geq 3}(\langle \boldsymbol{\sigma}^i, \boldsymbol{\sigma}^j \rangle_N)\right) \mu_0^{\otimes 2k}(\mathsf{d}\boldsymbol{\sigma}).$$

Applying Lemma 6.7.5, we have for $t > 0$ and $\varepsilon_N = N^{-1/2+c}$ that, with probability $1 - e^{-cN}$ over $\boldsymbol{W}^{(2)}$,

$$\frac{1}{(Z_{\leq 2})^{2k}} e^{-Nk\xi_{\geq 3}(1)} \mathbb{E}_{\geq 3} Z_{2k}\Big(A_{2k}(t+\varepsilon_N) \setminus A_{2k}(t)\Big) \leq \tag{6.135}$$

$$\begin{aligned} &\frac{1}{(Z_{\leq 2})^{2k}} \int_{\max_{i\neq j} |\langle \boldsymbol{\sigma}^i, \boldsymbol{\sigma}^j \rangle_N| \in [t, t+\varepsilon_N]} \exp\left(\sum_{i=1}^{2k} H_{\leq 2}(\boldsymbol{\sigma}^i) + N \sum_{i\neq j\in[2k]} \xi_{\geq 3}(\langle \boldsymbol{\sigma}^i, \boldsymbol{\sigma}^j \rangle_N)\right) \mu_0^{\otimes 2k}(\mathsf{d}\boldsymbol{\sigma}) \leq \\ &\leq \exp\Big\{-cN\big(t - \|\boldsymbol{u}\|_N^2\big)_+^2 + N(2k)^2 \xi_{\geq 3}(t+\varepsilon_N)\Big\}. \end{aligned} \tag{6.136}$$

Under the assumption $\|\boldsymbol{u}\|_N^2 \leq N^{c_0-1}$, $c_0 < c + 1/2$, summing over the range $N^{-1/2+c} < |t| \leq \delta$, we obtain the following with probability $1 - e^{-cN}$ over $\boldsymbol{W}^{(2)}$,

$$\begin{aligned} &\int_{\max_{i\neq j} |\langle \boldsymbol{\sigma}^i, \boldsymbol{\sigma}^j \rangle_N| \in [N^{-1/2+c}, \delta]} \exp\left(\sum_{i=1}^{2k} H_{\leq 2}(\boldsymbol{\sigma}^i) + N \sum_{i\neq j\in[2k]} \xi_{\geq 3}(\langle \boldsymbol{\sigma}^i, \boldsymbol{\sigma}^j \rangle_N)\right) \mu_0^{\otimes 2k}(\mathsf{d}\boldsymbol{\sigma}) \\ &\leq \exp(-N^c) \int_{S_N^{2k}} \exp\left(\sum_{i=1}^{2k} H_{\leq 2}(\boldsymbol{\sigma}^i)\right) \mu_0^{\otimes 2k}(\mathsf{d}\boldsymbol{\sigma}). \end{aligned}$$

This gives (6.133).
Proof of 3. Note that

$$
\begin{aligned}
e^{-Nk\xi_{\geq 3}(1)}\mathbb{E}_{\geq 3} Z_{2k}\big(A_{2k}(N^{-1/2+c})\big) = & \\
&= \int_{A_{2k}(N^{-1/2+c})} \exp\left(\sum_{i=1}^{2k} H_{\leq 2}(\boldsymbol{\sigma}^i) + N \sum_{i\neq j\in[2k]} \xi_{\geq 3}(\langle \boldsymbol{\sigma}^i, \boldsymbol{\sigma}^j\rangle_N)\right) \mu_0^{\otimes 2k}(\mathsf{d}\boldsymbol{\sigma}) \\
&= (1+O(N^{-1/2+3c}))\cdot Z_{\leq 2,2k}\Big(A_{2k}(N^{-1/2+c})\Big). \qquad (6.137)
\end{aligned}
$$

Therefore, using Eq. (6.132), we get

$$
\mathbb{E}_{\geq 3} Z_{2k}\big(A_{2k}(N^{-1/2+c})\big) = (1+O(N^{-1/2+3c}))e^{Nk\xi_{\geq 3}(1)}\big(Z_{\leq 2}\big)^{2k}. \qquad (6.138)
$$

Also,

$$
\mathbb{E} Z_{\leq 2,2k}\Big(A_{2k}(N^{-1/2+c})\Big) \leq e^{2k\|\boldsymbol{u}\|\sqrt{N}} \exp\big(k\beta_2^2 N + (2k)^2\beta_2^2 N^{2c}\big). \qquad (6.139)
$$

On the other hand, Lemma 6.7.3 readily implies that with probability at least $1-e^{-cN}$,

$$
(Z_{\leq 2})^{2k} \geq e^{-o(\sqrt{N})} \exp(k\beta_2^2 N). \qquad (6.140)
$$

Combining Eqs. (6.139) and (6.139), we get

$$
\mathbb{E} Z_{\leq 2,2k}\big(A_{2k}(N^{-1/2+c})\big) \leq e^{2k(1+\|\boldsymbol{u}\|)\sqrt{N}}(Z_{\leq 2})^{2k}. \qquad (6.141)
$$

Finally, using Eq. (6.138) together with the last display, we get

$$
\mathbb{E}_{\geq 3} Z_{2k}\big(A_{2k}(N^{-1/2+c})\big) \geq e^{-3k(\|\boldsymbol{u}\|+1)\sqrt{N}} \mathbb{E} Z_{2k}\big(A_{2k}(N^{-1/2+c})\big). \qquad (6.142)
$$

Combining this with Eq. (6.133) yields the claim.

□

Lemma 6.7.11. *For any $m \geq 2$, there exists a constant $c > 0$ such that, for $T = T(\delta)$,*

$$
Z_m\big(A_m(T,\delta)\big) \leq Z(T)^m \leq (1+e^{-cN})\cdot Z_m\big(A_m(\delta)\big) + e^{-cN+Nm\xi(1)/2}. \qquad (6.143)
$$

Proof. The left hand inequality is obvious since $A_m(T,\delta) \subseteq T^{\otimes m}$. For the right inequality consider first the case $m=2$. Then we have

$$
\begin{aligned}
Z(T)^2 &\leq Z_2\big(A_2(T,\delta)\big) + \int_{T\times T} \mathbf{1}_{|\langle \boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2\rangle_N|\geq \delta} e^{H(\boldsymbol{\sigma}^1)+H(\boldsymbol{\sigma}^2)} \mu_0^{\otimes 2}(\mathsf{d}\boldsymbol{\sigma}) && (6.144)\\
&\leq Z_2\big(A_2(T,\delta)\big) + \int_T e^{H(\boldsymbol{\sigma}^1)}\left[\int_T \mathbf{1}_{|\langle \boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2\rangle_N|\geq \delta} e^{H(\boldsymbol{\sigma}^2)}\mu_0(\mathsf{d}\boldsymbol{\sigma}^2)\right]\mu_0(\mathsf{d}\boldsymbol{\sigma}^1) && (6.145)\\
&\leq Z_2\big(A_2(T,\delta)\big) + \int_T e^{H(\boldsymbol{\sigma}^1)} e^{-\delta N + N\xi(1)/2}\mu_0(\mathsf{d}\boldsymbol{\sigma}^1) && (6.146)\\
&\leq Z_2\big(A_2(T,\delta)\big) + e^{-N\delta+N\xi(1)/2} Z(T) && (6.147)\\
&\leq Z_2\big(A_2(T,\delta)\big) + e^{-N\delta+N\xi(1)} + e^{-N\delta} Z(T)^2. && (6.148)
\end{aligned}
$$

where in the last step we used the AM-GM inequality. Solving this inequality for $Z(T)^2$, we get:

$$
Z(T)^2 \leq (1+e^{-cN}) Z_2\big(A_2(T,\delta)\big) + 2e^{-\delta N+N\xi(1)}. \qquad (6.149)
$$

which proves the claim for $m=2$.

Consider now $m \geq 2$. Note that

$$\begin{aligned}
&\int_{T(\delta)^m} e^{\sum_{i=1}^m H(\boldsymbol{\sigma}^i)} \mu_0^{\otimes m}(\mathsf{d}\boldsymbol{\sigma}) - \int_{A_m(T(\delta),\delta)} e^{\sum_{i=1}^m H(\boldsymbol{\sigma}^i)} \mu_0^{\otimes 2m}(\mathsf{d}\boldsymbol{\sigma}) \\
&\leq \sum_{i \neq j} \left(\int_{T(\delta)} e^{H(\boldsymbol{\sigma})} \mu_0(\mathsf{d}\boldsymbol{\sigma}) \right)^{m-2} \int_{\boldsymbol{\sigma}^i, \boldsymbol{\sigma}^j \in T(\delta) : |\langle \boldsymbol{\sigma}^i, \boldsymbol{\sigma}^j \rangle_N| > \delta} e^{H(\boldsymbol{\sigma}^i) + H(\boldsymbol{\sigma}^j)} \mu_0(\mathsf{d}\boldsymbol{\sigma}^i) \mu_0(\mathsf{d}\boldsymbol{\sigma}^j)
\end{aligned}$$

whence

$$\begin{aligned}
Z(T)^m - Z_m\big(A_m(T,\delta)\big) &\leq m^2 Z_2(T^{\otimes 2} \setminus A_2(T,\delta)) \cdot Z(T)^{m-2} \\
&\leq m^2 \cdot Z(T)^{m-1} \cdot e^{-N\delta + N\xi(1)/2} \,,
\end{aligned}$$

where in the last inequality we used Eq. (6.147). Using again the AM-GM inequality, we get

$$Z(T)^m - Z_m\big(A_m(T,\delta)\big) \leq m^2 e^{-N\delta} Z(T)^m + m^2 e^{-N\delta + Nm\xi(1)/2} \,,$$

which yields the claim. □

Proof of Lemma 6.7.8

We next prove Lemma 6.7.8. In the proof, we let c denote small absolute constants that can change from line to line. We will first prove the partition function estimate, Eq. (6.127) and then the magnetization estimate, Eq. (6.129).

Estimating the partition function, Eq. (6.127). By Eq. (6.133) in Lemma 6.7.10, with probability $1 - e^{-cN}$ over $\boldsymbol{W}^{(2)}$,

$$\begin{aligned}
&\int_{\max_{i \neq j} |\langle \boldsymbol{\sigma}^i, \boldsymbol{\sigma}^j \rangle_N| \in [N^{-1/2+c}, \delta]} \exp\left(\sum_{i=1}^{2k} H_{\leq 2}(\boldsymbol{\sigma}^i) + N \sum_{i<j \leq 2k} \xi_{\geq 3}(\langle \boldsymbol{\sigma}^i, \boldsymbol{\sigma}^j \rangle_N) \right) \mu_0^{\otimes 2k}(\mathsf{d}\boldsymbol{\sigma}) \\
&\leq \exp(-N^c) \int_{S_N^{2k}} \exp\left(\sum_{i=1}^{2k} H_{\leq 2}(\boldsymbol{\sigma}^i) \right) \mu_0^{\otimes 2k}(\mathsf{d}\boldsymbol{\sigma}). \qquad (6.150)
\end{aligned}$$

On $A_{2k}(N^{-1/2+c})) = \{|\langle \boldsymbol{\sigma}^i, \boldsymbol{\sigma}^j \rangle_N| \leq N^{-1/2+c}\} \ \forall i \neq j\}$, we can expand

$$\exp\left(N \sum_{i<j} \xi_{\geq 3}(\langle \boldsymbol{\sigma}^i, \boldsymbol{\sigma}^j \rangle_N) \right) = \sum_{\ell=0}^{L-1} \frac{1}{\ell!} \Big(N \sum_{i<j} \xi_{\geq 3}(\langle \boldsymbol{\sigma}^i, \boldsymbol{\sigma}^j \rangle_N) \Big)^\ell + O(N^{-L/2+3cL}).$$

Thus, for $T = T(\delta)$, the following holds with probability at least $1 - e^{-cN}$ over $\boldsymbol{W}_2$,

$$\begin{aligned}
&\mathbb{E}_{\geq 3}\Big\{ \big(Z(T) - \mathbb{E}_{\geq 3} Z(T) \big)^{2k} \Big\} \qquad (6.151) \\
&\overset{(a)}{\leq} \sum_{r \leq 2k} \binom{2k}{2k-r} (-1)^r \big(\mathbb{E}_{\geq 3} Z(T) \big)^{2k-r} \cdot \mathbb{E}_{\geq 3} Z\big(A_r(T,\delta)\big) \\
&\quad + e^{-Nc} \sum_{r \leq 2k} \binom{2k}{2k-r} \big(\mathbb{E}_{\geq 3} Z(T) \big)^{2k-r} \cdot \Big(\mathbb{E}_{\geq 3} Z\big(A_r(T,\delta)\big) + e^{Nr\xi(1)/2} \Big) \\
&\overset{(b)}{\leq} \sum_{r \leq 2k} \binom{2k}{2k-r} (-1)^r \big(\mathbb{E}_{\geq 3} Z(T) \big)^{2k-r} \cdot \mathbb{E}_{\geq 3} Z\big(A_r(T,\delta)\big) \qquad (6.152) \\
&\quad + e^{-Nc} \max_{r \leq 2k} e^{N(2k-r)\xi(1)/2} \cdot \Big(\mathbb{E}_{\geq 3} Z\big(A_r(T,\delta)\big) + e^{Nr\xi(1)/2} \Big) \,, \qquad (6.153)
\end{aligned}$$

where (a) follows from Lemma 6.7.10, (b) holds because $\mathbb{E}_{\geq 3} Z(T) \leq e^{c'N} \mathbb{E} Z$ with the claimed probability by Markov inequality.

We define the error terms

$$\mathsf{Err}_1 := e^{-cN+Nk\xi(1)} + e^{-cN} \max_{1\le r\le 2k} \Big(\mathbb{E}_{\ge 3} Z(A_r(T,\delta))\Big)^{2k/r} + \mathbb{E}_{\ge 3} Z(A_{2k}(\delta)\cap\{\boldsymbol{\sigma}^1\in T^c\}), \tag{6.154}$$

$$\mathsf{Err}_2 := N^{-L/2} e^{Nk\xi_{\ge 3}(1)} Z_{\le 2,2k}(A_{2k}(\delta))\,, \tag{6.155}$$

so that the bound (6.153) implies

$$\mathbb{E}_{\ge 3}\Big\{\big(Z(T)-\mathbb{E}_{\ge 3}Z(T)\big)^{2k}\Big\} \le \sum_{r\le 2k}\binom{2k}{2k-r}(-1)^r\big(\mathbb{E}_{\ge 3}Z(T)\big)^{2k-r}\cdot\mathbb{E}_{\ge 3}Z\big(A_r(T,\delta)\big) + O_k(\mathsf{Err}_1)\,. \tag{6.156}$$

Next note that

$$\begin{aligned}
&\big(\mathbb{E}_{\ge 3}Z(T)\big)^{2k-r}\cdot\mathbb{E}_{\ge 3}Z\big(A_r(T,\delta)\big)\\
&= e^{Nk\xi_{\ge 3}(1)}\left(\int_{T(\delta)} e^{H_{\le 2}(\boldsymbol{\sigma})}\mu_0(\mathsf{d}\boldsymbol{\sigma})\right)^{2k-r}\cdot\\
&\qquad\cdot\int_{A_r(\delta)} e^{\sum_{i=1}^r H_{\le 2}(\boldsymbol{\sigma}^i)}\exp\left(N\sum_{i<j}\xi_{\ge 3}(\langle\boldsymbol{\sigma}^i,\boldsymbol{\sigma}^j\rangle_N)\right)\mu_0^{\otimes r}(\mathsf{d}\boldsymbol{\sigma}) + O_k(\mathsf{Err}_1)\\
&= e^{Nk\xi_{\ge 3}(1)}\int_{A_{2k}(\delta)}\exp\left(\sum_{i'=1}^{2k-r}H_{\le 2}((\boldsymbol{\sigma}')^{i'}) + \sum_{i=1}^{r}H_{\le 2}(\boldsymbol{\sigma}^i)\right)\cdot\\
&\qquad\cdot\left\{\sum_{\ell=0}^{L-1}\frac{1}{\ell!}\left(N\sum_{i<j\le r}\xi_{\ge 3}(\langle\boldsymbol{\sigma}^i,\boldsymbol{\sigma}^j\rangle_N)\right)^{\ell}\right\}\mu_0^{\otimes r}(\mathsf{d}\boldsymbol{\sigma})\mu_0^{\otimes(2k-r)}(\mathsf{d}\boldsymbol{\sigma}') + O_k(\mathsf{Err}_1+\mathsf{Err}_2),
\end{aligned}$$

where the last inequality holds with probability $1-e^{-cN}$ over $\boldsymbol{W}^{(2)}$ by Eq. (6.133).

Substituting in Eq. (6.156), we get

$$\begin{aligned}
&\mathbb{E}_{\ge 3}\Big\{\big(Z(T)-\mathbb{E}_{\ge 3}Z(T)\big)^{2k}\Big\}\\
&\le e^{Nk\xi_{\ge 3}(1)}\int_{A_{2k}(\delta)} e^{\sum_{i=1}^{2k}H_{\le 2}(\boldsymbol{\sigma}^i)}\sum_{\ell\le L}\frac{1}{\ell!}\sum_{r\le 2k}(-1)^r\sum_{S\subseteq[2k]:|S|=r}\Big(N\sum_{i<j\in S}\xi_{\ge 3}(\langle\boldsymbol{\sigma}^i,\boldsymbol{\sigma}^j\rangle_N)\Big)^{\ell}\mu_0^{\otimes 2k}(\mathsf{d}\boldsymbol{\sigma})\\
&\qquad + O_k(\mathsf{Err}_1+\mathsf{Err}_2),
\end{aligned} \tag{6.157}$$

We can expand the ℓ-th power in (6.157), thus getting a sum indexed by sets of pairs $S=\{(i_t,j_t): t\le \ell\}\subseteq\binom{[2k]}{2}$. Denoting by $n(S)$ the number of distinct elements of $[2k]$ appearing in S, the coefficient of such therm is its coefficient is, for $n(S)<2k$,

$$\sum_{\ell\le r\le 2k}(-1)^r\binom{2k-n(S)}{r-n(S)} = 0$$

for $|\{i_t,j_t : t\le\ell\}|<2k$. Hence, taking $L<k$, we have

$$\mathbb{E}_{\ge 3}\Big\{\big(Z(T)-\mathbb{E}_{\ge 3}Z(T)\big)^{2k}\Big\} = O_k(\mathsf{Err}_1+\mathsf{Err}_2)\,. \tag{6.158}$$

We now estimate the error terms.

Error term Err_2. Using Lemma 6.7.7, we have

$$\begin{aligned}
\big(\mathbb{E}_{\ge 3}Z(T)\big)^{2k} &= \Big(\mathbb{E}_{\ge 3}Z - \mathbb{E}_{\ge 3}Z(T^c)\Big)^{2k}\\
&\ge (1-e^{-cN/8})\big(\mathbb{E}_{\ge 3}Z\big)^{2k}\\
&\ge c\, e^{Nk\xi_{\ge 3}(1)}\big(Z_{\le 2}\big)^{2k}\\
&\ge c\, e^{Nk\xi_{\ge 3}(1)}Z_{\le 2}\big(A_{2k}(\delta)\big).
\end{aligned}$$

From this estimate, we obtain with probability at least $1-\exp(-cN/8)$ over $\boldsymbol{W}^{(2)}$ that

$$\mathsf{Err}_2 \le C \cdot N^{-L/2} \cdot \big(\mathbb{E}_{\ge 3} Z(T)\big)^{2k}. \tag{6.159}$$

Error term Err_1. Using Lemma 6.7.9 by Markov inequality, with probability $1-\exp(-cN/2)$ over $\boldsymbol{W}^{(2)}$,

$$\mathbb{E}_{\ge 3} Z\big(A_{2k}(\delta) \cap \{\boldsymbol{\sigma}^1 \in T^c\}\big) \le e^{-cN/2} \mathbb{E} Z(A_{2k}(\delta)). \tag{6.160}$$

Further using Eq. (6.134) in Lemma 6.7.10, and using the assumption on $\|\boldsymbol{u}\|_2$, with probability $1-\exp(-cN/4)$ over $\boldsymbol{W}^{(2)}$,

$$\mathbb{E}_{\ge 3} Z\big(A_{2k}(\delta) \cap \{\boldsymbol{\sigma}^1 \in T^c\}\big) \le e^{-cN/2} \mathbb{E}_{\ge 3} Z(A_{2k}(\delta)).$$

Hence, with probability at least $1-\exp(-cN/8)$ over $\boldsymbol{W}^{(2)}$,

$$\mathsf{Err}_1 \le e^{-cN+Nk\xi(1)} + e^{-cN} \max_{1\le r\le 2k} \Big(\mathbb{E}_{\ge 3} Z(A_r(\delta))\Big)^{2k/r} \tag{6.161}$$

Further, with probability at least $1-\exp(-cN/8)$ over $\boldsymbol{W}^{(2)}$,

$$\begin{aligned}\mathbb{E}_{\ge 3} Z\big(A_r(\delta)\big) &= \mathbb{E}_{\ge 3} Z\big(A_r(N^{-1/2+c})\big) + \mathbb{E}_{\ge 3} Z\big(A_r(\delta) \setminus A_r(N^{-1/2+c})\big)\\ &\le 2\, e^{Nr\xi_{\ge 3}(1)/2} Z_{\le 2}\big(A_r(N^{-1/2+c})\big) + e^{-N^c + Nr\xi_{\ge 3}(1)/2} (Z_{\le 2})^r \,,\end{aligned} \tag{6.162}$$

where in the last line we used Eq. (6.133), and the fact that

$$\begin{aligned}\mathbb{E}_{\ge 3} Z\big(A_r(N^{-1/2+c})\big) &= \int_{A_r(N^{-1/2+c})} e^{\sum_{i=1}^r H_{\le 2}(\boldsymbol{\sigma}^i)} \exp\left(\frac{N}{2}\sum_{i,j\le r} \xi_{\ge 3}(\langle \boldsymbol{\sigma}^i, \boldsymbol{\sigma}^j\rangle_N)\right) \mu_0^{\otimes r}(\mathsf{d}\boldsymbol{\sigma})\\ &\le \big(1+O(N^{-1/2+3c})\big)\, e^{Nk\xi_{\ge 3}(1)} Z_{\le 2}\big(A_r(N^{-1/2+c})\big)\,.\end{aligned}$$

Using Eq. (6.106) in Eq. (6.162), we get

$$\mathbb{E}_{\ge 3} Z\big(A_r(\delta)\big) \le N^C e^{Nr\xi(1)/2}\,, \tag{6.163}$$

whence Eq. (6.161) simplifies to

$$\mathsf{Err}_1 \le e^{-cN+Nr\xi(1)}\,. \tag{6.164}$$

On the other hand, by Lemma 6.7.7 and Markov inequality, with probability $1-\exp(-cN/4)$ over $\boldsymbol{W}^{(2)}$,

$$\mathbb{E}_{\ge 3} \int_{T(\delta)^c} e^{H(\boldsymbol{\sigma})} \mu_0(\mathsf{d}\boldsymbol{\sigma}) \le e^{-cN/4} e^{N\xi(1)/2}.$$

Using Lemma 6.7.3, we obtain that, with probability at least $1-\exp(-cN/8)$ over $\boldsymbol{W}^{(2)}$,

$$\begin{aligned}\mathbb{E}_{\ge 3} Z &= \int e^{N\xi_{\ge 3}(1)/2} e^{H_{\le 2}(\boldsymbol{\sigma})} \mu_0(\mathsf{d}\boldsymbol{\sigma})\\ &\ge e^{N\xi(1)/2 - cN/10}\,,\end{aligned}$$

whence Eq. (6.164) yields

$$\mathsf{Err}_1 \le e^{-cN/16} \big(\mathbb{E}_{\ge 3} Z(T)\big)^{2k}. \tag{6.165}$$

We also note here the estimate

$$\mathbb{E}_{\ge 3} Z(T) = \mathbb{E}_{\ge 3} Z - \mathbb{E}_{\ge 3} Z(T^c) \ge (1-e^{-cN/10}) \mathbb{E}_{\ge 3} Z, \tag{6.166}$$

which holds with probability at least $1-\exp(-cN/8)$ over $\boldsymbol{W}^{(2)}$, as claimed in Eq. (6.128).

Combining the error estimates (6.165), (6.159) in the moment bound (6.158), we get, with probability at least $1-\exp(-Nc)$ with respect to $\boldsymbol{W}_2$,

$$\mathbb{E}_{\geq 3}\Big\{\big(Z(T)-\mathbb{E}_{\geq 3}Z(T)\big)^{2k}\Big\} \leq C\, N^{-L/2}\left(\mathbb{E}_{\geq 3}Z(T)\right)^{2k}. \tag{6.167}$$

Adjusting c, we have

$$\mathbb{P}\left(|Z-\mathbb{E}_{\geq 3}Z(T(\delta))| > \varepsilon\,\mathbb{E}_{\geq 3}Z(T(\delta))\right) \leq \varepsilon^{-2L}N^{-L/2}+e^{-cN}.$$

Estimating the magnetization, Eq. (6.129). We next apply the same argument to the magnetization. First, we note that

$$\mathbb{E}_{\geq 3}\left\{\left(\int_{T(\delta)}\sigma_1 e^{H(\boldsymbol{\sigma})}\mu_0(\mathsf{d}\boldsymbol{\sigma})\right)^{2k}\right\} = \mathbb{E}_{\geq 3}\int_{T(\delta)^{2k}}\prod_{i=1}^{2k}\sigma_1^i\, e^{\sum_{i=1}^{2k}H(\boldsymbol{\sigma}^i)}\,\mu_0^{\otimes 2k}(\mathsf{d}\boldsymbol{\sigma}) \tag{6.168}$$

$$= \mathbb{E}_{\geq 3}\int_{A_{2k}(\delta)}\prod_{i=1}^{2k}\sigma_1^i\, e^{\sum_{i=1}^{2k}H(\boldsymbol{\sigma}^i)}\mu^{\otimes 2k}(\mathsf{d}\boldsymbol{\sigma})+\mathsf{Err}_3\,, \tag{6.169}$$

where

$$\mathsf{Err}_3 := \mathbb{E}_{\geq 3}\int_{T(\delta)^{2k}}\prod_{i=1}^{2k}\sigma_1^i\, e^{\sum_{i=1}^{2k}H(\boldsymbol{\sigma}^i)}\mu^{\otimes 2k}(\mathsf{d}\boldsymbol{\sigma})-\mathbb{E}_{\geq 3}\int_{A_{2k}(\delta)}\prod_{i=1}^{2k}\sigma_1^i\, e^{\sum_{i=1}^{2k}H(\boldsymbol{\sigma}^i)}\mu^{\otimes 2k}(\mathsf{d}\boldsymbol{\sigma}). \tag{6.170}$$

We have

$$\begin{aligned}
&\left|\int_{T(\delta)^{2k}}\prod_{i=1}^{2k}\sigma_1^i e^{\sum_{i=1}^{2k}H(\boldsymbol{\sigma}^i)}\mu_0^{\otimes 2k}(\mathsf{d}\boldsymbol{\sigma})-\int_{A_{2k}(T(\delta),\delta)}\prod_{i=1}^{2k}\sigma_1^i e^{\sum_{i=1}^{2k}H(\boldsymbol{\sigma}^i)}\mu_0^{\otimes 2k}(\mathsf{d}\boldsymbol{\sigma})\right|\\
&=\left|\int_{T(\delta)^{2k}:\max_{i\neq j}|\langle\boldsymbol{\sigma}^i,\boldsymbol{\sigma}^j\rangle_N|>\delta}\prod_{i=1}^{2k}\sigma_1^i e^{\sum_{i=1}^{2k}H(\boldsymbol{\sigma}^i)}\mu_0^{\otimes 2k}(\mathsf{d}\boldsymbol{\sigma})\right|\\
&\overset{(a)}{\leq} N^k(2k)^2e^{-cN+N\xi(1)/2}Z\big(T(\delta)\big)^{2k-1}\\
&\overset{(b)}{\leq} e^{-cN/(2k)+Nk\xi(1)}+e^{-cN/(2k)}Z\big(T(\delta)\big)^{2k}\\
&\overset{(c)}{\leq} e^{-cN/(2k)+Nk\xi(1)}+e^{-cN/(2k)}\left(\mathbb{E}_{\geq 3}Z\big(T(\delta)\big)\right)^{2k}\,,
\end{aligned} \tag{6.171}$$

where in (a) we used Lemma 6.7.7, in (b) the AM-GM inequality, and (c) holds with probability at least $1-\exp(-cN)$ by Eq. (6.126).

Using Eq. (6.171) and Lemma 6.7.3 we obtain that, with probability at least $1-e^{-cN}$ over $\boldsymbol{W}^{(2)}$,

$$|\mathsf{Err}_3| \leq e^{-cN}\left(\mathbb{E}_{\geq 3}Z\big(T(\delta)\big)\right)^{2k}. \tag{6.172}$$

Turning to the main term in Eq. (6.169),

$$\begin{aligned}
&\mathbb{E}_{\geq 3}\int_{A_{2k}(\delta)}\prod_{i=1}^{2k}\sigma_1^i\, e^{\sum_{i=1}^{2k}H(\boldsymbol{\sigma}^i)}\mu^{\otimes 2k}(\mathsf{d}\boldsymbol{\sigma})\\
&= e^{Nk\xi_{\geq 3}(1)}\int_{A_{2k}(\delta)}\prod_{i=1}^{2k}\sigma_1^i\exp\left\{\sum_{i=1}^{2k}H_{\leq 2}(\boldsymbol{\sigma}^i)+\frac{N}{2}\sum_{i\neq j}\xi_{\geq 3}(\langle\boldsymbol{\sigma}^i,\boldsymbol{\sigma}^j\rangle_N)\right\}\mu_0^{\otimes 2k}(\mathsf{d}\boldsymbol{\sigma})
\end{aligned}$$

By Eqs. (6.132) and (6.133) in Lemma 6.7.10, we can bound

$$\begin{aligned}
&\mathbb{E}_{\geq 3} \int_{A_{2k}(\delta)} \prod_{i=1}^{2k} \sigma_1^i \, e^{\sum_{i=1}^{2k} H(\boldsymbol{\sigma}^i)} \mu^{\otimes 2k}(\mathsf{d}\boldsymbol{\sigma}) \\
&= e^{Nk\xi_{\geq 3}(1)} \int_{A_{2k}(N^{-1/2+c})} \prod_{i=1}^{2k} \sigma_1^i \exp\left\{ \sum_{i=1}^{2k} H_{\leq 2}(\boldsymbol{\sigma}^i) + \frac{N}{2} \sum_{i \neq j} \xi_{\geq 3}(\langle \boldsymbol{\sigma}^i, \boldsymbol{\sigma}^j \rangle_N) \right\} \mu_0^{\otimes 2k}(\mathsf{d}\boldsymbol{\sigma}) \\
&\qquad + O\Big(N^k e^{-N^c + Nk\xi_{\geq 3}(1)} Z_{\leq 2, 2k}\big(A_{2k}(\delta)\big) \Big) .
\end{aligned}$$

To bound the first term, using Lemma 6.7.5,

$$\begin{aligned}
&\int_{A_{2k}(\delta)} \prod_{i=1}^{2k} \sigma_1^i \exp\left\{ \sum_{i=1}^{2k} H_{\leq 2}(\boldsymbol{\sigma}^i) \right\} \mu_0^{\otimes 2k}(\mathsf{d}\boldsymbol{\sigma}) \\
&= \int_{S_N^{2k}} \prod_{i=1}^{2k} \sigma_1^i \exp\left\{ \sum_{i=1}^{2k} H_{\leq 2}(\boldsymbol{\sigma}^i) \right\} \mu_0^{\otimes 2k}(\mathsf{d}\boldsymbol{\sigma}) + O_k\left(N^k e^{-c\delta^2 N} \big(Z_{\leq 2}\big)^{2k} \right) \\
&= \left(\int_{S_N} \sigma_1 \exp\{H_{\leq 2}(\boldsymbol{\sigma})\} \, \mu_0(\mathsf{d}\boldsymbol{\sigma}) \right)^{2k} + O_k\left(N^k e^{-c\delta^2 N} \big(Z_{\leq 2}\big)^{2k} \right) .
\end{aligned}$$

By Lemma 6.7.3, we then obtain

$$\int_{A_{2k}(\delta)} \prod_{i=1}^{2k} \sigma_1^i \exp\left\{ \sum_{i=1}^{2k} H_{\leq 2}(\boldsymbol{\sigma}^i) \right\} \mu_0^{\otimes 2k}(\mathsf{d}\boldsymbol{\sigma}) \leq C_k \left(|u_1|^{2k} + N^k e^{-c\delta^2 N} \right) (Z_{\leq 2})^{2k} . \tag{6.173}$$

On the other hand, by taking the Taylor expansion of $\exp\left\{ \frac{N}{2} \sum_{i \neq j} \xi_{\geq 3}(\langle \boldsymbol{\sigma}^i, \boldsymbol{\sigma}^j \rangle_N) \right\}$ up to terms of order $L = Ck$ for $C > 2$, we obtain that, for $\xi_{\geq 3, \leq \ell}(s) = \sum_{3 \leq p \leq \ell} \beta_p^2 s^p$,

$$\begin{aligned}
&\frac{1}{(Z_{\leq 2})^{2k}} \int_{A_{2k}(N^{-1/2+c})} \prod_{i=1}^{2k} \sigma_1^i \left(\exp\left(N \sum_{i<j} \xi_{\geq 3}(\langle \boldsymbol{\sigma}^i, \boldsymbol{\sigma}^j \rangle_N) \right) - 1 \right) e^{\sum_{i=1}^{2k} H_{\leq 2}(\boldsymbol{\sigma}^i)} \mu_0^{\otimes 2k}(\mathsf{d}\boldsymbol{\sigma}) \\
&= O(N^{-k}) + \frac{1}{(Z_{\leq 2})^{2k}} \sum_{\ell \leq L} \frac{N^\ell}{\ell!} \int_{A_{2k}(N^{-1/2+c})} \prod_{i=1}^{2k} \sigma_1^i \left(\sum_{i<j} \xi_{\geq 3}(\langle \boldsymbol{\sigma}^i, \boldsymbol{\sigma}^j \rangle_N) \right)^\ell e^{\sum_{i=1}^{2k} H_{\leq 2}(\boldsymbol{\sigma}^i)} \mu_0^{\otimes 2k}(\mathsf{d}\boldsymbol{\sigma}) \\
&= O(N^{-k}) + \frac{1}{(Z_{\leq 2})^{2k}} \sum_{\ell \leq L} \frac{N^\ell}{\ell!} \int_{A_{2k}(N^{-1/2+c})} \prod_{i=1}^{2k} \sigma_1^i \left(\sum_{i<j} \xi_{\geq 3, \leq 4k}(\langle \boldsymbol{\sigma}^i, \boldsymbol{\sigma}^j \rangle_N) \right)^\ell e^{\sum_{i=1}^{2k} H_{\leq 2}(\boldsymbol{\sigma}^i)} \mu_0^{\otimes 2k}(\mathsf{d}\boldsymbol{\sigma}) \\
&\overset{(a)}{=} O(N^{-k} + e^{-N^c}) + \frac{1}{(Z_{\leq 2})^{2k}} \sum_{\ell \leq L} \frac{N^\ell}{\ell!} \int_{S_N^{2k}} \prod_{i=1}^{2k} \sigma_1^i \left(\sum_{i<j} \xi_{\geq 3, \leq 2k}(\langle \boldsymbol{\sigma}^i, \boldsymbol{\sigma}^j \rangle_N) \right)^\ell e^{\sum_{i=1}^{2k} H_{\leq 2}(\boldsymbol{\sigma}^i)} \mu_0^{\otimes 2k}(\mathsf{d}\boldsymbol{\sigma}) \\
&\overset{(b)}{=} O(N^{-k} + e^{-N^c}) + (1 + \|\boldsymbol{u}\|)^{O(k^2)} O_k \left(\sum_{\ell \leq k} |u_1|^{2k-2\ell} N^{-\ell} + \sum_{k < \ell \leq L} N^{-\ell/2 - k/2} \right) &(6.174) \\
&= (1 + \|\boldsymbol{u}\|)^{O(k^2)} O(N^{-k} + e^{-N^c} + |u_1|^{2k}) , &(6.175)
\end{aligned}$$

where in (a) we used again Lemma 6.7.10 and in (b) Lemma 6.7.4.

We thus have from Eqs. (6.172), (6.173), (6.175),

$$\begin{aligned}
&\mathbb{E}_{\geq 3} \left\{ \left(\int_{T(\delta)} \sigma_1 e^{H(\boldsymbol{\sigma})} \mu_0(\mathsf{d}\boldsymbol{\sigma}) \right)^{2k} \right\} \\
&\leq C_k (1 + \|\boldsymbol{u}\|)^{Ck^2} \left(|u_1|^{2k} + N^{-k} + N^k e^{-N^c} + e^{-cN} \right) \big(\mathbb{E}_{\geq 3} Z\big(T(\delta)\big) \big)^{2k} .
\end{aligned}$$

The desired claim (6.129) follows from Markov Inequality upon adjusting the constant c.

Magnetization in the band: proof of Lemma 6.7.2

In the remaining of this section, we denote by μ the Gibbs measure associated to $H(\boldsymbol{\sigma})$, i.e.

$$\mu(\mathrm{d}\boldsymbol{\sigma}) \propto \exp(H(\boldsymbol{\sigma}))\,\mu_0(\mathrm{d}\boldsymbol{\sigma}).$$

In the following we estimate the components of $\langle\boldsymbol{\sigma}\rangle = (\langle\sigma_1\rangle, \dots, \langle\sigma_N\rangle)$ in the basis of eigenvectors of the quadratic part of the Hamiltonian $\boldsymbol{W}^{(2)}$. For simplicity of notation, we consider the component $\langle\sigma_1\rangle$ but we emphasize that this does not necessarily correspond to the largest (or smallest) eigenvalue of $\boldsymbol{W}^{(2)}$. Defining $\boldsymbol{\sigma}_{-1} = (\sigma_2, \dots, \sigma_N)$, we have

$$\int \sigma_1 e^{H(\boldsymbol{\sigma})}\mu_0(\mathrm{d}\boldsymbol{\sigma}) = \frac{1}{Z}\int \sigma_1 e^{\sigma_1 u_1 + \Lambda_1\sigma_1^2}\hat{E}(\sigma_1)\mathrm{d}\sigma_1\,. \tag{6.176}$$

where we defined

$$\hat{E}(\sigma_1) = C_N(1-\sigma_1^2/N)^{(N-3)/2}\int \exp\Big(\frac{\sigma_1}{N}\sum_{i,j=2}^{N}\widetilde{g}_{1ij}\sigma_i\sigma_j\Big)e^{H_{\sigma_1}(\boldsymbol{\sigma}_{-1})}\mu_{0,\sqrt{N-\sigma_1^2}}(\mathrm{d}\boldsymbol{\sigma}_{-1})\,,$$

$\mu_{0,\rho}$ denotes the uniform measure over the sphere of radius ρ,

$$C_N := \frac{\Gamma(N-1)}{\Gamma((N-1)/2)^2 2^{N-2}\sqrt{N}} = \frac{1}{\sqrt{2\pi}} + O(N^{-1})\,,$$

and

$$H_{\sigma_1}(\boldsymbol{\sigma}_{-1}) := \sum_{i=2}^{N}(\sigma_i u_i + \Lambda_i\sigma_i^2) + N^{-1}\sum_{i,j,k>1} g^{(3)}_{ijk}\sigma_i\sigma_j\sigma_k + \sum_{p\geq 4}H_p(\boldsymbol{\sigma}).$$

Here $\widetilde{g}_{1ij}$ is the sum of g over permutations of $(1,i,j)$. In particular $\widetilde{g}_{1ij} = \widetilde{g}_{1ji}$

$$(\widetilde{g}_{1ij})_{1<i<j} \sim_{iid} \mathcal{N}(0, 3\beta_3^2/2)\,,\quad (\widetilde{g}_{1ii})_{1<i} \sim_{iid} \mathcal{N}(0, 3\beta_3^2)\,. \tag{6.177}$$

We set $\hat{E}(\sigma_1) = 0$ for $|\sigma_1| > \sqrt{N}$.

By Lemma 6.7.3 and Lemma 6.7.8, with probability $1 - e^{-cN} - N^{-C}$,

$$Z = (1 + O(N^{-c}))\sqrt{\frac{2}{G''(\gamma_*)}}\cdot\exp\Big\{N\Big[\xi_{\geq 3}(1) - \frac{1}{2}\log(2e) + G(\gamma_*)\Big]\Big\}\,,$$

where $G(\gamma)$ and γ_* where defined in Eqs. (6.102) and (6.103).

In estimating $\langle\sigma_1\rangle$, we first anneal over $\boldsymbol{g}_{\geq 4}$ and $\boldsymbol{g}_{3-} := (g_{ijk} : 1 < i < j < k)$. We have

$$E(\sigma_1) := \mathbb{E}_{\boldsymbol{g}_{3-},\boldsymbol{g}_{\geq 4}}[\hat{E}(\sigma_1)] = C_N\Big(1 - \frac{\sigma_1^2}{N}\Big)^{(N-3)/2}\int \exp\left(\frac{\sigma_1}{N}\sum_{i,j=2}^{N}\widetilde{g}_{1ij}\sigma_i\sigma_j\right)$$
$$\exp\Big\{H_{\leq 2}(\boldsymbol{\sigma}_{-1}) + N\xi_{\geq 4}(1)/2 + N\beta_3^2(1-\sigma_1^2/N)^3/2\Big\}\mu_{0,\sqrt{N-\sigma_1^2}}(\mathrm{d}\boldsymbol{\sigma}_{-1}).$$

The next lemma show that this expectation is an accurate approximation of $\hat{E}(\sigma_1)$.

Lemma 6.7.12. *We have for an appropriate $c \in (0, 1/8)$ that, with probability $1 - N^{-c}$,*

$$\int \sigma_1 e^{u_1\sigma_1 + \Lambda_1\sigma_1^2}\hat{E}(\sigma_1)\,\mathrm{d}\sigma_1 =$$
$$= \int \sigma_1 e^{u_1\sigma_1 + \Lambda_1\sigma_1^2}E(\sigma_1)\mathrm{d}\sigma_1 + O\left(N^{-1/2+c}(|u_1| + N^{-1/2})\int e^{u_1\sigma_1 + \Lambda_1\sigma_1^2}E(\sigma_1)\mathrm{d}\sigma_1\right).$$

Before proving Lemma 6.7.12, we use it to prove Lemma 6.7.1.

Proof of Lemma 6.7.1. For $U(\sigma_1) := N\xi_{\geq 4}(1)/2 + N\beta_3^2(1-\sigma_1^2/N)^3/2$, we have

$$E(\sigma_1) = C_N\Big(1-\frac{\sigma_1^2}{N}\Big)^{(N-3)/2}\int \exp\left(\frac{\sigma_1}{N}\sum_{i,j=2}^N \widetilde{g}_{1ij}\sigma_i\sigma_j\right) e^{H_{\leq 2}(\boldsymbol{\sigma}_{-1})+U(\sigma_1)}\,\mu_{0,\sqrt{N-\sigma_1^2}}(\mathsf{d}\boldsymbol{\sigma}_{-1})\,.$$

Again by Lemma 6.7.3, for $\boldsymbol{V} = \boldsymbol{V}(\sigma_1) := \boldsymbol{\Lambda}_{-1} + \boldsymbol{\Delta}$, where $\boldsymbol{\Lambda}_{-1}$ is the diagonal matrix with entries corresponding to the spectrum of $\boldsymbol{W}^{(2)}$, with Λ_1 replaced by 0, and $\boldsymbol{\Delta} := \sigma_1 N^{-1}\tilde{\boldsymbol{G}}$ with $\tilde{G}_{ij} = \widetilde{g}_{1ij}$,

$$E(\sigma_1) = (1+O(N^{-1}))\frac{1}{(2e)^{(N-1)/2}\sqrt{2\pi}}(1-\sigma_1^2/N)^{-1}\sqrt{\frac{2}{G''_{\sigma_1}(\gamma_*(\sigma_1))}}\exp\big(U(\sigma_1)+NG_{\sigma_1}(\gamma_*(\sigma_1))\big), \quad (6.178)$$

where we defined

$$G_{\sigma_1}(\gamma) := (1-\sigma_1^2/N)\gamma - \frac{1}{2N}\log\det(\gamma\boldsymbol{I}_{N-1}-\boldsymbol{V}) + \frac{1}{4N}\langle\boldsymbol{u},(\gamma\boldsymbol{I}_{N-1}-\boldsymbol{V})^{-1}\boldsymbol{u}\rangle\,, \quad (6.179)$$

$$\gamma_*(\sigma_1) = \arg\max G_{\sigma_1}(\gamma)\,. \quad (6.180)$$

By Lemma 6.7.12, we have

$$\int \sigma_1\,\mu(\mathsf{d}\boldsymbol{\sigma}) = \frac{\int \sigma_1 e^{u_1\sigma_1+\Lambda_1\sigma_1^2}E(\sigma_1)\mathsf{d}\sigma_1}{\int e^{u_1\sigma_1+\Lambda_1\sigma_1^2}E(\sigma_1)\mathsf{d}\sigma_1} + O\Big(N^{-1/2+c}(|u_1|+N^{-1/2})\Big)\,. \quad (6.181)$$

We next estimate these integrals by approximating their argument for small σ_1. Note that by Lemma 6.7.5 and Lemma 6.7.7, we can restrict these integrals to $|\sigma_1| \leq C\log N$ making a negligible error.

It is easy to see that, for $\sigma_1 = 0$, we recover $G_{\sigma_1}(\gamma) = G_0(\gamma)$, where $G_0(\gamma)$ is the same function defined in Eq. (6.102), with N replaced by $N-1$. To leading order, we can expand

$$\begin{aligned}
&G_{\sigma_1}(\gamma) = \\
&= (1-\sigma_1^2/N)\gamma - \frac{1}{2N}\log\det(\gamma\boldsymbol{I}-\boldsymbol{V}) + \frac{1}{4N}\langle\boldsymbol{u},(\gamma\boldsymbol{I}-\boldsymbol{V})^{-1}\boldsymbol{u}\rangle \\
&= (1-\sigma_1^2/N)\gamma - \frac{1}{2N}\log\det(\gamma\boldsymbol{I}-\boldsymbol{V}) + \frac{1}{4N}\langle\boldsymbol{u},(\boldsymbol{I}+(\gamma\boldsymbol{I}-\boldsymbol{\Lambda}_{-1})^{-1}\boldsymbol{\Delta}+\boldsymbol{E}_N)(\gamma\boldsymbol{I}-\boldsymbol{\Lambda}_{-1})^{-1}\boldsymbol{u}\rangle.
\end{aligned}$$

where $\|\boldsymbol{E}_N\|_{\mathsf{op}} = O(N^{-1})$ with probability $1-\exp(-cN)$ over $\boldsymbol{W}^{(3)}$.Therefore

$$\begin{aligned}
G_{\sigma_1}(\gamma) - G_0(\gamma) = &-\frac{\gamma\sigma_1^2}{N} + \frac{1}{2N}\log(\gamma-\Lambda_1) - \frac{1}{2N}\log\det\big(\boldsymbol{I}-(\gamma\boldsymbol{I}-\boldsymbol{\Lambda}_{-1})^{-1/2}\boldsymbol{\Delta}(\gamma\boldsymbol{I}-\boldsymbol{\Lambda}_{-1})^{-1/2}\big) \\
&+\frac{1}{4N}\langle\boldsymbol{u},(\gamma\boldsymbol{I}-\boldsymbol{\Lambda}_{-1})^{-1}\boldsymbol{\Delta}(\gamma\boldsymbol{I}-\boldsymbol{\Lambda}_{-1})^{-1}\boldsymbol{u}\rangle + O(\|\boldsymbol{u}\|^2/N^2)\,. \quad (6.182)
\end{aligned}$$

on $\gamma > \max_i \Lambda_i + \varepsilon$. Since the above difference (and its derivative with respect to λ) is of order $\sigma_1/\sqrt{N}$ and G is strongly convex in a neighborhood of γ_*, it follows that $\gamma_*(\sigma_1) = \gamma_* + O(\sigma_1^2/N)$. We will therefore restrict ourselves to $|\gamma-\gamma_*| \leq CN^{-1}(\log N)^2$.

We next expand the log-determinant term in the difference. Defining

$$D_2 := \sum_{i,j=1}^N \big(N^{-1}(\gamma\boldsymbol{I}-\boldsymbol{\Lambda}_{-1})^{-1/2}\tilde{\boldsymbol{G}}(\gamma\boldsymbol{I}-\boldsymbol{\Lambda}_{-1})^{-1/2}\big)_{ij}^2, \quad (6.183)$$

we have

$$\mathsf{Tr}\Big((\gamma\boldsymbol{I}-\boldsymbol{\Lambda}_{-1})^{-1/2}\boldsymbol{\Delta}(\gamma\boldsymbol{I}-\boldsymbol{\Lambda}_{-1})^{-1/2}\Big) = \frac{\sigma_1}{N}\sum_{i\neq 1}(\gamma-\Lambda_i)^{-1}\widetilde{g}_{1ii} + O(N^{-1}), \quad (6.184)$$

$$\mathsf{Tr}\Big(\big(\gamma\boldsymbol{I}-\boldsymbol{\Lambda}_{-1})^{-1/2}\boldsymbol{\Delta}(\gamma\boldsymbol{I}-\boldsymbol{\Lambda}_{-1})^{-1/2}\big)^2\Big) = D_2\sigma_1^2\,, \quad (6.185)$$

$$\mathsf{Tr}\Big(\big((\gamma\boldsymbol{I}-\boldsymbol{\Lambda}_{-1})^{-1/2}\boldsymbol{\Delta}(\gamma\boldsymbol{I}-\boldsymbol{\Lambda}_{-1})^{-1/2}\big)^k\Big) = O(N^{-1}) \qquad \text{for } k\geq 3. \quad (6.186)$$

Thus, with high probability,

$$\frac{1}{2N}\log\det\Big(\boldsymbol{I}-(\gamma\boldsymbol{I}-\boldsymbol{\Lambda}_{-1})^{-1/2}\boldsymbol{\Delta}(\gamma\boldsymbol{I}-\boldsymbol{\Lambda}_{-1})^{-1/2}\Big)=-\frac{\sigma_1}{2N^2}\sum_{i\neq 1}(\gamma-\Lambda_i)^{-1}\widetilde{g}_{1ii}-\frac{D_2}{4N}\sigma_1^2+O(N^{-2})\,.$$

For $\gamma=\gamma_*(\sigma_1)=\gamma_*+O(\sigma_1^2/N)$, we can compute

$$\mathbb{E}D_2=\frac{3\beta_3^2}{2N^2}\Big(\sum_{i\neq 1}(\gamma_*-\Lambda_i)^{-1}\Big)^2+O(N^{-1})$$

and

$$\mathrm{Var}(D_2)=\frac{\beta_3^4}{N^4}\,O\Big((\sum_{i\neq 1}\Big(\gamma_*-\Lambda_i)^{-2}\Big)^2\Big)=O(N^{-2}).$$

Furthermore, recalling the stationarity condition $G'(\gamma_*)=0$, which yields

$$\frac{1}{2N}\sum_{i=1}^N\frac{1}{\gamma_*-\Lambda_i}=1+\frac{1}{4N}\sum_{i=1}^N\frac{u_i^2}{(\gamma_*-\Lambda_i)^2}$$

which yields (for $\|\boldsymbol{u}\|\le N^{c_0}$) $\sum_{i\ge 1}(\gamma_*-\Lambda_i)^{-1}=2N+O(N^{2c_0})$, and therefore

$$\mathbb{E}D_2=6\beta_3^2+O(N^{-1})\,. \tag{6.187}$$

Substituting the above estimates in Eq. (6.182) the following holds with probability at least $1-\exp(-N^c)$, for $|\sigma_1|\le C\log N$,

$$\min_\gamma G_{\sigma_1}(\gamma)-G_0(\gamma_*)=-\frac{\gamma_*\sigma_1^2}{N}+\frac{1}{2N}\log(\gamma_*-\Lambda_1)+\frac{\sigma_1}{2N^2}\sum_{i=1}^N(\gamma_*-\Lambda_i)^{-1}\widetilde{g}_{1ii}+\frac{D_2}{4N}\sigma_1^2+O(N^{-2+3c_0}).$$

Letting $a_N:=C\log N$, and using Eq. (6.178),

$$\begin{aligned}
&\int\sigma_1 e^{\sigma_1u_1+\Lambda_1\sigma_1^2}E(\sigma_1)\mathsf{d}\sigma_1\\
&=\frac{1}{(2e)^{(N-1)/2}\sqrt{2\pi}}\sqrt{\frac{2}{G''(\gamma_*)}}\int_{[-a_N,a_N]}(1-\sigma_1^2/N)^{-1}\sigma_1\\
&\qquad\exp\Big\{NG_0(\gamma_*)+\frac{1}{2}\log(\gamma_*-\Lambda_1)+U(\sigma_1)+\sigma_1\Big(u_1+\frac{1}{2N}\sum_i(\gamma-\Lambda_i)^{-1}\widetilde{g}_{1ii}\Big)\\
&\qquad\qquad-\Big(-\Lambda_1-\frac{1}{4}D_2+\gamma_*\Big)\sigma_1^2+O(N^{-1+3c_0})\Big\}\mathsf{d}\sigma_1+\delta_N\\
&\overset{(a)}{=}(2e)^{-(N-1)/2}\sqrt{\frac{1}{\pi G''(\gamma_*)}}\int_{[-a_N,a_N]}\sigma_1\exp\Big\{NG_0(\gamma_*)+\frac{1}{2}\log(\gamma_*-\Lambda_1)\\
&\qquad+U(\sigma_1)+\sigma_1\Big(u_1+\frac{1}{2N}\sum_i(\gamma-\Lambda_i)^{-1}\widetilde{g}_{1ii}\Big)-\Big(-\Lambda_1-\frac{3}{2}\beta_3^2+\gamma_*\Big)\sigma_1^2+O(N^{-1+3c_0})\Big\}\mathsf{d}\sigma_1+\delta_N\\
&=(1+O(N^{-1}))(2e)^{-(N-1)/2}\sqrt{\frac{1}{\pi G''(\gamma_*)}}\int_{[-a_N,a_N]}\sigma_1\exp\Big\{NG_0(\gamma_*)+\frac{1}{2}\log(\gamma_*-\Lambda_1)\\
&\qquad+\frac{N}{2}(\xi_{\ge 4}(1)+\beta_3^2)/2+\sigma_1\Big(u_1+\frac{1}{2N}\sum_i(\gamma-\Lambda_i)^{-1}\widetilde{g}_{1ii}\Big)-(-\Lambda_1+\gamma_*)\sigma_1^2+O(N^{-1+3c_0})\Big\}\mathsf{d}\sigma_1+\delta_N\,,
\end{aligned}$$

where in (a) we used Eq. (6.187), and

$$|\delta_N|\le N^{-1}\int e^{\sigma_1u_1+\Lambda_1\sigma_1^2}E(\sigma_1)\mathsf{d}\sigma_1\,. \tag{6.188}$$

Therefore, we obtain

$$\frac{\int \sigma_1 \exp(\sigma_1 u_1 + \Lambda_1 \sigma_1^2) E(\sigma_1) \mathsf{d}\sigma_1}{\int \exp(\sigma_1 u_1 + \Lambda_1 \sigma_1^2) E(\sigma_1) \mathsf{d}\sigma_1} = \frac{u_1 + N^{-1} \sum_i (\gamma_* - \Lambda_i)^{-1} \widetilde{g}_{1ii}}{2(\gamma_* - \Lambda_1)} + O(N^{-1}). \tag{6.189}$$

which completes the proof using Eq. (6.181). □

Finally, we prove Lemma 6.7.12. The main idea is that the error in annealing can be controlled by accurate estimates of certain quantities involving overlap over the quadratic model on $\boldsymbol{\sigma}_{-1}$, which follows from Laplace transform and expansion of the dependence on σ_1.

Proof of Lemma 6.7.12. Define

$$\begin{aligned} W(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2) &:= \frac{1}{N} \mathbb{E}\Big\{ \big(H_3(\boldsymbol{\sigma}^1_{-1}) + H_{\geq 4}(\boldsymbol{\sigma}^1)\big)\big(H_3(\boldsymbol{\sigma}^2_{-1}) + H_{\geq 4}(\boldsymbol{\sigma}^2)\big) \Big\} \\ &= \beta_3^2 \langle \boldsymbol{\sigma}^1_{-1}, \boldsymbol{\sigma}^2_{-1} \rangle_N^3 + \xi_{\geq 4}\big(\langle \boldsymbol{\sigma}^1_{-1}, \boldsymbol{\sigma}^2_{-1} \rangle_N + \sigma_1^1 \sigma_1^2 / N\big) \,, \end{aligned}$$

where, with an abuse of notation, $H_3(\boldsymbol{\sigma}^a_{-1}) := N^{-1} \sum_{i,j,k=2}^N \sigma_i^a \sigma_j^a \sigma_k^a$ (and a similar notation will be used for $H_{\leq 2}(\boldsymbol{\sigma}^a_{-1})$ below). Note that $W(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^1) = \xi_{\geq 4}(1) + \beta_3^2 (1 - (\sigma_1^1)^2/N)^3$. For a Borel set $U \subseteq S_N^2$, define

$$\begin{aligned} Q(U) := \int_U & \sigma_1^1 \sigma_1^2 \, e^{u_1(\sigma_1^1 + \sigma_1^2) + \Lambda_1((\sigma_1^1)^2 + (\sigma_1^2)^2)} \cdot \\ & \cdot \exp\left\{ N^{-1}\Big(\sigma_1^1 \sum_{i,j=2}^N \widetilde{g}_{1ij} \sigma_i^1 \sigma_j^1 + \sigma_1^2 \sum_{i,j=2}^N \widetilde{g}_{1ij} \sigma_i^2 \sigma_j^2 \Big) + H_{\leq 2}(\boldsymbol{\sigma}^1_{-1}) + H_{\leq 2}(\boldsymbol{\sigma}^2_{-1}) \right\} \\ & e^{N[W(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^1) + W(\boldsymbol{\sigma}^2, \boldsymbol{\sigma}^2)]/2} \big\{ \exp[N W(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2)] - 1 \big\} \mu_0^{\otimes 2}(\mathsf{d}\boldsymbol{\sigma}) \,. \end{aligned}$$

Expanding the square and taking expectation, we obtain

$$\mathbb{E}_{\boldsymbol{g}_{\geq 4}, \boldsymbol{g}_{3-}} \left[\left\{ \int_{T(\delta)} \sigma_1 \left(e^{H(\boldsymbol{\sigma})} - \mathbb{E}_{\boldsymbol{g}_{\geq 4}, \boldsymbol{g}_{3-}} e^{H(\boldsymbol{\sigma})} \right) \mu_0(\mathsf{d}\boldsymbol{\sigma}) \right\}^2 \right] = Q\big(T(\delta) \times T(\delta)\big) \,.$$

Further, writing $T = T(\delta)$, and $A_2 = A_2(N^{-1/2+c})$, we obtain that, with probability at least $1 - \exp(-N^c)$,

$$\begin{aligned} \big| Q(A_2) - Q(T \times T) \big| &= N \cdot Q\Big(A_2 \setminus T \times T\Big) + N \cdot Q\Big(T \times T \setminus A_2\Big) \\ &\overset{(a)}{\leq} N \cdot Z_{\leq 2,2}\Big(A_2 \setminus T \times T\Big) e^{N \xi_{\geq 3}(1)} + N \cdot Z_{\leq 2,2}\Big(T \times T \setminus A_2\Big) e^{N \xi_{\geq 3}(1)} \\ &\overset{(b)}{\leq} e^{-cN} (Z_{\leq 2})^2 e^{N \xi_{\geq 3}(1)} + e^{-N^c} (Z_{\leq 2})^2 e^{N \xi_{\geq 3}(1)} \end{aligned}$$

where in (a) we used the fact that $|\sigma_1^1 \sigma_1^2| \leq N$, and in (b) the first term was bounded by using $Z_{\leq 2,2}((T \times T)^c) \leq 2 Z_{\leq 2}(T^c) Z_{\leq 2}$ and applying Lemma 6.7.7, see Eq. (6.124), and the second by $Z_{\leq 2,2}(T \times T \setminus A_2) \leq Z_{\leq 2,2}(A_2^c)$ and using Lemma 6.7.10, Eq. (6.132). Hence we conclude that

$$\begin{aligned} & \mathbb{E}_{\boldsymbol{g}_{\geq 4}, \boldsymbol{g}_{3-}} \left[\left\{ \int_{T(\delta)} \sigma_1 \left(e^{H(\boldsymbol{\sigma})} - \mathbb{E}_{\boldsymbol{g}_{\geq 4}, \boldsymbol{g}_{3-}} e^{H(\boldsymbol{\sigma})} \right) \mu_0(\mathsf{d}\boldsymbol{\sigma}) \right\}^2 \right] && (6.190) \\ & \qquad = Q(A_2(N^{-1/2+c})) + O\left(e^{-N^c + N \xi_{\geq 1}(1)} (Z_{\leq 2})^2 \right) . && (6.191) \end{aligned}$$

By Taylor expansion, always using the shorthand $A_2 = A_2(N^{-1/2+c})$,

$$\begin{aligned} Q(A_2) = \int_{A_2} & \sigma_1^1 \sigma_1^2 \, \exp\Big\{ \sum_{i=1}^2 \Big((u_1 \sigma_1^i + \Lambda_1 (\sigma_1^i)^2) + H_{\leq 2}(\boldsymbol{\sigma}^i_{-1}) + N W(\boldsymbol{\sigma}^i, \boldsymbol{\sigma}^i) \Big) \Big\} \\ & \exp\Big\{ \frac{\sigma_1^1}{N} \sum_{i,j=2}^N \widetilde{g}_{1ij} \sigma_i^1 \sigma_j^1 + \frac{\sigma_1^2}{N} \sum_{i,j=2}^N \widetilde{g}_{1ij} \sigma_i^2 \sigma_j^2 \Big\} \cdot \Big\{ \sum_{\ell=1}^L \frac{1}{\ell!} (N W(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2))^\ell + O(N^{-L/2+c}) \Big\} \mu_0^{\otimes 2}(\mathsf{d}\boldsymbol{\sigma}) \,. \end{aligned}$$

We estimate each term

$$T_\ell(a,b) := \int_{A_2} \sigma_1^1\sigma_1^2 \exp\Big\{\sum_{i=1}^2 \Big(\big(u_1\sigma_1^i + \Lambda_1(\sigma_1^i)^2\big) + H_{\le 2}(\boldsymbol{\sigma}_{-1}^i) + NW(\boldsymbol{\sigma}^i,\boldsymbol{\sigma}^i)\Big)\Big\} \tag{6.192}$$

$$\exp\Big\{\frac{\sigma_1^1}{N}\sum_{1<i<j}\widetilde{g}_{1ij}\sigma_i^1\sigma_j^1 + \frac{\sigma_1^2}{N}\sum_{i,j=2}^N \widetilde{g}_{1ij}\sigma_i^2\sigma_j^2\Big\}\cdot N^\ell\langle\boldsymbol{\sigma}_{-1}^1,\boldsymbol{\sigma}_{-1}^2\rangle_N^a(\sigma_1^1\sigma_1^2/N)^b\,\mu_0^{\otimes 2}(\mathsf{d}\boldsymbol{\sigma})\,.$$

We can restrict ourselves to terms with $a \ge 3\ell$ and $b = 0$, or $a + b \ge 3\ell + 1$, since these are the terms that can arise in $Q(A_2)$. Let

$$\widehat{T}_\ell(a,b) := \int_{S_N^2} \sigma_1^1\sigma_1^2 \exp\Big\{\sum_{i=1}^2 \Big(\big(u_1\sigma_1^i + \Lambda_1(\sigma_1^i)^2\big) + H_{\le 2}(\boldsymbol{\sigma}_{-1}^i) + NW(\boldsymbol{\sigma}^i,\boldsymbol{\sigma}^i)\Big)\Big\}$$

$$\exp\Big\{\frac{\sigma_1^1}{N}\sum_{i,j=2}^N\widetilde{g}_{1ij}\sigma_i^1\sigma_j^1 + \frac{\sigma_1^2}{N}\sum_{i,j=2}^N \widetilde{g}_{1ij}\sigma_i^2\sigma_j^2\Big\}\cdot N^\ell\langle\boldsymbol{\sigma}_{-1}^1,\boldsymbol{\sigma}_{-1}^2\rangle_N^a(\sigma_1^1\sigma_1^2/N)^b\,\mu_0^{\otimes 2}(\mathsf{d}\boldsymbol{\sigma})\,.$$

By Lemma 6.7.10, Eq. (6.132), we have

$$|T_\ell(a,b) - \widehat{T}_\ell(a,b)| \le e^{-N^c}.$$

Applying Lemma 6.7.6, we have, for appropriate $C_{i,j} = O(\|\boldsymbol{u}\|^{2a} + N^{\lfloor a/2\rfloor})$,

$$|\widehat{T}_\ell(a,b)| \le N^{\ell-b-a}\int_{S_N^2}(\sigma_1^1\sigma_1^2)^{b+1}\exp\Big\{\sum_{i=1}^2 \Big(\big(u_1\sigma_1^i + \Lambda_1(\sigma_1^i)^2\big) + H_{\le 2}(\boldsymbol{\sigma}_{-1}^i) + NW(\boldsymbol{\sigma}^i,\boldsymbol{\sigma}^i)\Big)\Big\}$$

$$\left\{C_{0,0} + \sum_{i,j=0,(i,j)\neq(0,0)}^L C_{i,j}N^{-(i+j)/2}(\sigma_1^1)^i(\sigma_1^2)^j + O_L(N^{-L/2})\right\}\mu_0^{\otimes 2}(\mathsf{d}\boldsymbol{\sigma}).$$

Note that when $b+1+i$ or $b+1+j$ is odd,

$$\frac{\int_{S_N^2}(\sigma_1^1\sigma_1^2)^{b+1}\exp\Big\{\sum_{i=1}^2\Big(\big(u_1\sigma_1^i+\Lambda_1(\sigma_1^i)^2\big)+H_{\le 2}(\boldsymbol{\sigma}_{-1}^i)+W(\boldsymbol{\sigma}^i,\boldsymbol{\sigma}^i)\Big)\Big\}C_{i,j}(\sigma_1^1)^i(\sigma_1^2)^j\,\mu_0^{\otimes 2}(\mathsf{d}\boldsymbol{\sigma})}{\int_{S_N^2}\exp\Big\{\sum_{i=1}^2\Big(\big(u_1\sigma_1^i+\Lambda_1(\sigma_1^i)^2\big)+H_{\le 2}(\boldsymbol{\sigma}_{-1}^i)+W(\boldsymbol{\sigma}^i,\boldsymbol{\sigma}^i)\Big)\Big\}\mu_0^{\otimes 2}(\mathsf{d}\boldsymbol{\sigma})}$$
$$= O_{b+i+j}\Big(|u_1|(1+|u_1|)^{2(b+1)+i+j}(\|\boldsymbol{u}\|^{2a}+N^{\lfloor a/2\rfloor})\Big)\,.$$

When both of them are odd,

$$\frac{\int_{S_N^2}(\sigma_1^1\sigma_1^2)^{b+1}\exp\Big\{\sum_{i=1}^2\Big(\big(u_1\sigma_1^i+\Lambda_1(\sigma_1^i)^2\big)+H_{\le 2}(\boldsymbol{\sigma}_{-1}^i)+W(\boldsymbol{\sigma}^i,\boldsymbol{\sigma}^i)\Big)\Big\}C_{i,j}(\sigma_1^1)^i(\sigma_1^2)^j\,\mu_0^{\otimes 2}(\mathsf{d}\boldsymbol{\sigma})}{\int_{S_N^2}\exp\Big\{\sum_{i=1}^2\Big(\big(u_1\sigma_1^i+\Lambda_1(\sigma_1^i)^2\big)+H_{\le 2}(\boldsymbol{\sigma}_{-1}^i)+W(\boldsymbol{\sigma}^i,\boldsymbol{\sigma}^i)\Big)\Big\}\mu_0^{\otimes 2}(\mathsf{d}\boldsymbol{\sigma})}$$
$$= O_{b+i+j}\Big(|u_1|^2(1+|u_1|)^{2(b+1)+i+j}(\|\boldsymbol{u}\|^{2a}+N^{\lfloor a/2\rfloor})\Big)\,.$$

Otherwise, when $b+1+i$ and $b+1+j$ are both even,

$$\frac{\Big|\int_{S_N^2}(\sigma_1^1\sigma_1^2)^{b+1}\exp\Big\{\sum_{i=1}^2\Big(\big(u_1\sigma_1^i+\Lambda_1(\sigma_1^i)^2\big)+H_{\le 2}(\boldsymbol{\sigma}_{-1}^i)+W(\boldsymbol{\sigma}^i,\boldsymbol{\sigma}^i)\Big)\Big\}C_{i,j}(\sigma_1^1)^i(\sigma_1^2)^j\,\mu_0^{\otimes 2}(\mathsf{d}\boldsymbol{\sigma})\Big|}{\int_{S_N^2}\exp\Big\{\sum_{i=1}^2\Big(\big(u_1\sigma_1^i+\Lambda_1(\sigma_1^i)^2\big)+H_{\le 2}(\boldsymbol{\sigma}_{-1}^i)+W(\boldsymbol{\sigma}^i,\boldsymbol{\sigma}^i)\Big)\Big\}\mu_0^{\otimes 2}(\mathsf{d}\boldsymbol{\sigma})}$$
$$\le O_{b+i+j}\Big((1+|u_1|)^{2(b+1)+i+j}(\|\boldsymbol{u}\|^{2a}+N^{\lfloor a/2\rfloor})\Big)\,.$$

Therefore, under the assumption $\|\boldsymbol{u}\| \le N^{c_0}$, for $\ell \le L$,

$$\frac{|\widehat{T}_\ell(a,b)|}{\int_{S_N^2} \exp\left\{\sum_{i=1}^2 \left((u_1\sigma_1^i + \Lambda_1(\sigma_1^i)^2) + H_{\le 2}(\boldsymbol{\sigma}_{-1}^i) + W(\boldsymbol{\sigma}^i,\boldsymbol{\sigma}^i)\right)\right\}\mu_0^{\otimes 2}(\mathsf{d}\boldsymbol{\sigma})}$$
$$\le (N^{\lfloor a/2\rfloor} + \|\boldsymbol{u}\|^{2a})\cdot\Bigg[O_L\Bigg(|u_1|^2\sum_{i,j\le L} N^{\ell-b-a-(i+j)/2}\Bigg) + \sum_{\substack{i,j\le L\\ i,j=b+1 \mod 2}} O_L\left(N^{\ell-b-a-(i+j)/2}\right)$$
$$+ \sum_{\substack{i,j\le L\\ i\ne j \mod 2}} O_L\left(|u_1|N^{\ell-b-a-(i+j)/2}\right)\Bigg]$$
$$= O_L(N^{-2} + N^{-3/2}|u_1| + N^{-1}|u_1|^2) = O_L(N^{-1}|u_1|^2 + N^{-2}),$$

where in the last step we used the fact that $\ell \ge 1$, and $a \ge 3\ell$ when $b = 0$, or $a + b \ge 3\ell + 1$, otherwise.

Take $L = 4$, and combining the terms in Eq. (6.192), we obtain

$$Q(A_2(N^{-1/2+c})) \le O\left((N^{-2} + N^{-1}|u_1|)(Z_{\le 2})^2 e^{N\xi_{\ge 3}(1)}\right),$$

and therefore, using Eq. (6.191)

$$\mathbb{E}_{\boldsymbol{g}_{\ge 4},\boldsymbol{g}_{3-}}\left[\left\{\int_{T(\delta)}\sigma_1\left(e^{H(\boldsymbol{\sigma})} - \mathbb{E}_{\boldsymbol{g}_{\ge 4},\boldsymbol{g}_{3-}} e^{H(\boldsymbol{\sigma})}\right)\mu_0(\mathsf{d}\boldsymbol{\sigma})\right\}^2\right] = O\left((N^{-2} + N^{-1}|u_1|^2)(Z_{\le 2})^2 e^{N\xi_{\ge 3}(1)}\right).$$

Thus, with probability at least $1 - N^{-c}$, we have

$$\left|\int_{T(\delta)}\sigma_1\left(e^{H(\boldsymbol{\sigma})} - \mathbb{E}_{\boldsymbol{g}_{\ge 4},\boldsymbol{g}_{3-}} e^{H(\boldsymbol{\sigma})}\right)\mu_0(\mathsf{d}\boldsymbol{\sigma})\right| \le Z_{\le 2}N^c(N^{-1} + N^{-1/2}|u_1|).$$

This yields the desired claim upon using Lemma 6.7.7. □

We note that (6.129) in Lemma 6.7.8 immediately gives the following high probability bound on the magnetization.

Lemma 6.7.13. *For any $\varepsilon, C > 0$, there exists $c_0 > 0$ such that, for $\|\boldsymbol{u}\| \le N^{c_0}$, with probability at least $1 - N^{-C}$, we have*

$$\left\|\int \boldsymbol{\sigma}\mu(\mathsf{d}\boldsymbol{\sigma})\right\|^2 \le N^{\varepsilon}, \tag{6.193}$$

for N sufficiently large.

Proof. We work, as before, in the basis of eigenvectors of the quadratic part $\boldsymbol{W}_2$ of the Hamiltonian. By (6.129), with $k = 4C/\varepsilon$, with probability at least $1 - N^{-2C}$,

$$\int_T \sigma_i e^{H(\boldsymbol{\sigma})}\mu_0(\mathsf{d}\boldsymbol{\sigma}) \le N^{\varepsilon/4}\|\boldsymbol{u}\|^{Ck}(|u_i| + CN^{-1/2})\mathbb{E}_{\ge 3}\int e^{H(\boldsymbol{\sigma})}\mu_0(\mathsf{d}\boldsymbol{\sigma}).$$

By (6.127) with $L = 4C$ and the union bound over $i \in [N]$, we then have, with probability at least $1 - \varepsilon^{-8C}N^{-C}$, for all $i \in [N]$,

$$\frac{1}{Z}\int \sigma_i e^{H(\boldsymbol{\sigma})}\mu_0(\mathsf{d}\boldsymbol{\sigma}) \le N^{\varepsilon/2}\|\boldsymbol{u}\|^{Ck}(|u_i| + CN^{-1/2}).$$

Assuming that c_0 is chosen so that $c_0 L < \varepsilon/4$, we then obtain (6.193). □

Lemma 6.7.2 now follows.

Proof of Lemma 6.7.2. Let $\hat{\boldsymbol{m}} = \boldsymbol{m} + \boldsymbol{\Delta}(\boldsymbol{m})$. From Lemma 6.7.1 we have, with probability at least $1 - N^{-c}$

$$\|\langle \boldsymbol{\sigma} \rangle - \hat{\boldsymbol{m}}\|^2 \le O\big(N^{-c} + N^{-c}\|\boldsymbol{u}\|^2\big)\,.$$

Therefore, using Lemma 6.7.13 and the trivial bound $\|\langle \boldsymbol{\sigma} \rangle\| \le \sqrt{N}$, we can pick $\varepsilon > 0$ sufficiently small and k sufficiently large such that, upon adjusting the constant c,

$$\begin{aligned}\mathbb{E}[\|\langle \boldsymbol{\sigma} \rangle - \hat{\boldsymbol{m}}\|^\alpha] &= O(N^{-c\alpha} + N^{-c\alpha}\|\boldsymbol{u}\|^\alpha) + N^{-C} + O(N^{\varepsilon\alpha} \cdot N^{-c}) \\ &= O(N^{-c/2}).\end{aligned}$$

□

6.7.2 Integrating over bands

Using the results in the previous section, we will complete the proof of Proposition 6.4.6. We will assume the setup of Proposition 6.4.6. We sample $\boldsymbol{x} \sim \mu_{\mathsf{unif}}$, $\boldsymbol{y} = t\boldsymbol{x} + \boldsymbol{B}_t$, and $\widetilde{H}(\,\cdot\,) \sim \mu_{\mathsf{null}}$ (the Gaussian process with covariance $\mathbb{E}\,\widetilde{H}(\boldsymbol{\sigma}^1)\widetilde{H}(\boldsymbol{\sigma}^2) = N\xi(\langle \boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2 \rangle))$ with $\boldsymbol{x}, \boldsymbol{B}, \widetilde{H}$ independent. We define the tilted disorder $H(\boldsymbol{\sigma}) = \widetilde{H}(\boldsymbol{\sigma}) + \langle \boldsymbol{y}, \boldsymbol{\sigma} \rangle + N\xi(\langle \boldsymbol{x}, \boldsymbol{\sigma} \rangle_N)$, so that $(\boldsymbol{x}, H, \boldsymbol{y}) \sim \mathbb{P}$ are distributed according to the planted model, cf. Eq. (6.3.2). (For simplicity of notation, we drop the dependence on t in the notation of $H, \boldsymbol{y}$ in this section compared to the notation in Section 6.4.) In this section, we will estimate the mean of the Gibbs measure given by H.

Recall that

$$\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m}) = N\xi(\langle \boldsymbol{x}, \boldsymbol{m} \rangle_N) + \widetilde{H}(\boldsymbol{m}) + \langle \boldsymbol{y}, \boldsymbol{m} \rangle + \frac{N}{2}\theta(\|\boldsymbol{m}\|_N^2) + \frac{N}{2}\log(1 - \|\boldsymbol{m}\|_N^2),$$

where $\theta(s) = \xi(1) - \xi(s) - (1-s)\xi'(s)$.

Let $\boldsymbol{m} \in \mathbb{R}^N$ and $q = \|\boldsymbol{m}\|_N^2$. The following lemma follows from standard calculations.

Lemma 6.7.14. *The distribution of $\widetilde{H}(\boldsymbol{\sigma})$ given $\nabla \mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m}) = 0$ is a Gaussian process with*

$$\begin{aligned}&N^{-1}\mathbb{E}[\widetilde{H}(\boldsymbol{\sigma}) \mid \nabla \mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m}) = 0, \boldsymbol{y}, \boldsymbol{x}] \\ &= \frac{\xi'(\langle \boldsymbol{m}, \boldsymbol{\sigma} \rangle_N)\langle \boldsymbol{z}, \boldsymbol{\sigma} \rangle_N}{\xi'(q)} - \frac{\xi''(q)\langle \boldsymbol{m}, \boldsymbol{z} \rangle_N}{\xi'(q)\zeta(q)}\xi'(\langle \boldsymbol{m}, \boldsymbol{\sigma} \rangle_N)\langle \boldsymbol{m}, \boldsymbol{\sigma} \rangle_N,\end{aligned} \tag{6.194}$$

with $\zeta(q) = \xi'(q) + q\xi''(q)$ and $\boldsymbol{z} = -\boldsymbol{y} - \xi'(\langle \boldsymbol{x}, \boldsymbol{m} \rangle_N)\boldsymbol{x} + (1-q)\xi''(q)\boldsymbol{m} + \frac{\boldsymbol{m}}{1-q}$, and covariance

$$\begin{aligned}&N^{-1}\mathsf{Cov}[\widetilde{H}(\boldsymbol{\sigma}^1), \widetilde{H}(\boldsymbol{\sigma}^2) \mid \nabla \mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m}) = 0, \boldsymbol{y}, \boldsymbol{x}] \\ &= \xi(\langle \boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2 \rangle_N) - \frac{\xi'(\langle \boldsymbol{m}, \boldsymbol{\sigma}^1 \rangle_N)\xi'(\langle \boldsymbol{m}, \boldsymbol{\sigma}^2 \rangle_N)}{\xi'(q)}\langle \boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2 \rangle_N \\ &\qquad + \frac{\xi''(q)\xi'(\langle \boldsymbol{m}, \boldsymbol{\sigma}^1 \rangle_N)\xi'(\langle \boldsymbol{m}, \boldsymbol{\sigma}^2 \rangle_N)\langle \boldsymbol{m}, \boldsymbol{\sigma}^1 \rangle_N\langle \boldsymbol{m}, \boldsymbol{\sigma}^2 \rangle_N}{\xi'(q)\zeta(q)}.\end{aligned} \tag{6.195}$$

Let $\boldsymbol{\sigma}^\perp = \mathsf{proj}_{\{\boldsymbol{x},\boldsymbol{m}\}^\perp}(\boldsymbol{\sigma})$ be the projection of $\boldsymbol{\sigma}$ on $\{\boldsymbol{x}, \boldsymbol{m}\}^\perp$, and similarly define $\boldsymbol{y}^\perp$, $\boldsymbol{z}^\perp$. Define the band

$$D_N(a,b) := \Big\{\boldsymbol{\sigma} \in S_N : \ \ \langle \boldsymbol{\sigma}, \boldsymbol{m} \rangle_N = aq \text{ and } \langle \boldsymbol{\sigma}, \boldsymbol{x} \rangle_N = b\Big\}, \tag{6.196}$$

and let $r(a,b) = \|\boldsymbol{\sigma} - \boldsymbol{\sigma}^\perp\|_N^2$ for $\boldsymbol{\sigma} \in D_N(a,b)$.

Throughout the rest of the section, we will condition on the event $\nabla \mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m}) = 0$, and on $\boldsymbol{y} - \boldsymbol{y}^\perp$ and $\boldsymbol{x}$. Conditional on $\nabla \mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m}) = 0, \boldsymbol{y} - \boldsymbol{y}^\perp, \boldsymbol{x}$, we can write

$$N^{-1}H(\boldsymbol{\sigma}) = \xi(b) + \frac{\xi'(aq)\langle \boldsymbol{z}, \boldsymbol{\sigma} \rangle_N}{\xi'(q)} - \frac{\xi''(q)\xi'(aq)aq\langle \boldsymbol{m}, \boldsymbol{z} \rangle_N}{\xi'(q)\zeta(q)} + N^{-1}\widehat{H}(\boldsymbol{\sigma}^\perp) + \langle \boldsymbol{y}^\perp, \boldsymbol{\sigma}^\perp \rangle_N + \langle \boldsymbol{y} - \boldsymbol{y}^\perp, \boldsymbol{\sigma} - \boldsymbol{\sigma}^\perp \rangle_N,$$

where $\widehat{H}$ is a centered Gaussian process with covariance

$$\begin{aligned}&N^{-1}\mathrm{Cov}(\widehat{H}(\boldsymbol{\sigma}^{\perp,1}), \widehat{H}(\boldsymbol{\sigma}^{\perp,2})) \\ &= \xi\left(r(a,b) + \langle \boldsymbol{\sigma}^{\perp,1}, \boldsymbol{\sigma}^{\perp,2} \rangle_N\right) - \frac{\xi'(aq)^2}{\xi'(q)}\langle \boldsymbol{\sigma}^{\perp,1}, \boldsymbol{\sigma}^{\perp,2} \rangle_N - \frac{\xi'(aq)^2 r(a,b)}{\xi'(q)} + \frac{\xi'(aq)^2\xi''(q)(aq)^2}{\zeta(q)\xi'(q)}.\end{aligned}$$

Let $\tilde{\boldsymbol{\sigma}} = \boldsymbol{\sigma}^\perp / \|\boldsymbol{\sigma}^\perp\|_N$. We can then write

$$
\begin{aligned}
&\int_{D_N(a,b)} e^{H(\boldsymbol{\sigma})} \mu_0^{a,b}(\mathrm{d}\boldsymbol{\sigma}) \\
&= \exp\left(N \left[\xi(b) + \frac{\xi'(aq)\langle \boldsymbol{z}+\boldsymbol{y}, \boldsymbol{\sigma}\rangle_N}{\xi'(q)} - \frac{\xi''(q)\xi'(aq) aq \langle \boldsymbol{m}, \boldsymbol{z}\rangle_N}{\xi'(q)\zeta(q)} + \left(1 - \frac{\xi'(aq)}{\xi'(q)}\right) \langle \boldsymbol{y} - \boldsymbol{y}^\perp, \boldsymbol{\sigma} - \boldsymbol{\sigma}^\perp \rangle_N \right] \right) \\
&\qquad \int_{S_{N-2}} \exp\left(N \left(1 - \frac{\xi'(aq)}{\xi'(q)}\right) (1 - r(a,b))^{1/2} \langle \boldsymbol{y}^\perp, \tilde{\boldsymbol{\sigma}} \rangle_N + \underline{\tilde{H}}(\tilde{\boldsymbol{\sigma}}) + \frac{N-3}{2} \log(1 - r(a,b)) \right) \mu_0(\mathrm{d}\tilde{\boldsymbol{\sigma}}) \\
&= \exp\left(N \Gamma_N(\boldsymbol{y}, \boldsymbol{m}; a, b) + \frac{N-3}{2} \log(1 - r(a,b)) \right) \int_{S_{N-2}} e^{N^{1/2} g_{a,b} + \underline{H}(\tilde{\boldsymbol{\sigma}})} \mu_0(\mathrm{d}\tilde{\boldsymbol{\sigma}}),
\end{aligned} \tag{6.197}
$$

where $\mu_0^{a,b}$ is the measure induced on $D_N(a,b)$ by μ_0, we defined Γ_N via

$$
\Gamma_N(\boldsymbol{y}, \boldsymbol{m}; a, b) := \xi(b) + \frac{\xi'(aq)\langle \boldsymbol{z}+\boldsymbol{y}, \boldsymbol{\sigma}\rangle_N}{\xi'(q)} - \frac{\xi''(q)\xi'(aq) aq \langle \boldsymbol{m}, \boldsymbol{z}\rangle_N}{\xi'(q)\zeta(q)} + \left(1 - \frac{\xi'(aq)}{\xi'(q)}\right) \langle \boldsymbol{y} - \boldsymbol{y}^\perp, \boldsymbol{\sigma} - \boldsymbol{\sigma}^\perp \rangle_N \,, \tag{6.198}
$$

and $\underline{H}$ is a Hamiltonian on $\tilde{\boldsymbol{\sigma}}$ with mixture $\tilde{\xi}(q) = \sum_{k\ge 1} \tilde{\xi}_k$ given by

$$
\tilde{\xi}_1 = (1 - r(a,b)) \left(\xi'(r(a,b)) - \frac{\xi'(aq)^2}{\xi'(q)} + \left(1 - \frac{\xi'(aq)}{\xi'(q)}\right)^2 t \right) =: \tilde{\gamma}_1^2, \tag{6.199}
$$

$$
\tilde{\xi}_2 = \frac{1}{2} \xi''(r(a,b))(1 - r(a,b))^2 =: \tilde{\gamma}_2^2, \tag{6.200}
$$

$$
\tilde{\xi}_p = \frac{1}{p!} \xi^{(p)}(r(a,b))(1 - r(a,b))^p, \quad p \ge 3\,. \tag{6.201}
$$

Finally, $g_{a,b}$ is a Gaussian independent of $\underline{H}$ with standard deviation $\tilde{\gamma}_0$ given by

$$
\tilde{\gamma}_0^2 := \xi(r(a,b)) - \frac{\xi'(aq)^2 r(a,b)}{\xi'(q)} + \frac{\xi'(aq)^2 \xi''(q)(aq)^2}{\zeta(q)\xi'(q)}\,. \tag{6.202}
$$

Note that

$$
\begin{aligned}
\tilde{\xi}_{\ge 2}(s) &= \sum_{p\ge 2} \frac{1}{p!} \xi^{(p)}(r(a,b))(1 - r(a,b))^p s^p \\
&= \xi(r(a,b) + (1 - r(a,b))s) - \xi(r(a,b)) - \xi'(r(a,b))(1 - r(a,b))s
\end{aligned}
$$

and therefore

$$
\begin{aligned}
\tilde{\xi}''_{\ge 2}(s) &= (1 - r(a,b))^2 \xi''(r(a,b) + (1 - r(a,b))s) \\
&\overset{(6.5)}{<} \frac{(1 - r(a,b))^2}{(1 - (r(a,b) + (1 - r(a,b))s))^2} = \frac{1}{(1-s)^2}.
\end{aligned} \tag{6.203}
$$

Integrating twice shows $\tilde{\xi}_{\ge 2}$ satisfies condition (6.101), and thus the results in Subsection 6.7.1 apply to $\tilde{\xi}_{\ge 2}$. Similarly, note that

$$
\tilde{\xi}_{\ge 3}(1) = \xi(1) - \xi(r(a,b)) - \xi'(r(a,b))(1 - r(a,b)) - \frac{1}{2}\xi''(r(a,b))(1 - r(a,b))^2. \tag{6.204}
$$

Following Subsection 6.7.1, we write the quadratic component of $\underline{H}$ as $\langle \boldsymbol{A}^{(2)}, \tilde{\boldsymbol{\sigma}}^{\otimes 2} \rangle$ for $\boldsymbol{A}^{(2)} = \boldsymbol{A}^{(2)}(a,b)$ a GOE matrix scaled by $\tilde{\gamma}_2/\sqrt{2}$. Recall the definition of $G(\gamma) = G(\gamma; \boldsymbol{A}, \boldsymbol{u})$ in Eq. (6.102). We take $\boldsymbol{u}$ to be the external field $\boldsymbol{u} = \tilde{\gamma}_1 \boldsymbol{g}$, and $\boldsymbol{A} = \boldsymbol{A}^{(2)}$. Note that $\boldsymbol{u}$ and $\boldsymbol{A}^{(2)}$ depend on the parameters a, b. Let

$\gamma_{a,b} = \arg\min_{z>z_*} G(z; \boldsymbol{A}^{(2)}, \boldsymbol{u})$, $z_* := \lambda_{\max}(\boldsymbol{A}^{(2)})$. From Lemma 6.7.8 Eqs. (6.127) and (6.128) and Lemma 6.7.3, when

$$\tilde{\gamma}_1^2 = (1-r(a,b))\left(\xi'(r(a,b)) - \frac{\xi'(aq)^2}{\xi'(q)} + \left(1-\frac{\xi'(aq)}{\xi'(q)}\right)^2 t\right) \le N^{c_0-1},$$

we have (with probability at least $1-N^{-c}$, conditional on $\nabla\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m}) = 0, \boldsymbol{y}-\boldsymbol{y}^\perp, \boldsymbol{x}$) that

$$\begin{aligned}
&\int_{D_N(a,b)} e^{H(\boldsymbol{\sigma})}\mu_0^{a,b}(\mathrm{d}\boldsymbol{\sigma}) \\
&= (1+O(N^{-c}))(2e)^{-(N-2)/2}\sqrt{\frac{2}{NG''(\gamma_{a,b}; \boldsymbol{A}^{(2)}, \boldsymbol{u})}} \\
&\qquad \exp\Bigg(N\Bigg[N^{-1/2}g_{a,b} + \Gamma_N(\boldsymbol{y}, \boldsymbol{m}; a, b) + \frac{N-3}{2N}\log(1-r(a,b)) + \min_{z>z_*} G(z; \boldsymbol{A}^{(2)}, \boldsymbol{u}) \\
&\qquad\qquad + \frac{1}{2}\left(\xi(1) - \xi(r(a,b)) - \xi'(r(a,b))(1-r(a,b)) - \frac{1}{2}\xi''(r(a,b))(1-r(a,b))^2\right)\Bigg]\Bigg), \quad (6.205)
\end{aligned}$$

where we have simplified using Eq. (6.204). By independence of $\boldsymbol{u}, \tilde{W}^{(2)}$, and the fact that $\tilde{W}^{(2)}$ is a GOE matrix scaled by $\tilde{\gamma}_2/\sqrt{2}$, the following holds with probability at least $1-\exp(-N^c)$ provided $z > \tilde{\gamma}_2\sqrt{2}+\delta$ for some constant $\delta > 0$

$$G(z; \boldsymbol{A}^{(2)}, \boldsymbol{u}) = G_{a,b}(z) + O(1/N), \tag{6.206}$$

where

$$G_{a,b}(z) := z - \frac{1}{2}\left(\psi(z\sqrt{2}/\tilde{\gamma}_2) + \log(\tilde{\gamma}_2/\sqrt{2})\right) + \frac{1}{4}\left(\tilde{\gamma}_1^2 + (1-r(a,b))\left(1-\frac{\xi'(aq)}{\xi'(q)}\right)^2 t\right)\phi(z\sqrt{2}/\tilde{\gamma}_2), \tag{6.207}$$

and, for $x > 2$,

$$\phi(x) = \frac{1}{2}(x-\sqrt{x^2-4}), \qquad \psi(x) = \frac{1}{2}((x-\sqrt{x^2-4})/2)^2 - \log((x-\sqrt{x^2-4})/2).$$

Note that $\phi(x) = \int(x-u)^{-1}\mu_{\mathsf{sc}}(\mathrm{d}u)$ and $\psi(x) = \int\log(x-u)\mu_{\mathsf{sc}}(\mathrm{d}u)$ where μ_{sc} is the semicircular law. Moreover, $\psi'(x) = \phi(x)$.

Thus,

$$\int_{D_N(a,b)} e^{H(\boldsymbol{\sigma})}\mu_0^{a,b}(\mathrm{d}\boldsymbol{\sigma}) = \sqrt{\frac{2}{NG''(\gamma_{a,b}; \boldsymbol{A}^{(2)}, \boldsymbol{u})}}\exp\left(NE(a,b) + N^{1/2}g_{a,b} + O(1)\right), \tag{6.208}$$

where we define

$$\begin{aligned}
E(a,b) := &-\frac{N-2}{2N}\ln(2e) + \frac{N-3}{2N}\log(1-r(a,b)) + \min_{z>\tilde{\gamma}_2\sqrt{2}} G_{a,b}(z) + \Gamma_N(\boldsymbol{y}, \boldsymbol{m}; a, b) \\
&+ \frac{1}{2}\left(\xi(1) - \xi(r(a,b)) - \xi'(r(a,b))(1-r(a,b)) - \frac{1}{2}\xi''(r(a,b))(1-r(a,b))^2\right). \qquad (6.209)
\end{aligned}$$

Let $b_* = \langle \boldsymbol{x}, \boldsymbol{m}\rangle_N$. Note that $r(1, b_*) = q$. Furthermore, we have

$$r(a,b) = a^2 q + \frac{(b - a\langle \boldsymbol{x}, \boldsymbol{m}\rangle_N)^2}{1-\langle \boldsymbol{x}, \boldsymbol{m}\rangle_N^2/q}. \tag{6.210}$$

We will next verify several properties of $E(a,b)$, starting with the observation that $(a,b) = (1, b_*)$ is a stationary point of E.

Lemma 6.7.15. *We have $\nabla E(a,b)|_{(a,b)=(1,b_*)} = 0$. (Here ∇ denotes gradient with respect to (a,b).)*

Proof. We first compute $\nabla \min_z G_{a,b}(z)$. Let $z_*(a,b) = \arg\min_z G_{a,b}(z)$ and $z_* = \arg\min_z G_{1,b_*}(z)$ so

$$\partial_z G_{a,b}(z)|_{z_*(a,b)} = 0 \Leftrightarrow 1 - \frac{1}{\sqrt{2}\tilde{\gamma}_2}\phi(z\sqrt{2}/\tilde{\gamma}_2) = 0. \tag{6.211}$$

For $\alpha \in \{a, b\}$,

$$\partial_\alpha z_*(a,b) = (\partial_z^2 G_{a,b}(z))^{-1}|_{(a,b,z_*(a,b))}\partial_\alpha \partial_z G_{a,b}(z)|_{(a,b,z_*(a,b))}. \tag{6.212}$$

A quick calculation shows that $z_* = 1/2 + \tilde{\gamma}_2^2$ when $(a,b) = (1,b_*)$, and for $\alpha \in \{a,b\}$,

$$\partial_\alpha \min_z G_{a,b}(z)|_{(a,b)} = \partial_\alpha G_{a,b}(z_*(a,b))|_{(a,b)}.$$

Also note that

$$\nabla \left(\tilde{\gamma}_1^2 + (1 - r(a,b))\left(1 - \frac{\xi'(aq)}{\xi'(q)}\right)^2 t\right)\Bigg|_{(1,b_*)} = 0; \quad \left(\tilde{\gamma}_1^2 + (1 - r(a,b))\left(1 - \frac{\xi'(aq)}{\xi'(q)}\right)^2 t\right)\Bigg|_{(1,b_*)} = 0. \tag{6.213}$$

From the definition of G and the stationary condition (6.211), we obtain that

$$\nabla \min_z G_{a,b}(z)|_{(1,b_*)} = \frac{1}{2}\left(-\tilde{\gamma}_2^{-1} + \sqrt{2}\tilde{\gamma}_2^{-2} z_* \phi(z_*\sqrt{2}/\tilde{\gamma}_2)\right)\nabla\tilde{\gamma}_2 = \tilde{\gamma}_2 \nabla \tilde{\gamma}_2 = \frac{1}{2}\nabla(\tilde{\gamma}_2^2). \tag{6.214}$$

Furthermore,

$$\nabla \tilde{\gamma}_2^2 = -\xi''(r(a,b))(1 - r(a,b))\nabla r(a,b) + \frac{1}{2}\xi'''(r(a,b))(1 - r(a,b))^2 \nabla r(a,b).$$

We have

$$\begin{aligned}
&\nabla\left(\frac{1}{2}\log(1 - r(a,b)) + \frac{1}{2}(\xi(1) - \xi(r(a,b)) - \xi'(r(a,b))(1 - r(a,b)) - \frac{1}{2}\xi''(r(a,b))(1 - r(a,b))^2)\right) \\
&= \frac{-1}{2(1 - r(a,b))}\nabla r(a,b) - \frac{1}{4}\xi'''(r(a,b))(1 - r(a,b))^2 \nabla r(a,b),
\end{aligned} \tag{6.215}$$

and furthermore $\partial_a r(a,b)|_{(1,b_*)} = 2q$, $\partial_b r(a,b)|_{(1,b_*)} = 0$ and $r(1,b_*) = q$.

Recall

$$\langle \boldsymbol{z} + \boldsymbol{y}, \boldsymbol{\sigma}\rangle_N = -\xi'(\langle \boldsymbol{x}, \boldsymbol{m}\rangle_N)b + aq\left((1-q)\xi''(q) + \frac{1}{1-q}\right).$$

Moreover,

$$\langle \boldsymbol{y} - \boldsymbol{y}^\perp, \boldsymbol{\sigma} - \boldsymbol{\sigma}^\perp\rangle_N = a\langle \boldsymbol{y}, \boldsymbol{m}\rangle_N + \frac{b - a\langle \boldsymbol{x}, \boldsymbol{m}\rangle_N}{1 - \langle \boldsymbol{x}, \boldsymbol{m}\rangle_N^2/q}\left(\langle \boldsymbol{y}, \boldsymbol{x}\rangle_N - \langle \boldsymbol{x}, \boldsymbol{m}\rangle_N \langle \boldsymbol{y}, \boldsymbol{m}\rangle_N/q\right).$$

Hence,

$$\begin{aligned}
&\frac{\xi'(aq)\langle \boldsymbol{y} + \boldsymbol{z}, \boldsymbol{\sigma}\rangle_N}{\xi'(q)} - \frac{\xi''(q)\xi'(aq)aq\langle \boldsymbol{m}, \boldsymbol{z}\rangle_N}{\xi'(q)\zeta(q)} + \left(1 - \frac{\xi'(aq)}{\xi'(q)}\right)\langle \boldsymbol{y} - \boldsymbol{y}^\perp, \boldsymbol{\sigma} - \boldsymbol{\sigma}^\perp\rangle_N \\
&= \frac{\xi'(aq)(-\xi'(\langle \boldsymbol{x}, \boldsymbol{m}\rangle_N)b + aq((1-q)\xi''(q) + \frac{1}{1-q}))}{\xi'(q)} - \frac{\xi''(q)\xi'(aq)aq\langle \boldsymbol{m}, \boldsymbol{z}\rangle_N}{\xi'(q)\zeta(q)} \\
&\qquad + \left(1 - \frac{\xi'(aq)}{\xi'(q)}\right)\left(a\langle \boldsymbol{y}, \boldsymbol{m}\rangle_N + \frac{b - a\langle \boldsymbol{x}, \boldsymbol{m}\rangle_N}{1 - \langle \boldsymbol{x}, \boldsymbol{m}\rangle_N^2/q}\left(\langle \boldsymbol{y}, \boldsymbol{x}\rangle_N - \langle \boldsymbol{x}, \boldsymbol{m}\rangle_N\langle \boldsymbol{y}, \boldsymbol{m}\rangle_N/q\right)\right) \\
&= \frac{\xi'(aq)(-\xi'(\langle \boldsymbol{x}, \boldsymbol{m}\rangle_N)b + aq((1-q)\xi''(q) + \frac{1}{1-q}))}{\xi'(q)} \\
&\qquad + \left(1 - \frac{\xi'(aq)}{\xi'(q)}\right)\frac{b - a\langle \boldsymbol{x}, \boldsymbol{m}\rangle_N}{1 - \langle \boldsymbol{x}, \boldsymbol{m}\rangle_N^2/q}(\langle \boldsymbol{y}, \boldsymbol{x}\rangle_N - \langle \boldsymbol{x}, \boldsymbol{m}\rangle_N\langle \boldsymbol{y}, \boldsymbol{m}\rangle_N/q) \\
&\qquad\quad + \frac{\xi''(q)\xi'(aq)aq\left(\xi'(\langle \boldsymbol{x}, \boldsymbol{m}\rangle_N)\langle \boldsymbol{x}, \boldsymbol{m}\rangle_N - q\xi''(q)(1-q) - \frac{q}{1-q}\right)}{\xi'(q)\zeta(q)} \\
&\qquad\quad + \langle \boldsymbol{y}, \boldsymbol{m}\rangle_N\left(\frac{\xi''(q)\xi'(aq)aq}{\xi'(q)\zeta(q)} + a\left(1 - \frac{\xi'(aq)}{\xi'(q)}\right)\right).
\end{aligned}$$

Note that

$$\partial_a\left(\left(1-\frac{\xi'(aq)}{\xi'(q)}\right)\frac{b-a\langle \boldsymbol{x},\boldsymbol{m}\rangle_N}{1-\langle \boldsymbol{x},\boldsymbol{m}\rangle_N^2/q}\right)\Big|_{(1,b_*)}=\partial_b\left(\left(1-\frac{\xi'(aq)}{\xi'(q)}\right)\frac{b-a\langle \boldsymbol{x},\boldsymbol{m}\rangle_N}{1-\langle \boldsymbol{x},\boldsymbol{m}\rangle_N^2/q}\right)\Big|_{(1,b_*)}=0,$$

and $\partial_a(\xi'(aq)aq)|_{(1,q)}=q\zeta(q)$ so

$$\partial_a\left(\frac{\xi''(q)\xi'(aq)aq}{\xi'(q)\zeta(q)}+a\left(1-\frac{\xi'(aq)}{\xi'(q)}\right)\right)|_{(1,b_*)}=\partial_b\left(\frac{\xi''(q)\xi'(aq)aq}{\xi'(q)\zeta(q)}+a\left(1-\frac{\xi'(aq)}{\xi'(q)}\right)\right)|_{(1,b_*)}=0.$$

Thus, we can compute

$$\begin{aligned}&\partial_a\left(\frac{\xi'(aq)\langle \boldsymbol{y}+\boldsymbol{z},\boldsymbol{\sigma}\rangle_N}{\xi'(q)}-\frac{\xi''(q)\xi'(aq)aq\langle \boldsymbol{m},\boldsymbol{z}\rangle_N}{\xi'(q)\zeta(q)}+\left(1-\frac{\xi'(aq)}{\xi'(q)}\right)\langle \boldsymbol{y}-\boldsymbol{y}^\perp,\boldsymbol{\sigma}-\boldsymbol{\sigma}^\perp\rangle_N\right)|_{(1,b_*)}\\&=q(1-q)\xi''(q)+\frac{q}{1-q}.\end{aligned}\tag{6.216}$$

Similarly,

$$\begin{aligned}&\partial_b\left(\frac{\xi'(aq)\langle \boldsymbol{y}+\boldsymbol{z},\boldsymbol{\sigma}\rangle_N}{\xi'(q)}-\frac{\xi''(q)\xi'(aq)aq\langle \boldsymbol{m},\boldsymbol{z}\rangle_N}{\xi'(q)\zeta(q)}+\left(1-\frac{\xi'(aq)}{\xi'(q)}\right)\langle \boldsymbol{y}-\boldsymbol{y}^\perp,\boldsymbol{\sigma}-\boldsymbol{\sigma}^\perp\rangle_N\right)|_{(1,b_*)}\\&=-\xi'(\langle \boldsymbol{x},\boldsymbol{m}\rangle_N).\end{aligned}\tag{6.217}$$

Combining (6.214), (6.215), (6.216), (6.217), we obtain the desired claim that $\nabla E(a,b)|_{(a,b)=(1,b_*)}=0$. □

Lemma 6.7.16. *We have*

$$\mathbb{E}\left[\int_{D_N(a,b)}e^{H(\boldsymbol{\sigma})}\mu_0^{a,b}(\mathsf{d}\boldsymbol{\sigma})\Big|\nabla\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m})=\boldsymbol{0},\boldsymbol{x},\boldsymbol{y}-\boldsymbol{y}^\perp,g_{a,b}\right]=\exp\left\{N\hat{E}(a,b)-\log(2e)+\sqrt{N}g_{a,b}\right\},\tag{6.218}$$

where

$$\begin{aligned}\hat{E}(a,b):=&\frac{1}{N}\ln(2e)+\frac{N-3}{2N}\log(1-r(a,b))+\left(\frac{1}{2}\tilde{\gamma}_2^2+\frac{1}{2}\tilde{\gamma}_1^2\right)+\Gamma_N(\boldsymbol{y},\boldsymbol{m};a,b)\\&+\frac{1}{2}\left(\xi(1)-\xi(r(a,b))-\xi'(r(a,b))(1-r(a,b))-\frac{1}{2}\xi''(r(a,b))(1-r(a,b))^2\right).\end{aligned}\tag{6.219}$$

Furthermore, $E(a,b)$ is uniformly upper bounded by $\hat{E}(a,b)$, $E(1,b_*)=\hat{E}(1,b_*)$, and $\nabla E(1,b_*)=\nabla\hat{E}(1,b_*)=0$.

Proof. Eq (6.218) follows from a direct calculation. For the last claim, let $\tilde{\gamma}_2=\tilde{\gamma}_2(a,b)$ and $\tilde{\gamma}_1=\tilde{\gamma}_1(a,b)$. Given a quadratic Hamiltonian $H_{\le 2}(\boldsymbol{\sigma})=\tilde{\gamma}_2\langle\boldsymbol{\sigma},\boldsymbol{A}\boldsymbol{\sigma}\rangle/\sqrt{2}+\tilde{\gamma}_1\langle\boldsymbol{g},\boldsymbol{\sigma}\rangle$ where $\boldsymbol{A}$ is a GOE matrix and $\boldsymbol{g}\sim\mathcal{N}(0,\boldsymbol{I}_N)$,

$$\mathbb{E}\left[\int e^{H_{\le 2}(\boldsymbol{\sigma})}\mu_0(\mathsf{d}\boldsymbol{\sigma})\right]=e^{N\tilde{\gamma}_2^2/2+N\tilde{\gamma}_1^2/2}.$$

On the other hand, by Lemma 6.7.3, we have, for $\tilde{\gamma}_1$ sufficiently small, with probability at least $1-e^{-cN}$,

$$\int e^{H_{\le 2}(\boldsymbol{\sigma})}\mu_0(\mathsf{d}\boldsymbol{\sigma})\ge\exp\left\{(1-o(1))N\Big(\min_{z>\tilde{\gamma}_2\sqrt{2}}G_{a,b}(z)-\frac{1}{2}\log(2e)\Big)\right\}.$$

Since this holds for all N, Markov inequality implies

$$\min_{z>\tilde{\gamma}_2\sqrt{2}}G_{a,b}(z)-\frac{1}{2}\log(2e)\le\frac{1}{2}\tilde{\gamma}_2^2+\frac{1}{2}\tilde{\gamma}_1^2\,.$$

The last claim follows immediately upon this observation. □

When $\langle \boldsymbol{x}, \boldsymbol{m}\rangle_N = q$, $\langle \boldsymbol{y}, \boldsymbol{m}\rangle_N = t$, $\langle \boldsymbol{y}, \boldsymbol{x}\rangle_N = t$, $\|\boldsymbol{y}\|^2 = t + t^2$, under the constraint $\xi'(q) + t = \frac{q}{1-q}$, we can simplify $\hat{E}(a, b)$ as

$$\tilde{E}(a,b) := \frac{1}{N}\ln(2e) + \frac{N-3}{2N}\log(1 - r(a,b)) + \frac{1}{2}\tilde{\gamma}_1^2 + \xi(b) - b\xi'(aq) + \frac{\xi'(aq)aq}{(1-q)\xi'(q)} + at\left(1 - \frac{\xi'(aq)}{\xi'(q)}\right)$$
$$+ \frac{1}{2}\left(\xi(1) - \xi(r(a,b)) - \xi'(r(a,b))(1 - r(a,b))\right). \tag{6.220}$$

Indeed, under these values and constraints,

$$\begin{aligned}\Gamma_N(\boldsymbol{y}, \boldsymbol{m}; a, b) &= \xi(b) + \frac{\xi'(aq)\langle \boldsymbol{y} + \boldsymbol{z}, \boldsymbol{\sigma}\rangle}{\xi'(q)} - \frac{\xi''(q)\xi'(aq)aq\langle \boldsymbol{m}, \boldsymbol{z}\rangle}{\xi'(q)\theta(q)} + \left(1 - \frac{\xi'(aq)}{\xi'(q)}\right)\langle \boldsymbol{y} - \boldsymbol{y}^{\perp}, \boldsymbol{\sigma} - \boldsymbol{\sigma}^{\perp}\rangle \\ &= \xi(b) - \frac{\xi''(q)\xi'(aq)aq(-t - \xi'(q)q + \frac{q}{1-q} + q(1-q)\xi''(q))}{\xi'(q)\theta(q)} \\ &\qquad + \frac{\xi'(aq)(-\xi'(q)b + aq((1-q)\xi''(q) + \frac{1}{1-q}))}{\xi'(q)} + t\left(1 - \frac{\xi'(aq)}{\xi'(q)}\right) \\ &= \xi(b) - b\xi'(aq) + \frac{\xi'(aq)aq}{(1-q)\xi'(q)} + at\left(1 - \frac{\xi'(aq)}{\xi'(q)}\right).\end{aligned}$$

Furthermore, in this case, $b_* = q$.

Lemma 6.7.17. *For $\varepsilon > 0$ sufficiently small, there is $\eta > 0$ such that for all (a, b) satisfying $|aq-q|+|b-q| \leq \varepsilon$, we have $\nabla^2\tilde{E}(a,b) \preceq -\eta \boldsymbol{I}_2$.*

Proof. We have

$$\begin{aligned}\partial_b^2\tilde{E}(a,b)|_{(1,q)} &= 2\xi''(q) - \left(\frac{1}{2(1 - r(a,b))} + \frac{1}{2}\xi''(r(a,b))(1 - r(a,b))\right)\partial_b^2(r(a,b))|_{(1,q)} \\ &= -\frac{1}{(1-q)^2} + \xi''(q),\end{aligned} \tag{6.221}$$

and

$$\begin{aligned}\partial_b\partial_a\tilde{E}(a,b)|_{(1,q)} &= -\left(\frac{1}{2(1 - r(a,b))} + \frac{1}{2}\xi''(r(a,b))(1 - r(a,b))\right)\partial_{a,b}(r(a,b))|_{(1,q)} \\ &= -q\xi''(q) + \frac{q}{(1-q)^2}.\end{aligned} \tag{6.222}$$

Finally, we compute

$$\begin{aligned}&\partial_a^2\tilde{E}(a,b)|_{(1,q)} \\ &= -q^3\xi'''(q) + q(2\xi''(q) + q\xi'''(q)) + (1-q)(q^2\xi'''(q) + q\xi''(q) - q^2\xi''(q)^2/\xi'(q) + q^2v\xi''(q)^2/\xi'(q)^2) \\ &\quad - \frac{q(2q+1)}{(1-q)^2} - 2q^2(1-q)\xi'''(q) + q(2q-1)\xi''(q) \\ &= -\frac{q(2q+1)}{(1-q)^2} + q(q+2)\xi''(q) + (1-q)(-q^2\xi''(q)^2/\xi'(q) + q^2v\xi''(q)^2/\xi'(q)^2).\end{aligned} \tag{6.223}$$

Using the constraints $v = t - \frac{t^2(1-q)}{q}$, and that $t = \frac{q}{1-q} - \xi'(q)$, we can simplify

$$\partial_a^2\tilde{E}(a,b)|_{(1,q)} = -\frac{q(2q+1)}{(1-q)^2} + q(q+2)\xi''(q) - q(1-q)^2\xi''(q)^2. \tag{6.224}$$

Consider a change of variable $\tilde{a} = aq$ and let $\underline{E}(\tilde{a}, b) = \tilde{E}(\tilde{a}/q, b)$. Combining (6.221), (6.222), (6.224), under the condition $\xi''(q) < \frac{1}{(1-q)^2}$, that $\underline{E}(\tilde{a}, b)$ is strictly concave at $(\tilde{a}, b) = (q, q)$ is equivalent to

$$\begin{aligned}&\frac{1}{q}\left((1-q)^2\xi''(q)^2 - (q+2)\xi''(q) + \frac{(2q+1)}{(1-q)^2}\right) > \left(\frac{1}{(1-q)^2} - \xi''(q)\right) \\ &\Leftrightarrow \frac{1}{q}\left((1-q)\xi''(q) - \frac{1}{1-q}\right)^2 + \frac{1}{(1-q)^2} > 0.\end{aligned}$$

□

Notice that, for (a,b) in a neighborhood of $(1,b_*)$, the Hessian of $\hat{E}(a,b)$ is a continuous rational function of q, $\xi(b)$, $\xi(q)$, $\xi'(q)$, $\xi''(q)$, $\xi'''(q)$, $\xi'(\langle \boldsymbol{x}, \boldsymbol{m}\rangle_N)$, and $\langle \boldsymbol{x}, \boldsymbol{y}\rangle_N, \langle \boldsymbol{x}, \boldsymbol{m}\rangle_N, \langle \boldsymbol{y}, \boldsymbol{m}\rangle_N, \|\boldsymbol{y}\|_N^2$. Hence, we have the following implication of the previous lemma.

Corollary 6.7.18. *There exist $\varepsilon, \eta > 0$ such that, for $|\langle \boldsymbol{x}, \boldsymbol{m}\rangle_N - q| \le \varepsilon$, $|\langle \boldsymbol{y}, \boldsymbol{m}\rangle_N - t| \le \varepsilon$, $|\langle \boldsymbol{y}, \boldsymbol{x}\rangle_N - t| \le \varepsilon$, and $|\|\boldsymbol{y}\|_N^2 - t| \le \varepsilon$, all (a,b) such that $|aq - a_*q| + |b - b_*| \le \varepsilon$, we have $\nabla^2 \hat{E}(a,b) \preceq -\eta \boldsymbol{I}_2$. (Here $(a_*, b_*) = (1, q)$.)*

We will next prove several simple preliminary estimates before giving the proof of Proposition 6.4.6.

Recall that on $D_N(a,b)$, we have defined the Hamiltonian $\underline{H}(\tilde{\boldsymbol{\sigma}})$, which is a spin glass with mixture given by Eqs. (6.199) to (6.201). Let $\boldsymbol{A}^{(p)}(a,b) = \nabla^p \underline{H}(\boldsymbol{0})$ and $\boldsymbol{u}(a,b) = \nabla \underline{H}(\boldsymbol{0})$.

By Lemma 6.7.15, Lemma 6.7.16, and Corollary 6.7.18 and the preceding remark, there is a unique local maxima $(a_*, b_*) = (1, b_*)$ of $E(a,b)$ and $\widehat{E}(a,b)$ with $|qa_* - q| + |b_* - q| \le \varepsilon$, and $\widehat{E}(a,b)$ is strongly concave at (a_*, b_*). In particular, there is $\eta > 0$ such that, for sufficiently small ε and (a,b) such that $|qa - qa_*| + |b - b_*| \le \varepsilon$, we have

$$E(a,b) \le E(a_*, b_*) - \eta(|qa - qa_*|^2 + |b - b_*|^2). \tag{6.225}$$

For each a, b let $\boldsymbol{m}(a,b)$ be the unique point in $V := \mathsf{span}(\boldsymbol{m}, \boldsymbol{x})$ with $\|\boldsymbol{m}(a,b)\|_N^2 = qa$ and $\langle \boldsymbol{m}(a,b), \boldsymbol{x}\rangle_N = b$.

The following lemma follows from standard control on suprema of Gaussian processes (see, e.g. [MS23, Lemma A.3]).

Lemma 6.7.19. *For $\varepsilon > 0$ sufficiently small there exist $c = c(\varepsilon)$, $C = C(\varepsilon) > 0$ depending uniquely on ε such that the following holds with probability at least $1 - e^{-cN}$ conditional on $\nabla \mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m}) = 0$. For (a,b) such that $|qa - qa_*| + |b - b_*| \le \varepsilon$, we have*

$$\nabla H(\boldsymbol{m}(a,b)) = \nabla H(\boldsymbol{m}) + \nabla^2 H(\boldsymbol{m})(\boldsymbol{m}(a,b) - \boldsymbol{m}) + \mathsf{Err},$$

where $\|\mathsf{Err}\| \le CN^{-1/2}\|\boldsymbol{m}(a,b) - \boldsymbol{m}\|^2$. Furthermore, $\mathsf{proj}_V^{\perp}(\nabla H(\boldsymbol{m})) = 0$, and for $\boldsymbol{v} \in V$, $\|\nabla^2 H(\boldsymbol{m})\boldsymbol{v}\|^2 \le C\|\boldsymbol{v}\|^2$. (Here $V = \mathsf{span}(\boldsymbol{m}, \boldsymbol{x})$.)

As a corollary of Lemma 6.7.19, we obtain the following control on the effective fields $\boldsymbol{u}(a,b) = (1 - r(a,b))^{1/2}\mathsf{proj}_V^{\perp}(\nabla H(\boldsymbol{m}))$.

Lemma 6.7.20. *For $\varepsilon, \delta > 0$ sufficiently small, the following holds with probability at least $1 - e^{-cN}$. There exists $\boldsymbol{u}_1, \boldsymbol{u}_2 \in \mathbb{R}^{N-3}$ with $\|\boldsymbol{u}_1\|, \|\boldsymbol{u}_2\| = O(N^{1/2})$ such that, for any (a,b) with $|qa - qa_*| + |b - b_*| \le \varepsilon$, we have $\|\boldsymbol{u}(a,b) - (qa - qa_*)\boldsymbol{u}_1 - (b - b_*)\boldsymbol{u}_2\| \le CN^{1/2}(|qa - qa_*|^2 + |b - b_*|^2)$.*

Furthermore, for $\gamma > \delta + \mathbb{E}\lambda_{\max}(\boldsymbol{A}^{(2)}(a_*, b_*))$ and $i, j \in \{1, 2\}$, there is $c = c(\delta) > 0$ such that, with probability $1 - e^{-N^c}$, $\langle \boldsymbol{u}_i, (\gamma \boldsymbol{I} - \boldsymbol{A}^{(2)}(a_*, b_*))^{-1}\boldsymbol{u}_j\rangle$, concentrates in a window of size $O(N^{1/2+c})$ around its expectation.

Proof. The first part follows from Lemma 6.7.19, using $\mathsf{proj}_V^{\perp}(\nabla H(\boldsymbol{m})) = \boldsymbol{0}$.

The second part holds by concentration of Lipschitz functions of Gaussian random variables. Indeed note that $\boldsymbol{u}_i$ depend linearly on $\boldsymbol{A}^{(2)}(a_*, b_*)$ as well as on independent Gaussian random variables. Under the high probability event $\mathbb{E}\lambda_{\max}(\boldsymbol{A}^{(2)}(a_*, b_*)) + \delta/2 \le \lambda_{\max}(\boldsymbol{A}^{(2)}(a_*, b_*)) \le C$, the quantity $\langle \boldsymbol{u}_i, (\gamma \boldsymbol{I} - \boldsymbol{A}^{(2)}(a_*, b_*))^{-1}\boldsymbol{u}_j\rangle$ is indeed Lipschitz in these Gaussians as well as on $\boldsymbol{A}^{(2)}(a_*, b_*)$. □

Let

$$R := \left\{(a,b) : q|a - a_*| + |b - b_*| \le N^{-1/2+c}\right\}. \tag{6.226}$$

Recall the random shifts $g_{a,b}$ in Eq. (6.197). We have the following control on $g_{a,b}$, again from standard control on Gaussian processes.

Lemma 6.7.21. *We have that $g_{a,b}$, for $(a,b) \in R$, forms a Gaussian process with $\mathbb{E}[(g_{a,b} - g_{a',b'})^2] = O(\|\boldsymbol{m}(a,b) - \boldsymbol{m}(a',b')\|_N^2)$. Furthermore, with probability at least $1 - e^{-cN}$, we have, for all $(a,b) \in R$ that*

$$|g_{a,b} - g_{a_*,b_*}| \le C(|q(a - a_*)| + |b - b_*|).$$

Proof. The first claim follows from a standard calculation, and the second claim follows Sudakov-Fernique inequality, comparing with the linear process $\langle \boldsymbol{g}, \boldsymbol{m}(a,b) - \boldsymbol{m}(a_*, b_*)\rangle/\sqrt{N}$ for $\boldsymbol{g}$ a standard normal vector. $\square$

Lemma 6.7.22. *The scaled GOE matrices $\boldsymbol{A}^{(2)}(a,b)$ for $(a,b) \in R$ form a Gaussian process with metric*

$$\mathbb{E}\Big\{\big\|\boldsymbol{A}^{(2)}(a,b) - \boldsymbol{A}^{(2)}(a',b')\big\|_{\mathrm{F}}^2\Big\} \le CN(|qa - qa'|^2 + |b - b'|^2)\,.$$

Furthermore, for any $\eta > 0$ there exist constants $c, C > 0$ such that, with probability at least $1 - e^{-cN}$,

$$\|\boldsymbol{A}^{(2)}(a,b) - \boldsymbol{A}^{(2)}(a',b')\|_{\mathrm{F}}^2 \le CN(|qa - qa'|^2 + |b - b'|^2)^{1+\eta}\,, \quad \forall (a,b), (a',b') \in R\,, \tag{6.227}$$

$$\|\boldsymbol{A}^{(2)}(a,b) - \boldsymbol{A}^{(2)}(a',b')\|_{\mathsf{op}}^2 \le C(|qa - qa_*|^2 + |b - b_*|^2)^{1+\eta}\,, \quad \forall (a,b), (a',b') \in R\,. \tag{6.228}$$

and

$$\sup_{(a,b)\in R} \|\boldsymbol{A}^{(2)}(a,b) - \boldsymbol{A}^{(2)}(a_*,b_*)\|_{\mathrm{F}}^2 \le CN^{2c}\,, \tag{6.229}$$

$$\sup_{(a,b)\in R} \|\boldsymbol{A}^{(2)}(a,b) - \boldsymbol{A}^{(2)}(a_*,b_*)\|_{\mathsf{op}}^2 \le CN^{-1+2c}\,. \tag{6.230}$$

Proof. The bound on the canonical distance of $\boldsymbol{A}^{(2)}$ follows from a straightforward calculation.

The bounds (6.227) and (6.229) follow from chaining on R, together with the standard bound on chi-squared random variables

$$\mathbb{P}\left(\|\boldsymbol{A}^{(2)}(a,b) - \boldsymbol{A}^{(2)}(a',b')\|_{\mathrm{F}}^2 > \kappa\,\mathbb{E}\|\boldsymbol{A}^{(2)}(a,b) - \boldsymbol{A}^{(2)}(a',b')\|_{\mathrm{F}}^2\right) \le 2e^{-cN^2[(\kappa-1)\wedge(\kappa-1)^2]},$$

for $\kappa > 1$.

The bound (6.228) and (6.230) follow from a similar chaining argument. Indeed, $\boldsymbol{A}^{(2)}(a,b) - \boldsymbol{A}^{(2)}(a',b')$ is a matrix with independent entries with variance bounded by $C(|qa - qa_*|^2 + |b - b_*|^2)/N$, whence by standard estimates on the norm of Gaussian random matrices, the following holds for all $\kappa > \kappa_0$

$$\mathbb{P}\left(\|\boldsymbol{A}^{(2)}(a,b) - \boldsymbol{A}^{(2)}(a',b')\|_{\mathsf{op}}^2 > \kappa\, N^{-1}\mathbb{E}\|\boldsymbol{A}^{(2)}(a,b) - \boldsymbol{A}^{(2)}(a',b')\|_{\mathrm{F}}^2\right) \le 2e^{-cN\kappa}.$$

$\square$

Recall $G(\gamma; \boldsymbol{A}, \boldsymbol{u})$ in (6.102) and $G_{a,b}(\gamma)$ in (6.207). The next lemma gives control over G and $G_{a,b}$ for $(a,b) \in R$.

Lemma 6.7.23. *Given a compact interval $I \subseteq [M, \infty)$, $M := \varepsilon + \mathbb{E}\lambda_{\max}(\boldsymbol{A}^{(2)}(a_*, b_*))$, the following holds with probability at least $1 - e^{-N^\delta}$ for appropriate $C_0, c, \delta > 0$ depending on $\varepsilon > 0$:*

1. *We have*

$$\sup_{\gamma\in I,(a,b)\in R} \Big| N(G(\gamma; \boldsymbol{A}^{(2)}(a,b), \boldsymbol{u}(a,b)) - G_{a,b}(\gamma)) - N(G(\gamma; \boldsymbol{A}^{(2)}(a_*,b_*), \boldsymbol{u}(a_*,b_*)) - G_{a_*,b_*}(\gamma)) \Big|$$
$$= O(N^{-1/2+c}). \tag{6.231}$$

2. *We have*

$$\sup_{\substack{(a,b),(a',b')\in R:(a,b)+(a',b')=2(a_*,b_*)\\ \gamma\in I}} \Big| NG(\gamma; \boldsymbol{A}^{(2)}(a,b), \boldsymbol{u}(a,b)) - NG(\gamma; \boldsymbol{A}^{(2)}(a',b'), \boldsymbol{u}(a',b')) \Big|$$
$$= O(N^{-1/2+c}). \tag{6.232}$$

3. *The event in Lemma 6.7.3 holds uniformly in $(a,b) \in R$. Namely,*

$$Z_{\le 2}(a,b) = \int e^{\langle \boldsymbol{\sigma}, \boldsymbol{A}^{(2)}(a,b)\boldsymbol{\sigma}\rangle + \langle \boldsymbol{u}(a,b), \boldsymbol{\sigma}\rangle} \mu_0^{a,b}(\mathsf{d}\boldsymbol{\sigma})$$
$$= (1 + \mathsf{Err}_{a,b}(N))(2e)^{-(N-2)/2}\sqrt{\frac{2}{G''(\gamma_{a,b}; \boldsymbol{A}^{(2)}(a,b), \boldsymbol{u}(a,b))}}\, e^{NG(\gamma_{a,b}; \boldsymbol{A}^{(2)}(a,b), \boldsymbol{u}(a,b))}\,, \tag{6.233}$$

where $\sup_{(a,b)\in R} |\mathsf{Err}_{a,b}(N)| \le C_0, N^{-c}$.

Proof. We can represent $\boldsymbol{A}^{(2)}(a,b) = \boldsymbol{A}^{(2)}(a_*,b_*)+\boldsymbol{\Delta}(a,b)$ where each entry of $\boldsymbol{\Delta}(a,b)$ forms an independent Gaussian process with metric $\mathbb{E}[(\Delta_{i,j}(a,b)-\Delta_{i,j}(a',b'))^2]^{1/2} \le CN^{-1/2}(q|a-a'|+|b-b'|)$, and $\boldsymbol{\Delta}(a_*,b_*) = 0$.

Letting $\boldsymbol{Q}_*(\gamma) = \gamma\boldsymbol{I} - \boldsymbol{A}^{(2)}(a_*,b_*)$, we can expand

$$
\begin{aligned}
&G(\gamma;\boldsymbol{A}^{(2)}(a,b),\boldsymbol{u}(a,b)) = \\
&= \gamma - \frac{1}{2N}\log\det(\gamma\boldsymbol{I} - \boldsymbol{A}^{(2)}(a,b)) + \frac{1}{4N}\langle \boldsymbol{u}(a,b), (\gamma\boldsymbol{I}-\boldsymbol{A}^{(2)}(a,b))^{-1}\boldsymbol{u}(a,b)\rangle \\
&= G(\gamma;\boldsymbol{A}^{(2)}(a_*,b_*),\boldsymbol{u}(a_*,b_*)) - \frac{1}{2N}\log\det\left(\boldsymbol{I} - \boldsymbol{Q}_*(\gamma)^{-1/2}\boldsymbol{\Delta}(a,b)\boldsymbol{Q}_*(\gamma)^{-1/2}\right) \\
&\qquad + \frac{1}{4N}\langle \boldsymbol{u}(a,b), (\boldsymbol{Q}_*(\gamma)-\boldsymbol{\Delta}(a,b))^{-1}\boldsymbol{u}(a,b)\rangle - \frac{1}{4N}\langle \boldsymbol{u}(a,b), \boldsymbol{Q}_*(\gamma)^{-1}\boldsymbol{u}(a,b)\rangle.
\end{aligned}
\tag{6.234}
$$

Next, for $k \ge 2$, let

$$
X_k(a,b) = \mathsf{Tr}\left(\left(\boldsymbol{Q}_*(\gamma)^{-1/2}\boldsymbol{\Delta}(a,b)\boldsymbol{Q}_*(\gamma)^{-1/2}\right)^k\right). \tag{6.235}
$$

We have

$$
|X_k(a,b) - X_k(a',b')| \le C_k N^{-(k-1)(1/2-c)+1/2}\|\boldsymbol{A}^{(2)}(a,b) - \boldsymbol{A}^{(2)}(a',b')\|_{\mathrm{F}},
$$

under the event in Lemma 6.7.22. Recall that this also guarantees

$$
\sup_{(a,b)\in R}\left\{\|\boldsymbol{\Delta}(a,b)\|_{\mathsf{op}} \vee \|\boldsymbol{Q}_*(\gamma)^{-1/2}\boldsymbol{\Delta}(a,b)\boldsymbol{Q}_*(\gamma)^{-1/2}\|_{\mathsf{op}}\right\} \le CN^{-1/2+c}\,, \tag{6.236}
$$

Hence, under the event in Lemma 6.7.22, we have $|X_k(a,b) - X_k(a',b')| \le N^{-(k-1)(1/2-c)-L+1}$ whenever $q|a-a'|+|b-b'| < N^{-L}$.

Let $\mathbb{P}_{\boldsymbol{\Delta}}$ and $\mathbb{E}_{\boldsymbol{\Delta}}$ denote probability and expectation with respect to $\boldsymbol{\Delta}(a,b)$ only, i.e. conditional on $\boldsymbol{A}^{(2)}(a_*,b_*)$. Also let

$$
\boldsymbol{M}(a,b) = \boldsymbol{Q}_*(\gamma)^{-1/2}\boldsymbol{\Delta}(a,b)\boldsymbol{Q}_*(\gamma)^{-1/2}
$$

be the matrix appearing in $X_k(a,b)$. On the $\boldsymbol{A}^{(2)}(a_*,b_*)$-measurable, probability $1-e^{-cN}$ event that $\boldsymbol{Q}_*(\gamma)^{-1/2}$ is bounded in operator norm, $\boldsymbol{M}(a,b)$ is (conditional on $\boldsymbol{A}^{(2)}(a_*,b_*)$, in a suitable basis) a random matrix with independent centered gaussian entries, with variances not equal but bounded uniformly by N^{-2+2c}. It is well known, cf. [AGZ10, Chapter 2] that tracial moments of $\boldsymbol{M}(a,b)$ amount to certain (weighted) cycle counts, and from this a routine calculation gives

$$
\mathbb{E}_{\boldsymbol{\Delta}}X_k(a,b) = \begin{cases} 0 & k \text{ odd}, \\ O_k(N^{1-k(1/2-c)}) & k \text{ even}, \end{cases} \qquad \mathsf{Var}_{\boldsymbol{\Delta}}[X_k(a,b)] = O_k(N^{-k(1-2c)}).
$$

(The last estimate amounts to computing cycle counts for $\mathbb{E}[\mathsf{Tr}(\boldsymbol{M}(a,b)^k)^2]$ and $\mathbb{E}[\mathsf{Tr}(\boldsymbol{M}(a,b)^k)]^2$, cf. [AGZ10, Proof of Lemma 2.1.7].) For any fixed $(a,b)\in R$, Gaussian hypercontractivity gives

$$
\mathbb{P}_{\boldsymbol{\Delta}}\left\{|X_k(a,b) - \mathbb{E}_{\boldsymbol{\Delta}}X_k(a,b)| \ge t\sqrt{\mathsf{Var}_{\boldsymbol{\Delta}}(X_k(a,b))}\right\} \le \exp(-t^{c_k}),
$$

because $X_k(a,b)$ is a degree k-polynomial in the entries of $\boldsymbol{M}(a,b)$. By a union bound over a N^{-L}-net of R (of size N^{2L}), with probability $1-e^{-N^{c_k\delta}}$ over $\boldsymbol{\Delta}(a,b)$ the following holds.

For k even, uniformly in $(a,b)\in R$

$$
\mathsf{Tr}\left(\left(\boldsymbol{Q}_*(\gamma)^{-1/2}\boldsymbol{\Delta}(a,b)\boldsymbol{Q}_*(\gamma)^{-1/2}\right)^k\right) - \mathbb{E}\mathsf{Tr}\left(\left(\boldsymbol{Q}_*(\gamma)^{-1/2}\boldsymbol{\Delta}(a,b)\boldsymbol{Q}_*(\gamma)^{-1/2}\right)^k\right) = O_k(N^{\delta-k(1/2-c)}),
$$

and for k odd,

$$
\sup_{(a,b)\in R}\mathsf{Tr}\left(\left(\boldsymbol{Q}_*(\gamma)^{-1/2}\boldsymbol{\Delta}(a,b)\boldsymbol{Q}_*(\gamma)^{-1/2}\right)^k\right) = O_k(N^{\delta-k(1/2-c)}).
$$

Recall that $|\log(1-x)+x+x^2/2| \le |x|^3$ for all $|x| \le 1/4$. Therefore, uniformly in $(a,b) \in R$, for all δ, c small enough

$$\begin{aligned}
&\log\det\Big(\boldsymbol{I} - \boldsymbol{Q}_*(\gamma)^{-1/2}\boldsymbol{\Delta}(a,b)\boldsymbol{Q}_*(\gamma)^{-1/2}\Big) \\
&= -\sum_{k=1}^{2}\frac{1}{k}\mathsf{Tr}\left(\left(\boldsymbol{Q}_*(\gamma)^{-1/2}\boldsymbol{\Delta}(a,b)\boldsymbol{Q}_*(\gamma)^{-1/2}\right)^k\right) + O(N^{-1+c}) \\
&= -\mathbb{E}\mathsf{Tr}\left(\left(\boldsymbol{Q}_*(\gamma)^{-1/2}\boldsymbol{\Delta}(a,b)\boldsymbol{Q}_*(\gamma)^{-1/2}\right)^2\right) + O(N^{-1/2+c+\delta}).
\end{aligned} \tag{6.237}$$

We next turn to the term $\langle \boldsymbol{u}(a,b), (\boldsymbol{Q}_*(\gamma) - \boldsymbol{\Delta}(a,b))^{-1}\boldsymbol{u}(a,b)\rangle$ in Eq. (6.234). From Lemma 6.7.20, there are $\boldsymbol{u}_1, \boldsymbol{u}_2$ with $\|\boldsymbol{u}_1\|, \|\boldsymbol{u}_2\| = O(N^{1/2})$ such that letting $\boldsymbol{u}_0(a,b) := q(a-a_*)\boldsymbol{u}_1 + (b-b_*)\boldsymbol{u}_2$, we have $\|\boldsymbol{u}(a,b) - \boldsymbol{u}_0(a,b)\| \le CN^{-1/2+2c}$ and $\|\boldsymbol{u}_0(a,b)\| \le CN^c$ for any $(a,b) \in R$ Therefore, with probability $1-e^{-N^\delta}$,

$$\begin{aligned}
&\langle \boldsymbol{u}_0(a,b), (\boldsymbol{Q}_*(\gamma) - \boldsymbol{\Delta}(a,b))^{-1}\boldsymbol{u}_0(a,b)\rangle \\
&= \langle \boldsymbol{u}_0(a,b), \boldsymbol{Q}_*(\gamma)^{-1}\boldsymbol{u}_0(a,b)\rangle + \langle \boldsymbol{u}_0(a,b), \boldsymbol{Q}_*(\gamma)^{-1}\boldsymbol{\Delta}(a,b)\boldsymbol{Q}_*(\gamma)^{-1}\boldsymbol{u}_0(a,b)\rangle + O(N^{-1+3c}) \\
&= \langle \boldsymbol{u}_0(a,b), \boldsymbol{Q}_*(\gamma)^{-1}\boldsymbol{u}_0(a,b)\rangle + O(N^{-1+3c+\delta}).
\end{aligned}$$

where the first estimate follows from Eq. (6.236), and the second from independence of $\boldsymbol{\Delta}(a,b)$ and $\boldsymbol{u}_1, \boldsymbol{u}_2$, together with the fact that the entries of $\boldsymbol{\Delta}(a,b)$ have variance bounded by N^{-2+2c}. Therefore, we obtain that, with probability $1-e^{-N^\delta}$,

$$\begin{aligned}
&\langle \boldsymbol{u}(a,b), (\boldsymbol{Q}_*(\gamma) - \boldsymbol{\Delta}(a,b))^{-1}\boldsymbol{u}(a,b)\rangle \\
&= \langle q(a-a_*)\boldsymbol{u}_1 + (b-b_*)\boldsymbol{u}_2, \boldsymbol{Q}_*(\gamma)^{-1}(q(a-a_*)\boldsymbol{u}_1 + (b-b_*)\boldsymbol{u}_2)\rangle + O(N^{-1+4c+\delta}).
\end{aligned} \tag{6.238}$$

By similarly taking a union bound over a net of R of radius N^{-L} and using the continuity in Lemma 6.7.22, we can guarantee Eq. (6.238) uniformly in $(a,b) \in R$.

Combining the last conclusion in Lemma 6.7.20, Eqs. (6.237) and (6.238) and a union bound over γ, over any compact interval of γ, upon changing δ, with probability at least $1-e^{-N^\delta}$, that

$$\begin{aligned}
&\sup_{\gamma,(a,b)\in R}\Big|N(G(\gamma;\boldsymbol{A}^{(2)}(a,b),\boldsymbol{u}(a,b)) - G_{a,b}(\gamma) - G(\gamma;\boldsymbol{A}^{(2)}(a_*,b_*),\boldsymbol{u}(a_*,b_*)) + G_{a_*,b_*}(\gamma))\Big| \\
&\le O(N^{-1/2+c}).
\end{aligned}$$

Thus, we have, with probability at least $1-e^{-N^\delta}$, uniformly in $(a,b) \in R$ and γ, that Eq. (6.231) holds.

Given $(a,b), (a',b') \in R$ such that $(a,b) + (a',b') = 2(a_*,b_*)$, we have that

$$\mathbb{E}\mathsf{Tr}\left(\left(\boldsymbol{Q}_*(\gamma)^{-1/2}\boldsymbol{\Delta}(a,b)\boldsymbol{Q}_*(\gamma)^{-1/2}\right)^2\right) = \mathbb{E}\mathsf{Tr}\left(\left(\boldsymbol{Q}_*(\gamma)^{-1/2}\boldsymbol{\Delta}(a',b')\boldsymbol{Q}_*(\gamma)^{-1/2}\right)^2\right).$$

Combining with Eqs. (6.237) and (6.238), we then obtain Eq. (6.232).

Finally, we recall that for each $(a,b) \in R$, the event in Lemma 6.7.3 holds with probability at least $1-e^{-cN}$. On the other hand, for appropriate $C > C' > 1$, using Lemma 6.7.20 and Lemma 6.7.22, with probability $1-e^{-cN}$, uniformly over $(a,b), (a',b') \in R$ with $|q(a-a')| + |b-b'| \le N^{-C}$, we have $Z_{\le 2}(a,b) = (1+O(N^{-C'}))Z_{\le 2}(a',b')$. Similar continuity estimates hold for the right hand side of Eq. (6.233). Taking a net of radius N^{-C} of R and apply the union bound, we obtain Eq. (6.233) uniformly in $(a,b) \in R$. □

Proof of Proposition 6.4.6. Consider appropriate constants $c > c' > 0$. Define

$$D_N(\varepsilon) := \left\{\boldsymbol{\sigma} \in S_N :\ N^{-1/2+c} \le |\langle \boldsymbol{\sigma}, \boldsymbol{m}\rangle_N - a_*q| + |\langle \boldsymbol{\sigma}, \boldsymbol{x}\rangle_N - b_*| \le \varepsilon\right\}, \tag{6.239}$$

$$\hat{D}_N(\varepsilon) := \left\{(a,b) \in \mathbb{R}^2 :\ N^{-1/2+c} \le |aq - a_*q| + |b-b_*| \le \varepsilon\right\}, \tag{6.240}$$

Using Markov's Inequality and that the annealing upper bound $\widehat{E}(a,b)$ is strongly concave for $(a,b) \in \hat{D}_N(\varepsilon)$ (see Lemma 6.7.16 and Corollary 6.7.18), we obtain that with probability $1-\exp(-N^{c'})$, for all sufficiently small $\varepsilon > 0$, there is $\eta > 0$ such that

$$\begin{aligned}\int_{D_N(\varepsilon)} e^{H(\boldsymbol{\sigma})}\mu_0(\mathsf{d}\boldsymbol{\sigma}) &= \int_{\hat{D}_N(\varepsilon)} \left\{\int e^{H(\boldsymbol{\sigma})}\mu_0^{a,b}(\mathsf{d}\boldsymbol{\sigma})\right\} \mathsf{d}a\mathsf{d}b \\ &\le e^{-\eta(|qa-qa_*|^2+|b-b_*|^2)+N\widehat{E}(a_*,b_*)+\sqrt{N}g_{a,b}} \\ &= e^{-\eta(|qa-qa_*|^2+|b-b_*|^2)+NE(a_*,b_*)+\sqrt{N}g_{ab}}. \end{aligned} \tag{6.241}$$

Denote by $\langle\,\cdot\,\rangle_{a,b}$ the average with respect to the Gibbs measure restricted to band $D_N(a,b)$, namely with respect to $\mu^{a,b}(\mathsf{d}\boldsymbol{\sigma}) \propto \exp\{NH(\boldsymbol{\sigma})\}\,\mu_0^{a,b}(\mathsf{d}\boldsymbol{\sigma})$. Note that $\boldsymbol{u}(a_*,b_*)=0$, and for $|qa-qa_*|+|b-b_*| \le N^{-1/2+c}$, we have $\|\boldsymbol{u}(a,b)\| = O(N^{c'})$.

Recall $\gamma_{a,b} = \arg\min_{z>\lambda_{\max}(\boldsymbol{A}^{(2)}(a,b))} G(z;\boldsymbol{A}^{(2)}(a,b),\boldsymbol{u}(a,b))$. Let

$$\begin{aligned}\Delta(a,b) := \frac{1}{2}(\gamma_{a,b} - \boldsymbol{A}^{(2)}(a,b))^{-1}\boldsymbol{u}(a,b) \\ + \frac{1}{2}(\gamma_{a,b} - \boldsymbol{A}^{(2)}(a,b))^{-1}\langle \boldsymbol{A}^{(3)}(a,b),(\gamma_{a,b} - \boldsymbol{A}^{(2)}(a,b))^{-1}\rangle.\end{aligned}$$

We have $\Delta(a_*,b_*) = \Delta(\boldsymbol{m}) + O(N^{-c})$, where $\Delta(\boldsymbol{m})$ is defined as per Eq. (6.23).

Let

$$Z(a,b) := \int e^{H(\boldsymbol{\sigma})}\mu_0^{a,b}(\boldsymbol{\sigma}),$$

and

$$Z := \int_{\mathsf{Band}_*(2\iota)} \exp(H_{N,t}(\boldsymbol{\sigma}))\;\mu_0(\mathsf{d}\boldsymbol{\sigma}).$$

Recall the definitions

$$\begin{aligned}R &= \left\{(a,b) \in \mathbb{R}^2 : |qa-qa_*| + |b-b_*| \le N^{-1/2+c}\right\}, \\ R_+ &= \left\{(a,b) \in \mathbb{R}^2 : |qa-qa_*| + |b-b_*| \le N^{-1/2}\right\}.\end{aligned}$$

Using Eq. (6.241), we have that, with probability at least $1-\exp(-N^{c'})$,

$$\begin{aligned}\widetilde{\boldsymbol{m}}_{2\iota}(\boldsymbol{m}) &= \frac{\int_{\mathsf{Band}_*(2\iota)} \boldsymbol{\sigma}\exp(H_{N,t}(\boldsymbol{\sigma}))\;\mu_0(\mathsf{d}\boldsymbol{\sigma})}{\int_{\mathsf{Band}_*(2\iota)} \exp(H_{N,t}(\boldsymbol{\sigma}))\;\mu_0(\mathsf{d}\boldsymbol{\sigma})} \\ &= O(e^{-\eta N^{2c}}) + \int_R \frac{Z(a,b)}{Z}\langle\boldsymbol{\sigma}\rangle_{a,b}\mathsf{d}(a,b).\end{aligned}$$

Let

$$Z_T(a,b) = \int \mathbf{1}\{\boldsymbol{\sigma} \in T(a,b)\}e^{H(\boldsymbol{\sigma})}\mu_0^{a,b}(\mathsf{d}\boldsymbol{\sigma}), \qquad Z_T = \int_{\mathsf{Band}_*(2\iota)} \mathbf{1}\{\boldsymbol{\sigma} \in T(a,b)\}e^{H(\boldsymbol{\sigma})}\mu_0(\mathsf{d}\boldsymbol{\sigma}),$$

where $T(a,b) \subseteq D_N(a,b)$ is the typical set (6.120) defined for the effective model on $D_N(a,b)$. Recall from Lemma 6.7.7 that for each $(a,b) \in R$, with probability $1-e^{-cN}$,

$$Z_T(a,b) \ge (1-e^{-cN})Z(a,b), \qquad \mathbb{E}_{\ge 3}Z_T(a,b) \ge (1-e^{-cN})\mathbb{E}_{\ge 3}Z(a,b). \tag{6.242}$$

By a union bound over a $e^{-cN/10}$-net of R and standard continuity properties of H, with probability $1-e^{-cN/2}$ this holds simultaneously for all $(a,b) \in R$. By integrating, on this event we also have

$$Z_T \ge (1-e^{-cN})Z, \qquad \mathbb{E}_{\ge 3}Z_T \ge (1-e^{-cN})\mathbb{E}_{\ge 3}Z. \tag{6.243}$$

Note that, by Eq. (6.126) in Lemma 6.7.8, for $k > L \geq 1$, the following holds with probability at least $1 - e^{-cN}$,

$$\mathbb{E}_{\geq 3}\left[(Z_T(a,b) - \mathbb{E}_{\geq 3}Z_T(a,b))^{2k}\right] \leq C_L N^{-L} \left(\mathbb{E}_{\geq 3}Z_T(a,b)\right)^{2k}, \tag{6.244}$$

and therefore

$$\mathbb{E}_{\geq 3}\left[Z_T(a,b)^{2k}\right] \leq (1 + C_L N^{-L}) \left(\mathbb{E}_{\geq 3}Z_T(a,b)\right)^{2k}. \tag{6.245}$$

Again by standard continuity estimates and the union bound over a net of R of radius $e^{-c'N}$, the above estimates hold uniformly in $(a,b) \in R$ with probability at least $1 - e^{-cn/2}$. By Eq. (6.243), the same estimates hold for Z in place of Z_T uniformly in $(a,b) \in R$ with probability at least $1 - e^{-c'N}$.

By Eqs. (6.231), (6.233) of Lemma 6.7.23, together with Eqs. (6.197) and Lemma 6.7.21 (which implies that $e^{\sqrt{N}(g_{a,b} - g_{a',b'})} = O(1)$ for all $(a,b), (a',b') \in R_+$), with probability at least $1 - e^{-N^\delta}$, uniformly in $(a,b) \in R_+$,

$$\mathbb{E}_{\geq 3}Z(a,b) \geq \Omega\left(\mathbb{E}_{\geq 3}Z(a_*, b_*)\right). \tag{6.246}$$

From strict concavity of $E(a,b)$ at (a_*, b_*) (see (6.225)), and the simple estimate

$$\begin{aligned} &\sup_{a,b}\Big(N^{1/2}(|q(a-a_*)| + |b - b_*|) - \eta N(|q(a-a_*)|^2 + |b-b_*|^2)\Big) \\ &= O_\eta(1) - \eta N(|q(a-a_*)|^2 + |b-b_*|^2), \end{aligned} \tag{6.247}$$

we obtain that, uniformly in $(a,b) \in R$,

$$\mathbb{E}_{\geq 3}Z(a,b) \leq O\left(\mathbb{E}_{\geq 3}Z(a_*, b_*) \cdot e^{-\eta N(|qa - qa_*|^2 + |b-b_*|^2)/2}\right). \tag{6.248}$$

Further, by Lemma 6.7.8, we also have

$$\mathbb{P}\left\{\left|Z_T - \mathbb{E}_{\geq 3}Z_T\right| > \frac{1}{2}\,\mathbb{E}_{\geq 3}Z_T\right\} \leq CN^{-L/2}. \tag{6.249}$$

Since R_+ has volume $\Theta(N^{-1})$, on the event in (6.246) we have

$$\mathbb{E}_{\geq 3}Z \geq \int_{R_+} \mathbb{E}_{\geq 3}Z(a,b)\ \mathsf{d}(a,b) = \Omega(N^{-1}\mathbb{E}_{\geq 3}Z(a_*, b_*)).$$

Furthermore, when (6.248) holds we have

$$\mathbb{E}_{\geq 3}Z = \int_R \mathbb{E}_{\geq 3}Z(a,b)\ \mathsf{d}(a,b) \leq O(N^{-1}\mathbb{E}_{\geq 3}Z(a_*, b_*))),$$

where the N^{-1} comes from integrating the exponential in (6.248). Thus, with probability at least $1 - e^{-N^\delta}$,

$$\mathbb{E}_{\geq 3}Z = \Theta(N^{-1}\mathbb{E}_{\geq 3}Z(a_*, b_*)). \tag{6.250}$$

Let $\mathcal{E}$ denote the event that estimates (6.242), (6.244), (6.245), (6.246), (6.248), (6.249), (6.250) all hold. By the above, we have $\mathbb{P}(\mathcal{E}) \geq 1 - CN^{-L/2}$. Further, for $\hat{Z}(a,b) = e^{NE(a,b)}$ and $\hat{Z} = \int_R \hat{Z}(a,b)\mathsf{d}(a,b)$,

$$\int_R \mathbb{E}\left[\mathbf{1}_{\mathcal{E}}\left(\frac{Z(a,b)}{Z}\right)^{2k}\right]\mathsf{d}(a,b) = O\left(\int_R \mathbb{E}\left[\mathbf{1}_{\mathcal{E}}\left(\frac{\hat{Z}(a,b)}{\hat{Z}}\right)^{2k}\right]\mathsf{d}(a,b)\right) \tag{6.251}$$

Under the event $\mathcal{E}$, from Eqs. (6.246), (6.248), we thus obtain

$$\int_R \mathbb{E}\left[\mathbf{1}_{\mathcal{E}}\left(\frac{Z(a,b)}{Z}\right)^2\right]^{1/2}\mathsf{d}(a,b) = O(1), \tag{6.252}$$

By Jensen and Cauchy-Schwarz inequality,

$$\begin{aligned}
&\mathbb{E}\left[\mathbf{1}_{\mathcal{E}}\left\|\int_R \frac{Z(a,b)}{Z}(\langle\boldsymbol{\sigma}\rangle_{a,b}-\boldsymbol{\Delta}(a,b)-\boldsymbol{m}(a,b))\mathsf{d}(a,b)\right\|^{2+\delta}\right]\\
&\le \int_R \mathbb{E}\left[\mathbf{1}_{\mathcal{E}}\frac{Z(a,b)}{Z}\left\|(\langle\boldsymbol{\sigma}\rangle_{a,b}-\boldsymbol{\Delta}(a,b)-\boldsymbol{m}(a,b))\right\|^{2+\delta}\right]\mathsf{d}(a,b)\\
&\le \int_R \mathbb{E}\left[\mathbf{1}_{\mathcal{E}}\left(\frac{Z(a,b)}{Z}\right)^2\right]^{1/2}\mathbb{E}\left[\left\|(\langle\boldsymbol{\sigma}\rangle_{a,b}-\boldsymbol{\Delta}(a,b)-\boldsymbol{m}(a,b))\right\|^{2(2+\delta)}\right]^{1/2}\mathsf{d}(a,b).
\end{aligned}$$

By Lemma 6.7.2, we have

$$\mathbb{E}\left[\left\|\langle\boldsymbol{\sigma}\rangle_{a,b}-\Delta(a,b)-\boldsymbol{m}(a,b)\right\|^{2(2+\delta)}\right]\le N^{-c}. \tag{6.253}$$

Combining with Eq. (6.252), we obtain that

$$\mathbb{E}\left[\mathbf{1}_{\mathcal{E}}\left\|\int_R \frac{Z(a,b)}{Z}(\langle\boldsymbol{\sigma}\rangle_{a,b}-\Delta(a,b)-\boldsymbol{m}(a,b))\mathsf{d}(a,b)\right\|^{2+\delta}\right]\le O(N^{-c}). \tag{6.254}$$

On the other hand, by Lemma 6.7.20, with probability $1-e^{-cN}$, there are $\boldsymbol{u}_1,\boldsymbol{u}_2$ with $\|\boldsymbol{u}_1\|,\|\boldsymbol{u}_2\|=O(N^{1/2})$ such that, for $|qa-qa_*|+|b-b_*|\le N^{-1/2+c}$,

$$\|\boldsymbol{u}(a,b)-(q(a-a_*)\boldsymbol{u}_1+(b-b_*)\boldsymbol{u}_2)\|=O(N^{-1/2+2c}). \tag{6.255}$$

Using this, letting $\overline{Z}=\mathbb{E}_{\ge 3}Z(a_*,b_*)$ and defining $\overline{a}:=a-a_*$, $\overline{b}:=b-b_*$, we have

$$\begin{aligned}
&\int_R \frac{Z(a,b)}{Z}\boldsymbol{u}(a,b)\mathsf{d}(a,b)\\
&=\int_R \frac{Z(a,b)-\overline{Z}}{Z}(q\overline{a}\boldsymbol{u}_1+\overline{b}\boldsymbol{u}_2)\mathsf{d}(a,b)+\int_R\frac{\overline{Z}}{Z}(q\overline{a}\boldsymbol{u}_1+\overline{b}\boldsymbol{u}_2)\mathsf{d}(a,b)+O(N^{-1/2+2c})\\
&=\int_R \frac{Z(a,b)-\overline{Z}}{Z}(q\overline{a}\boldsymbol{u}_1+\overline{b}\boldsymbol{u}_2)\mathsf{d}(a,b)+O(N^{-1/2+2c})\\
&=\int_R \frac{Z(a,b)-\mathbb{E}_{\ge 3}Z(a,b)}{Z}(q\overline{a}\boldsymbol{u}_1+\overline{b}\boldsymbol{u}_2)\mathsf{d}(a,b)\\
&\qquad+\int_R \frac{\mathbb{E}_{\ge 3}Z(a,b)-\mathbb{E}_{\ge 3}Z(a_*,b_*)}{Z}(q\overline{a}\boldsymbol{u}_1+\overline{b}\boldsymbol{u}_2)\mathsf{d}(a,b)+O(N^{-1/2+2c}).
\end{aligned}$$

Furthermore, by Hölder's inequality on the measure $\frac{\mathbf{1}\{(a,b)\in R\}\mathsf{d}(a,b)}{\mathsf{Vol}(R)}$,

$$\begin{aligned}
&\mathbb{E}\left[\mathbf{1}_{\mathcal{E}}\left\|\int_R \frac{Z(a,b)-\mathbb{E}_{\ge 3}Z(a,b)}{Z}(q(a-a_*)\boldsymbol{u}_1+(b-b_*)\boldsymbol{u}_2)\mathsf{d}(a,b)\right\|^{2+\delta}\right]\\
&\le \int_R \mathbb{E}\left[\left(\mathbf{1}_{\mathcal{E}}\left|\frac{Z(a,b)-\mathbb{E}_{\ge 3}Z(a,b)}{Z}\right|\|q(a-a_*)\boldsymbol{u}_1+(b-b_*)\boldsymbol{u}_2\|\,\mathsf{Vol}(R)\right)^{2+\delta}\right]\frac{\mathsf{d}(a,b)}{\mathsf{Vol}(R)}
\end{aligned}$$

By (6.248) and (6.250), $Z=\Omega(N^{-1}Z_{\ge 3}(a,b))$. Moreover, we have the estimates $|q(a-a_*)|,|b-b_*|\le N^{-1/2+c}$ by definition of R, $\mathsf{Vol}(R)\le N^{-1+2c}$, and $\|\boldsymbol{u}_1\|,\|\boldsymbol{u}_2\|\le\sqrt{N}$. Combining these estimates, the last display is bounded by

$$O(N^{3c(2+\delta)})\int_R\mathbb{E}\left[\mathbf{1}_{\mathcal{E}}\left|\frac{Z(a,b)-\mathbb{E}_{\ge 3}Z(a,b)}{Z_{\ge 3}(a,b)}\right|^{2+\delta}\right]\frac{\mathsf{d}(a,b)}{\mathsf{Vol}(R)} \tag{6.256}$$

Finally, since $\mathcal{E}$ contains the event that (6.242), (6.244) holds for (a,b), for any $(a,b) \in R$ we have the estimate

$$\begin{aligned}
&\mathbb{E}\left[\mathbf{1}_{\mathcal{E}}\left|\frac{Z(a,b)-\mathbb{E}_{\geq 3}Z(a,b)}{Z_{\geq 3}(a,b)}\right|^{2+\delta}\right] \\
&\quad\leq \mathbb{E}\left[\mathbf{1}((6.242),\ (6.244)\text{ holds for }(a,b))\mathbb{E}_{\geq 3}\left|\frac{Z(a,b)-\mathbb{E}_{\geq 3}Z(a,b)}{Z_{\geq 3}(a,b)}\right|^{2+\delta}\right] \\
&\quad\leq \mathbb{E}\left[\mathbf{1}((6.242),\ (6.244)\text{ holds for }(a,b))\frac{(\mathbb{E}_{\geq 3}|Z_T(a,b)-\mathbb{E}_{\geq 3}Z_T(a,b)|^{2k})^{(2+\delta)/2k}}{|Z_{\geq 3}(a,b)|^{2+\delta}}\right]+e^{-cN},
\end{aligned}$$

and by (6.244) this is bounded by $N^{-1/2}$. Then, for c small enough, (6.256) is bounded by $O(N^{-c})$.

By Eqs. (6.233) and (6.232) of Lemma 6.7.23,

$$\left\|\int_R \frac{\mathbb{E}_{\geq 3}Z(a,b)-\mathbb{E}_{\geq 3}Z(a_*,b_*)}{Z}(q\bar{a}\boldsymbol{u}_1+\bar{b}\boldsymbol{u}_2)\mathsf{d}(a,b)\right\| = O(N^{-c}).$$

Similarly, we have

$$\mathbb{E}\left[\left\|\int_R \frac{Z(a,b)}{Z}\boldsymbol{m}(a,b)\mathsf{d}(a,b)-\int_R\frac{\mathbb{E}_{\geq 3}Z(a,b)}{Z}\boldsymbol{m}(a,b)\mathsf{d}(a,b)\right\|^{2+\delta}\right] = O(N^{-c}),$$

Again by Eqs. (6.233) and (6.232) of Lemma 6.7.23, noting that $\boldsymbol{m}(a,b)+\boldsymbol{m}(a',b') = 2\boldsymbol{m}$ if $(a,b)+(a',b') = 2(a_*,b_*)$,

$$\mathbb{E}\left[\left\|\int_R \frac{\mathbb{E}_{\geq 3}Z(a,b)}{Z}\boldsymbol{m}(a,b)\mathsf{d}(a,b)-\boldsymbol{m}\right\|^{2+\delta}\right] = O(N^{-c}),$$

Thus, we obtain

$$\mathbb{E}\left[\|\widetilde{\boldsymbol{m}}_{2\iota}(\boldsymbol{m})-\boldsymbol{m}-\Delta(\boldsymbol{m})\|^{2+\delta}\right] \leq O(N^{1+\delta/2}e^{-\eta N^c})+O(N^{-c})+O(N^{1+\delta/2}\cdot N^{-L}) = O(N^{-c}). \tag{6.257}$$

Proposition 6.4.6 then follows. □

6.8 Lognormal fluctuations of partition function

Proof of Lemma 6.3.2. Recall that $H_{N,2}$ denotes the degree-2 part of H_N, which is of the form

$$H_{N,2}(\boldsymbol{\sigma}) = \frac{\xi''(0)^{1/2}}{2}\langle \boldsymbol{G}\boldsymbol{\sigma},\boldsymbol{\sigma}\rangle,$$

for $\boldsymbol{G} \sim \mathsf{GOE}(N)$. Let

$$Z_{N,2} = \int_{S_N} \exp H_{N,2}(\boldsymbol{\sigma})\ \mathsf{d}\mu_0(\boldsymbol{\sigma}).$$

It follows from [BL16, Theorem 1.2] (with $w_2 = 2, W_4 = 3$) that, with $\sigma^2 = -\frac{1}{2}\log(1-\xi''(0))$ and $W \sim \mathcal{N}(-\frac{1}{2}\sigma^2,\sigma^2)$,

$$\frac{Z_{N,2}}{\mathbb{E}\,Z_{N,2}} = \frac{Z_{N,2}}{\exp(N\xi''(0)/2)} \xrightarrow{d} \exp(W). \tag{6.258}$$

Recall that the results in Section 6.7.1 only assume (6.31) rather than (6.5), and thus apply in the present proof. Let $\delta > 0$ be small and $T = T(\delta)$ as in (6.120), and recall the restricted partition function

$$Z_N(T) = \int_T \exp H_N(\boldsymbol{\sigma})\ \mathsf{d}\mu_0(\boldsymbol{\sigma}).$$

By (6.125), in Lemma 6.7.7, we have

$$\mathbb{E}[Z_N - Z_N(T)] \le e^{-cN}\,\mathbb{E}[Z_N].$$

By Markov's inequality, applied respectively to the randomness of H_N and $H_{N,2}$, with probability $1-e^{-cN}$,

$$(Z_N - Z_N(T)) \vee \mathbb{E}_{\ge 3}[Z_N - Z_N(T)] \le e^{-cN}\,\mathbb{E}[Z_N].$$

By (6.158) (for $k=2$), we also have, with probability $1-o(1)$,

$$\mathbb{E}_{\ge 3}\left[(Z_N(T) - \mathbb{E}_{\ge 3}[Z_N(T)])^2\right] = o(1)\,\mathbb{E}[Z_N]^2.$$

Thus with probability $1-o(1)$,

$$|Z_N(T) - \mathbb{E}_{\ge 3} Z_N(T)| \le o(1)\,\mathbb{E}[Z_N].$$

On the intersection of these events,

$$\left|\frac{Z_N}{\mathbb{E}\,Z_N} - \frac{\mathbb{E}_{\ge 3}\,Z_N}{\mathbb{E}\,Z_N}\right| \le \frac{|Z_N - Z_N(T)|}{\mathbb{E}\,Z_N} + \frac{|Z_N(T) - \mathbb{E}_{\ge 3}\,Z_N(T)|}{\mathbb{E}\,Z_N} + \frac{|\,\mathbb{E}_{\ge 3}\,Z_N - \mathbb{E}_{\ge 3}\,Z_N(T)|}{\mathbb{E}\,Z_N} = o(1).$$

Since

$$\frac{\mathbb{E}_{\ge 3}\,Z_N}{\mathbb{E}\,Z_N} = \frac{Z_{N,2}}{\mathbb{E}\,Z_{N,2}},$$

the result follows from (6.258). $\square$

6.9 Completing the proof of Theorem 6.2.1

The following two propositions are the final ingredients in the proof of Theorem 6.2.1. Let δ, L be as in Algorithm 2 and $T' = \delta L$.

Proposition 6.9.1. *Let $(H_N, \boldsymbol{y}_{T'})$ be sampled from the marginal of the planted distribution $\mathbb{P}$ (as defined in Eq. (6.33)). Let $\boldsymbol{y}^L$ be generated as in Algorithm 2, run on input H_N. Then,*

$$\mathbb{E}_{H_N}\,\mathrm{TV}\left(\mathcal{L}(\boldsymbol{y}_{T'}|H_N), \mathcal{L}(\boldsymbol{y}^L|H_N)\right) = o_N(1).$$

Proposition 6.9.2. *Let $(H_N, \boldsymbol{\sigma}, \boldsymbol{y}_{T'})$ be sampled from the marginal of the planted distribution $\mathbb{P}$. Let $\boldsymbol{\rho}^{\mathsf{MALA}}$ be the (random) output of* MALA *run on $\widetilde{\nu}^{\mathsf{proj}}_{H_N,\boldsymbol{y}_{T'}}$ (recall Eq. (6.20)) and $\widehat{\boldsymbol{\sigma}} = \boldsymbol{\sigma}_{\boldsymbol{y}_{T'}}(\boldsymbol{\rho}^{\mathsf{MALA}})$ (recall Eq. (6.16)). Then,*

$$\mathbb{E}_{H_N,\boldsymbol{y}_{T'}}\,\mathrm{TV}\left(\mathcal{L}(\boldsymbol{\sigma}|H_N,\boldsymbol{y}_{T'}), \mathcal{L}(\widehat{\boldsymbol{\sigma}}|H_N,\boldsymbol{y}_{T'})\right) = o_N(1).$$

Proof of Theorem 6.2.1. Let $\mathcal{K} : \mathscr{H}_N \times \mathbb{R}^N \to \mathbb{R}^N$ be the random map that, given input $(H_N, \boldsymbol{y})$, generates $\boldsymbol{\rho}^{\mathsf{MALA}}$ by running MALA on $\nu_{H_N,\boldsymbol{y}}$ and outputs $\widehat{\boldsymbol{\sigma}} = \boldsymbol{\sigma}_{\boldsymbol{y}}(\boldsymbol{\rho}^{\mathsf{MALA}})$. Let $(H_N, \boldsymbol{\sigma}, \boldsymbol{y}_{T'})$ be sampled from the marginal of $\mathbb{P}$. Let $\mathbb{P}_{\mathsf{alg},H_N}$ denote the law of the output of $\boldsymbol{y}^L$ generated by Algorithm 2 on input H_N. Then,

$$\begin{aligned}
\mathbb{E}_{H_N\sim\mathbb{P}}\mathrm{TV}(\mu_{H_N}, \mu^{\mathsf{alg}}) &= \mathbb{E}_{H_N\sim\mathbb{P}}\mathrm{TV}\left(\mathbb{E}_{\boldsymbol{y}_{T'}\sim\mathbb{P}(\cdot|H_N)}\mathcal{L}(\boldsymbol{\sigma}|H_N,\boldsymbol{y}_{T'}), \mathbb{E}_{\boldsymbol{y}^L\sim\mathbb{P}_{\mathsf{alg},H_N}}\mathcal{L}(\mathcal{K}(H_N,\boldsymbol{y}^L))\right) \\
&\le \mathbb{E}_{H_N\sim\mathbb{P}}\mathrm{TV}\left(\mathbb{E}_{\boldsymbol{y}_{T'}\sim\mathbb{P}(\cdot|H_N)}\mathcal{L}(\boldsymbol{\sigma}|H_N,\boldsymbol{y}_{T'}), \mathbb{E}_{\boldsymbol{y}_{T'}\sim\mathbb{P}(\cdot|H_N)}\mathcal{L}(\mathcal{K}(H_N,\boldsymbol{y}_{T'}))\right) \\
&\quad + \mathbb{E}_{H_N\sim\mathbb{P}}\mathrm{TV}\left(\mathbb{E}_{\boldsymbol{y}_{T'}\sim\mathbb{P}(\cdot|H_N)}\mathcal{L}(\mathcal{K}(H_N,\boldsymbol{y}_{T'})), \mathbb{E}_{\boldsymbol{y}^L\sim\mathbb{P}_{\mathsf{alg},H_N}}\mathcal{L}(\mathcal{K}(H_N,\boldsymbol{y}^L))\right) \\
&\le \mathbb{E}_{(H_N,\boldsymbol{y}_{T'})\sim\mathbb{P}}\mathrm{TV}\left(\mathcal{L}(\boldsymbol{\sigma}|H_N,\boldsymbol{y}_{T'}), \mathcal{L}(\mathcal{K}(H_N,\boldsymbol{y}_{T'}))\right) \\
&\quad + \mathbb{E}_{H_N\sim\mathbb{P}}\mathrm{TV}\left(\mathcal{L}(\boldsymbol{y}_{T'}|H_N), \mathcal{L}(\boldsymbol{y}^L|H_N)\right).
\end{aligned}$$

The last inequality is by data processing. By Propositions 6.9.1 and 6.9.2, the final bound is $o_N(1)$. Thus, with probability $1-o_N(1)$ over $H_N\sim\mathbb{P}$, $\mathrm{TV}(\mu_{H_N}, \mu^{\mathsf{alg}}) = o_N(1)$. By Corollary 6.3.5, the same is true for $H_N\sim\mathbb{Q}$. $\square$

6.9.1 TV-closeness of Euler discretization: Proof of Proposition 6.9.1

We prove Proposition 6.9.1 by an application of Girsanov's theorem, an approach introduced [CCL+23] in a related context. For all $0 \le \ell \le L-1$, define

$$\widehat{\boldsymbol{m}}(\boldsymbol{y}, \ell\delta) = \boldsymbol{m}^{\mathsf{alg}}(H_N, \boldsymbol{y}, \ell\delta)$$

to be the output of Algorithm 1 with these inputs. Then, define the process $(\widehat{\boldsymbol{y}}_t)_{t\in[0,T]}$ by $\widehat{\boldsymbol{y}}_0 = \boldsymbol{0}$ and, for $t \in [\ell\delta, (\ell+1)\delta)$,

$$\mathsf{d}\widehat{\boldsymbol{y}}_t = \widehat{\boldsymbol{m}}(\widehat{\boldsymbol{y}}_{\ell\delta}, \ell\delta)\ \mathsf{d}t + \mathsf{d}\boldsymbol{B}_t. \tag{6.259}$$

On each interval $[\ell\delta, (\ell+1)\delta)$, the drift in (6.259) is constant, so this SDE can be integrated directly: conditional on $H_N, \widehat{\boldsymbol{y}}_{\ell\delta}$,

$$\widehat{\boldsymbol{y}}_{(\ell+1)\delta} = \widehat{\boldsymbol{y}}_{\ell\delta} + \delta \boldsymbol{m}^{\mathsf{alg}}(H_N, \widehat{\boldsymbol{y}}_{\ell\delta}, \ell\delta) + \boldsymbol{B}_{(\ell+1)\delta} - \boldsymbol{B}_{\ell\delta}.$$

Note that $\boldsymbol{B}_{(\ell+1)\delta} - \boldsymbol{B}_{\ell\delta} =_d \sqrt{\delta}\boldsymbol{w}^\ell$ for $\boldsymbol{w}^\ell \sim \mathcal{N}(0, \boldsymbol{I}_N)$, so this is precisely the Euler discretization in Algorithm 2. It follows that

$$\mathcal{L}(\widehat{\boldsymbol{y}}_T | H_N) = \mathcal{L}(\boldsymbol{y}^L | H_N). \tag{6.260}$$

Lemma 6.9.3. *Given H_N, let $(\boldsymbol{y}_t)_{t\in[0,T]}$ be sampled from (6.3) and $(\widehat{\boldsymbol{y}}_t)_{t\in[0,T]}$ be sampled from (6.259). Then,*

$$\mathbb{E}_{H_N\sim\mathbb{P}} \mathsf{KL}(\mathcal{L}(\boldsymbol{y}_T|H_N), \mathcal{L}(\widehat{\boldsymbol{y}}_T|H_N)) \le \frac{1}{2}\sum_{\ell=0}^{L-1} \int_{\ell\delta}^{(\ell+1)\delta} \mathbb{E}_{\mathbb{P}} \|\widehat{\boldsymbol{m}}(\boldsymbol{y}_{\ell\delta}, \ell\delta) - \boldsymbol{m}(\boldsymbol{y}_t, t)\|^2\ \mathsf{d}t.$$

Proof. Fix any realization of H_N. For $0 \le \ell \le L-1$ and $t \in [\ell\delta, (\ell+1)\delta)$, define the process

$$\boldsymbol{b}_t = \widehat{\boldsymbol{m}}(\boldsymbol{y}_{\ell\delta}, \ell\delta) - \boldsymbol{m}(\boldsymbol{y}_t, t).$$

Let

$$\mathcal{E}_t = \exp\left(\int_0^t \langle \boldsymbol{b}_s, \mathsf{d}\boldsymbol{B}_s\rangle - \frac{1}{2}\int_0^t \|\boldsymbol{b}_s\|^2\ \mathsf{d}s\right).$$

Let Q be the probability measure (conditional on H_N) under which $(\boldsymbol{B}_t)_{t\in[0,T]}$ is a Brownian motion and let P be the probability measure with $\frac{\mathsf{d}P}{\mathsf{d}Q} = \mathcal{E}_T$. By Girsanov's theorem [LG16, Theorem 5.22],

$$\boldsymbol{\beta}_t = \boldsymbol{B}_t - \int_0^t \boldsymbol{b}_s\ \mathsf{d}s$$

is a Brownian motion under P. (Since $\|\widehat{\boldsymbol{m}}(\boldsymbol{y}_{\ell\delta}, \ell\delta)\|, \|\boldsymbol{m}(\boldsymbol{y}_t, t)\| \le \sqrt{N}$, $\boldsymbol{b}_t$ is a.s. bounded, and thus the conditions of Girsanov's theorem are satisfied.) The SDE (6.3) rearranges as

$$\mathsf{d}\boldsymbol{y}_t = (\boldsymbol{m}(\boldsymbol{y}_t, t) + \boldsymbol{b}_t)\ \mathsf{d}t + \mathsf{d}\boldsymbol{\beta}_t = \widehat{\boldsymbol{m}}(\boldsymbol{y}_{\ell\delta}, \ell\delta)\ \mathsf{d}t + \mathsf{d}\boldsymbol{\beta}_t, \qquad t \in [\ell\delta, (\ell+1)\delta).$$

Thus, under P, the law of $(\boldsymbol{y}_t)_{t\in[0,T]}$ is that of $(\widehat{\boldsymbol{y}}_t)_{t\in[0,T]}$. By data processing,

$$\mathsf{KL}(\mathcal{L}(\boldsymbol{y}_T|H_N), \mathcal{L}(\widehat{\boldsymbol{y}}_T|H_N)) \le \mathsf{KL}(Q, P) = \mathbb{E}_Q \log \frac{\mathsf{d}Q}{\mathsf{d}P} = \frac{1}{2}\int_0^T \mathbb{E}_Q \|\boldsymbol{b}_t\|^2\ \mathsf{d}t.$$

The result follows by taking expectation over H_N. □

Lemma 6.9.4. *For all $0 \le \ell \le L-1$, $t \in [\ell\delta, (\ell+1)\delta)$, we have $\mathbb{E}_{\mathbb{P}}\|\widehat{\boldsymbol{m}}(\boldsymbol{y}_{\ell\delta}, \ell\delta) - \boldsymbol{m}(\boldsymbol{y}_t, t)\|^2 = o_N(1)$.*

Proof. We first estimate

$$\mathbb{E}_{\mathbb{P}}\|\widehat{\boldsymbol{m}}(\boldsymbol{y}_{\ell\delta}, \ell\delta) - \boldsymbol{m}(\boldsymbol{y}_t, t)\|^2 \le 2\mathbb{E}_{\mathbb{P}}\|\widehat{\boldsymbol{m}}(\boldsymbol{y}_{\ell\delta}, \ell\delta) - \boldsymbol{m}(\boldsymbol{y}_{\ell\delta}, \ell\delta)\|^2 + 2\mathbb{E}_{\mathbb{P}}\|\boldsymbol{m}(\boldsymbol{y}_{\ell\delta}, \ell\delta) - \boldsymbol{m}(\boldsymbol{y}_t, t)\|^2.$$

The first term on the right-hand side is $o_N(1)$ by Theorem 6.4.1, so it suffices to bound the second term. Recall that for $(H_N, \boldsymbol{x}, (\boldsymbol{y}_t)_{t\in[0,T]}) \sim \mathbb{P}$, conditional on $(H_N, \boldsymbol{y}_t)$ the posterior law on $\boldsymbol{x}$ is $\mu_t(\boldsymbol{\sigma}) \propto e^{H_{N,t}(\boldsymbol{\sigma})}$, for $H_{N,t}(\boldsymbol{\sigma})$ as in (6.38). Furthermore, for $s = t - \ell\delta$, $\boldsymbol{g} \sim \mathcal{N}(0, I_N)$,

$$H_{N,t}(\boldsymbol{\sigma}) = H_{N,\ell\delta}(\boldsymbol{\sigma}) + \langle s\boldsymbol{x} + \sqrt{s}\boldsymbol{g}, \boldsymbol{\sigma}\rangle.$$

Let $\boldsymbol{\Delta}_{t,\ell\delta}(\boldsymbol{\sigma}) = H_{N,t}(\boldsymbol{\sigma}) - H_{N,\ell\delta}(\boldsymbol{\sigma})$. With probability $1 - e^{-cN}$, $\|\boldsymbol{g}\| \le 2\sqrt{N}$. Let $\mathcal{E}$ denote this event. On $\mathcal{E}$,

$$\sup_{\boldsymbol{\sigma}\in S_N} \|\boldsymbol{\Delta}_{t,\ell\delta}(\boldsymbol{\sigma})\| \le \delta\sqrt{N}\|\boldsymbol{x}\| + \sqrt{\delta N}\|\boldsymbol{g}\| \le 3\sqrt{\delta}N = 3/N. \tag{6.261}$$

So,

$$\begin{aligned}
\boldsymbol{m}(\boldsymbol{y}_{\ell\delta}, \ell\delta) - \boldsymbol{m}(\boldsymbol{y}_t, t) &= \frac{\int_{S_N} \boldsymbol{\sigma} e^{H_{N,\ell\delta}(\boldsymbol{\sigma})}}{\int_{S_N} e^{H_{N,\ell\delta}(\boldsymbol{\sigma})}} - \frac{\int_{S_N} \boldsymbol{\sigma} e^{H_{N,t}(\boldsymbol{\sigma})}}{\int_{S_N} e^{H_{N,t}(\boldsymbol{\sigma})}} \\
&= \frac{\iint \boldsymbol{\sigma}^1 (e^{H_{N,\ell\delta}(\boldsymbol{\sigma}^1)+H_{N,t}(\boldsymbol{\sigma}^2)} - e^{H_{N,\ell\delta}(\boldsymbol{\sigma}^2)+H_{N,t}(\boldsymbol{\sigma}^1)})\ \mu_0^{\otimes 2}(\mathsf{d}\boldsymbol{\sigma})}{\iint e^{H_{N,\ell\delta}(\boldsymbol{\sigma}^1)+H_{N,t}(\boldsymbol{\sigma}^2)}\ \mu_0^{\otimes 2}(\mathsf{d}\boldsymbol{\sigma})} \\
&= \frac{\iint \boldsymbol{\sigma}^1 (e^{\boldsymbol{\Delta}_{t,\ell\delta}(\boldsymbol{\sigma}^1)} - e^{\boldsymbol{\Delta}_{t,\ell\delta}(\boldsymbol{\sigma}^2)}) e^{H_{N,\ell\delta}(\boldsymbol{\sigma}^1)+H_{N,\ell\delta}(\boldsymbol{\sigma}^2)}\ \mu_0^{\otimes 2}(\mathsf{d}\boldsymbol{\sigma})}{\iint e^{\boldsymbol{\Delta}_{t,\ell\delta}(\boldsymbol{\sigma}^2)} e^{H_{N,\ell\delta}(\boldsymbol{\sigma}^1)+H_{N,\ell\delta}(\boldsymbol{\sigma}^2)}\ \mu_0^{\otimes 2}(\mathsf{d}\boldsymbol{\sigma})}.
\end{aligned}$$

By (6.261),

$$\|\boldsymbol{\sigma}^1\| |e^{\boldsymbol{\Delta}_{t,\ell\delta}(\boldsymbol{\sigma}^1)} - e^{\boldsymbol{\Delta}_{t,\ell\delta}(\boldsymbol{\sigma}^2)}| = O(N^{-1/2})$$

for all $\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2 \in S_N$, and thus $\|\boldsymbol{m}(\boldsymbol{y}_{\ell\delta}, \ell\delta) - \boldsymbol{m}(\boldsymbol{y}_t, t)\| = O(N^{-1/2})$. So,

$$\begin{aligned}
\mathbb{E}_{\mathbb{P}} \|\boldsymbol{m}(\boldsymbol{y}_{\ell\delta}, \ell\delta) - \boldsymbol{m}(\boldsymbol{y}_t, t)\|^2 &\le \mathbb{E}_{\mathbb{P}} \mathbf{1}\{\mathcal{E}\} \|\boldsymbol{m}(\boldsymbol{y}_{\ell\delta}, \ell\delta) - \boldsymbol{m}(\boldsymbol{y}_t, t)\|^2 + \mathbb{E}_{\mathbb{P}} \mathbf{1}\{\mathcal{E}^c\} \|\boldsymbol{m}(\boldsymbol{y}_{\ell\delta}, \ell\delta) - \boldsymbol{m}(\boldsymbol{y}_t, t)\|^2 \\
&\le O(N^{-1/2}) + e^{-cN} \cdot 4N = o_N(1).
\end{aligned}$$

□

Proof of Proposition 6.9.1. By (6.260) and Lemmas 6.9.3 and 6.9.4,

$$\mathbb{E}_{H_N\sim\mathbb{P}} \mathsf{KL}(\mathcal{L}(\boldsymbol{y}_T|H_N), \mathcal{L}(\boldsymbol{y}^L|H_N)) = o_N(1).$$

The result follows from Pinsker's inequality and Jensen's inequality:

$$\begin{aligned}
\mathbb{E}_{H_N\sim\mathbb{P}} \mathsf{KL}(\mathcal{L}(\boldsymbol{y}_T|H_N), \mathcal{L}(\boldsymbol{y}^L|H_N)) &\ge 2\mathbb{E}_{H_N\sim\mathbb{P}} \left[\mathrm{TV}(\mathcal{L}(\boldsymbol{y}_T|H_N), \mathcal{L}(\boldsymbol{y}^L|H_N))^2\right] \\
&\ge 2\left[\mathbb{E}_{H_N\sim\mathbb{P}} \mathrm{TV}(\mathcal{L}(\boldsymbol{y}_T|H_N), \mathcal{L}(\boldsymbol{y}^L|H_N))\right]^2.
\end{aligned}$$

□

6.9.2 Log-concavity of late measures

In this subsection, we prove Proposition 6.9.2. Let $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_N$ be the standard basis. By a change of coordinates, we may assume without loss of generality that $\widehat{\boldsymbol{y}} = \boldsymbol{y}/\|\boldsymbol{y}\|_N = \boldsymbol{e}_N\sqrt{N}$ and $\boldsymbol{U} = (\boldsymbol{e}_1, \ldots, \boldsymbol{e}_{N-1})$.

Lemma 6.9.5. *For any $\boldsymbol{y} \neq \mathbf{0}$, the push-forward of $\mu_{H_N,\boldsymbol{y}}(\,\cdot\,|\langle\boldsymbol{\sigma}, \boldsymbol{y}\rangle > 0)$ under the stereographic projection $\boldsymbol{T}_{\boldsymbol{y}}$ is $\nu^{\mathsf{proj}}_{H_N,\boldsymbol{y}}$, defined in (6.18).*

Proof. Note that (denoting by $\boldsymbol{D}F$ the Jacobian of map F):

$$\boldsymbol{D}\boldsymbol{\sigma}_{\boldsymbol{y}}(\boldsymbol{\rho})^\top = \frac{[\boldsymbol{I}_{N-1}, \mathbf{0}]}{\sqrt{1 + \|\boldsymbol{\rho}\|_N^2}} - \frac{\boldsymbol{\rho}\boldsymbol{\sigma}_{\boldsymbol{y}}(\boldsymbol{\rho})^\top/N}{1 + \|\boldsymbol{\rho}\|_N^2}.$$

Since $[\boldsymbol{I}_{N-1}, \boldsymbol{0}]\boldsymbol{\sigma}_{\boldsymbol{y}}(\boldsymbol{\rho}) = \frac{\boldsymbol{\rho}}{\sqrt{1+\|\boldsymbol{\rho}\|_N^2}}$, we have

$$\boldsymbol{D}\boldsymbol{\sigma}_{\boldsymbol{y}}(\boldsymbol{\rho})^\top \boldsymbol{D}\boldsymbol{\sigma}_{\boldsymbol{y}}(\boldsymbol{\rho}) = \frac{\boldsymbol{I}_{N-1}}{1+\|\boldsymbol{\rho}\|_N^2} - \frac{\boldsymbol{\rho}\boldsymbol{\rho}^\top/N}{(1+\|\boldsymbol{\rho}\|_N^2)^2} = \frac{\boldsymbol{I}_{N-1}}{1+\|\boldsymbol{\rho}\|_N^2}\left(\boldsymbol{I}_{N-1} - \frac{\boldsymbol{\rho}\boldsymbol{\rho}^\top/N}{1+\|\boldsymbol{\rho}\|_N^2}\right).$$

The stereographic projection thus incurs a change of density factor of

$$\det(\boldsymbol{D}\boldsymbol{\sigma}_{\boldsymbol{y}}(\boldsymbol{\rho})^\top \boldsymbol{D}\boldsymbol{\sigma}_{\boldsymbol{y}}(\boldsymbol{\rho}))^{1/2} = (1+\|\boldsymbol{\rho}\|_N^2)^{-N/2}.$$

This precisely accounts for the term $-\frac{N}{2}\log(1+\|\boldsymbol{\rho}\|_N^2)$ in (6.17). ☐

Lemma 6.9.6. *For sufficiently large T, with probability $1-o_N(1)$ over $(H_N, \boldsymbol{y}_T)$ as in Proposition 6.9.2, $\nu^{\mathsf{proj}}_{H_N,\boldsymbol{y}_T}(\|\boldsymbol{\rho}\|_N^2 \le \varepsilon_0) = 1-o_N(1)$ and $\mu_{H_N,\boldsymbol{y}_T}(\langle \boldsymbol{\sigma}, \boldsymbol{y}_T\rangle_N \le 0) = o_N(1)$.*

Proof. Let $(H_N, \boldsymbol{x}, \boldsymbol{y}_T)$ be a sample from $\mathbb{P}$, and let $q_* = q_*(T)$ be as in Fact 6.4.2. Note that $q_* > 1-\frac{1}{T}$, as

$$\xi_T'\big(1-1/T\big) \ge T + \xi_T'\big(1-1/T\big) \ge T > T-1 = \frac{1-1/T}{1/T}.$$

By Proposition 6.5.12, with probability $1-o_N(1)$,

$$\mu_{H_N,\boldsymbol{y}_T}(\langle \boldsymbol{\sigma}, \boldsymbol{x}\rangle_N \ge 1-1/T) = 1-o_N(1).$$

With probability $1-o_N(1)$, we have $\|\boldsymbol{y}\|_N = \sqrt{T(T+1)} + o_N(1)$, so

$$\langle \boldsymbol{x}, \widehat{\boldsymbol{y}}\rangle = \frac{\langle \boldsymbol{x}, \boldsymbol{y}\rangle}{\|\boldsymbol{y}\|_N} = \sqrt{1-\frac{1}{T+1}} + o_N(1).$$

On this event, $\{\boldsymbol{\sigma} \in S_N : \langle \boldsymbol{\sigma}, \boldsymbol{x}\rangle_N \ge 1-1/T\} \subseteq \{\boldsymbol{\sigma} \in S_N : \langle \boldsymbol{\sigma}, \widehat{\boldsymbol{y}}\rangle_N \ge 1-2/T\}$. So, with probability $1-o_N(1)$,

$$\mu_{H_N,\boldsymbol{y}_T}(\langle \boldsymbol{\sigma}, \widehat{\boldsymbol{y}}\rangle_N \ge 1-2/T) = 1-o_N(1).$$

(This of course implies $\mu_{H_N,\boldsymbol{y}_T}(\langle \boldsymbol{\sigma}, \boldsymbol{y}_T\rangle_N \le 0) = o_N(1)$.) For sufficiently large T, the stereographic projection $\boldsymbol{T}_{\boldsymbol{y}}$ maps $\{\boldsymbol{\sigma} \in S_N : \langle \boldsymbol{\sigma}, \widehat{\boldsymbol{y}}\rangle_N \ge 1-2/T\}$ into $\{\boldsymbol{\rho} \in \mathbb{R}^{N-1} : \|\boldsymbol{\rho}\|_N^2 \le \varepsilon_0\}$. The conclusion follows from Lemma 6.9.5. ☐

Corollary 6.9.7. *Recall definition (6.20) of $\nu^{\mathsf{proj}}_{H_N,\boldsymbol{y}_T}$, $\widetilde{\nu}^{\mathsf{proj}}_{H_N,\boldsymbol{y}_T}$. For sufficiently large T, with probability $1-o_N(1)$ over $(H_N, \boldsymbol{y}_T)$, $\mathrm{TV}(\nu^{\mathsf{proj}}_{H_N,\boldsymbol{y}_T}, \widetilde{\nu}^{\mathsf{proj}}_{H_N,\boldsymbol{y}_T}) = o_N(1)$.*

Proof. Since $\varphi(x) = 0$ for $x \in [0, \varepsilon_0]$, and $\varphi(x) \ge 0$ for $x > \varepsilon_0$, we have

$$\int_{\|\boldsymbol{\rho}\|_N^2 \le \varepsilon_0} \exp \widetilde{H}^{\mathsf{proj}}_{N,\boldsymbol{y}_T}(\boldsymbol{\rho})\ \mathsf{d}\boldsymbol{\rho} = \int_{\|\boldsymbol{\rho}\|_N^2 \le \varepsilon_0} \exp H^{\mathsf{proj}}_{N,\boldsymbol{y}_T}(\boldsymbol{\rho})\ \mathsf{d}\boldsymbol{\rho},$$
$$\int_{\|\boldsymbol{\rho}\|_N^2 > \varepsilon_0} \exp \widetilde{H}^{\mathsf{proj}}_{N,\boldsymbol{y}_T}(\boldsymbol{\rho})\ \mathsf{d}\boldsymbol{\rho} \le \int_{\|\boldsymbol{\rho}\|_N^2 > \varepsilon_0} \exp H^{\mathsf{proj}}_{N,\boldsymbol{y}_T}(\boldsymbol{\rho})\ \mathsf{d}\boldsymbol{\rho}.$$

Combined with Lemma 6.9.6, it follows that with probability $1-o_N(1)$,

$$\widetilde{\nu}^{\mathsf{proj}}_{H_N,\boldsymbol{y}_T}(\|\boldsymbol{\rho}\|_N^2 \le \varepsilon_0) \ge \nu^{\mathsf{proj}}_{H_N,\boldsymbol{y}_T}(\|\boldsymbol{\rho}\|_N^2 \le \varepsilon_0) \ge 1-o_N(1).$$

Since $\widetilde{\nu}^{\mathsf{proj}}_{H_N,\boldsymbol{y}_T}$ and $\nu^{\mathsf{proj}}_{H_N,\boldsymbol{y}_T}$ are furthermore proportional on $\{\|\boldsymbol{\rho}\|_N^2 \le \varepsilon_0\}$, the conclusion follows. ☐

Proposition 6.9.8. *For sufficiently large T, there exist $C_{\min}, C_{\max} > 0$ (depending on T) such that with probability $1-o_N(1)$, for all $\boldsymbol{\rho} \in \mathbb{R}^{N-1}$,*

$$-C_{\max}\boldsymbol{I}_{N-1} \preceq \nabla^2 \widetilde{H}^{\mathsf{proj}}_{N,\boldsymbol{y}_T}(\boldsymbol{\rho}) \preceq -C_{\min}\boldsymbol{I}_{N-1}.$$

Proof. Let $\boldsymbol{y} = \boldsymbol{y}_T$ $\widehat{\boldsymbol{y}} = \boldsymbol{y}/\|\boldsymbol{y}\|_N$, and assume without loss of generality $\widehat{\boldsymbol{y}} = \sqrt{N}\,\boldsymbol{e}_N$. Let $\boldsymbol{U}^\top = [\boldsymbol{I}_{N-1}, \boldsymbol{0}] \in \mathbb{R}^{(N-1)\times N}$ be the projection onto the orthogonal complement of $\widehat{\boldsymbol{y}}$.

A direct calculation shows

$$\begin{aligned}
\nabla^2 \widetilde{H}^{\mathsf{proj}}_{N,\boldsymbol{y}}(\boldsymbol{\rho}) &= \frac{\langle \nabla H_{N,\boldsymbol{y}}(\boldsymbol{\sigma}_{\boldsymbol{y}}(\boldsymbol{\rho})), \boldsymbol{\sigma}_{\boldsymbol{y}}(\boldsymbol{\rho})\rangle}{N(1+\|\boldsymbol{\rho}\|_N^2)} \left(-\boldsymbol{I}_{N-1} + \frac{3\boldsymbol{\rho}\boldsymbol{\rho}^\top}{N(1+\|\boldsymbol{\rho}\|_N^2)}\right) \\
&+ \frac{\langle \nabla^2 H_{N,\boldsymbol{y}}(\boldsymbol{\sigma}_{\boldsymbol{y}}(\boldsymbol{\rho})), \boldsymbol{\sigma}_{\boldsymbol{y}}(\boldsymbol{\rho})^{\otimes 2}\rangle}{N(1+\|\boldsymbol{\rho}\|_N^2)^2} \cdot \frac{\boldsymbol{\rho}\boldsymbol{\rho}^\top}{N} + \frac{\boldsymbol{U}^\top \nabla^2 H_{N,\boldsymbol{y}}(\boldsymbol{\sigma}_{\boldsymbol{y}}(\boldsymbol{\rho}))\boldsymbol{U}}{1+\|\boldsymbol{\rho}\|_N^2} \\
&- \frac{\boldsymbol{\rho}\boldsymbol{\sigma}_{\boldsymbol{y}}(\boldsymbol{\rho})^\top \nabla^2 H_{N,\boldsymbol{y}}(\boldsymbol{\sigma}_{\boldsymbol{y}}(\boldsymbol{\rho}))\boldsymbol{U} + \boldsymbol{U}^\top \nabla^2 H_{N,\boldsymbol{y}}(\boldsymbol{\sigma}_{\boldsymbol{y}}(\boldsymbol{\rho}))\boldsymbol{\sigma}_{\boldsymbol{y}}(\boldsymbol{\rho})\boldsymbol{\rho}^\top}{N(1+\|\boldsymbol{\rho}\|_N^2)} \\
&- \frac{\boldsymbol{\rho}\nabla H_{N,\boldsymbol{y}}(\boldsymbol{\sigma}_{\boldsymbol{y}}(\boldsymbol{\rho}))^\top \boldsymbol{U} + \boldsymbol{U}^\top \nabla H_{N,\boldsymbol{y}}(\boldsymbol{\sigma}_{\boldsymbol{y}}(\boldsymbol{\rho}))\boldsymbol{\rho}^\top}{N(1+\|\boldsymbol{\rho}\|_N^2)^{3/2}} \\
&- \left(T\varphi'(\|\boldsymbol{\rho}\|_N^2) + \frac{1}{1+\|\boldsymbol{\rho}\|_N^2}\right)\boldsymbol{I}_{N-1} - \left(T\varphi''(\|\boldsymbol{\rho}\|_N^2) - \frac{1}{(1+\|\boldsymbol{\rho}\|_N^2)^2}\right)\frac{2\boldsymbol{\rho}\boldsymbol{\rho}^\top}{N}.
\end{aligned}$$

By Proposition 6.3.6, there exists $C > 0$ (independent of T) such that with probability $1 - o_N(1)$,

$$\sup_{\boldsymbol{\sigma}\in S_N} \|\nabla H_N(\boldsymbol{\sigma})\|_N,\ \sup_{\boldsymbol{\sigma}\in S_N} \left\|\nabla^2 H_N(\boldsymbol{\sigma})\right\|_{\mathsf{op}} \le C.$$

We will show that on this event,

$$\begin{aligned}
\nabla^2 \widetilde{H}^{\mathsf{proj}}_{N,\boldsymbol{y}}(\boldsymbol{\rho}) &= \frac{\|\boldsymbol{y}\|_N}{(1+\|\boldsymbol{\rho}\|_N^2)^{3/2}} \left(-\boldsymbol{I}_{N-1} + \frac{3\boldsymbol{\rho}\boldsymbol{\rho}^\top}{N(1+\|\boldsymbol{\rho}\|_N^2)}\right) - T\varphi'(\|\boldsymbol{\rho}\|_N^2)\boldsymbol{I}_{N-1} \\
&\quad - T\varphi''(\|\boldsymbol{\rho}\|_N^2) \cdot \frac{2\boldsymbol{\rho}\boldsymbol{\rho}^\top}{N} + O(1),
\end{aligned} \tag{6.262}$$

where $O(1)$ denotes a matrix of operator norm $O(1)$, independent of T. Note that

$$\frac{\langle \nabla H_{N,\boldsymbol{y}}(\boldsymbol{\sigma}_{\boldsymbol{y}}(\boldsymbol{\rho})), \boldsymbol{\sigma}_{\boldsymbol{y}}(\boldsymbol{\rho})\rangle}{N(1+\|\boldsymbol{\rho}\|_N^2)} = \frac{\langle \nabla H_N(\boldsymbol{\sigma}_{\boldsymbol{y}}(\boldsymbol{\rho})) + \boldsymbol{y}, \boldsymbol{\sigma}_{\boldsymbol{y}}(\boldsymbol{\rho})\rangle}{N(1+\|\boldsymbol{\rho}\|_N^2)} = \frac{\langle \nabla H_N(\boldsymbol{\sigma}_{\boldsymbol{y}}(\boldsymbol{\rho})), \boldsymbol{\sigma}_{\boldsymbol{y}}(\boldsymbol{\rho})\rangle}{N(1+\|\boldsymbol{\rho}\|_N^2)} + \frac{\|\boldsymbol{y}\|_N}{(1+\|\boldsymbol{\rho}\|_N^2)^{3/2}}.$$

The first term on the right-hand side is bounded independently of T, as

$$\frac{|\langle \nabla H_N(\boldsymbol{\sigma}_{\boldsymbol{y}}(\boldsymbol{\rho})), \boldsymbol{\sigma}_{\boldsymbol{y}}(\boldsymbol{\rho})\rangle|}{N} \le \|\nabla H_N(\boldsymbol{\sigma}_{\boldsymbol{y}}(\boldsymbol{\rho}))\|_N \|\boldsymbol{\sigma}_{\boldsymbol{y}}(\boldsymbol{\rho})\|_N.$$

Similarly, all other terms in the expansion of $\nabla^2 \widetilde{H}^{\mathsf{proj}}_{N,\boldsymbol{y}}(\boldsymbol{\rho})$ above, aside from $T\varphi'(\|\boldsymbol{\rho}\|_N^2)\boldsymbol{I}_{N-1}$ and $T\varphi''(\|\boldsymbol{\rho}\|_N^2)\cdot \frac{2\boldsymbol{\rho}\boldsymbol{\rho}^\top}{N}$, are bounded independently of T, due to the following inequalities:

$$\begin{aligned}
\left\|\nabla^2 H_{N,\boldsymbol{y}}(\boldsymbol{\sigma}_{\boldsymbol{y}}(\boldsymbol{\rho}))\right\|_{\mathsf{op}} &= \left\|\nabla^2 H_N(\boldsymbol{\sigma}_{\boldsymbol{y}}(\boldsymbol{\rho}))\right\|_{\mathsf{op}} = O(1), \\
\frac{\left\|\boldsymbol{U}^\top \nabla H_{N,\boldsymbol{y}}(\boldsymbol{\sigma}_{\boldsymbol{y}}(\boldsymbol{\rho}))\boldsymbol{\rho}^\top\right\|_{\mathsf{op}}}{N} &\le \left\|\boldsymbol{U}^\top \nabla H_{N,\boldsymbol{y}}(\boldsymbol{\sigma}_{\boldsymbol{y}}(\boldsymbol{\rho}))\right\|_N \|\boldsymbol{\rho}\|_N \\
&= \left\|\boldsymbol{U}^\top \nabla H_N(\boldsymbol{\sigma}_{\boldsymbol{y}}(\boldsymbol{\rho}))\right\|_N \|\boldsymbol{\rho}\|_N \le O(1)\|\boldsymbol{\rho}\|_N,
\end{aligned}$$

and $\left\|\boldsymbol{\rho}\boldsymbol{\rho}^\top\right\|_{\mathsf{op}}/N = \|\boldsymbol{\rho}\|_N^2$, $\left\|\boldsymbol{\rho}\boldsymbol{\sigma}_{\boldsymbol{y}}(\boldsymbol{\rho})^\top\right\|_{\mathsf{op}}/N = \|\boldsymbol{\rho}\|_N$. (Note that each of these terms, each copy of $\|\boldsymbol{\rho}\|_N^2$ in the resulting bound is compensated by at least one copy of $1+\|\boldsymbol{\rho}\|_N^2$ in the denominator.) This proves (6.262).

With probability $1 - o_N(1)$, we have $\|\boldsymbol{y}\|_N = \sqrt{T(T+1)} + o_N(1)$. On this event, (6.262) yields

$$\nabla^2 \widetilde{H}^{\mathsf{proj}}_{N,\boldsymbol{y}}(\boldsymbol{\rho}) = T(-\boldsymbol{M}(\boldsymbol{\rho}) + o_T(1)),$$

where $o_T(1)$ denotes a matrix with operator norm vanishing with T and

$$\boldsymbol{M}(\boldsymbol{\rho}) = \frac{\boldsymbol{I}_{N-1}}{(1+\|\boldsymbol{\rho}\|_N^2)^{3/2}} - \frac{3\boldsymbol{\rho}\boldsymbol{\rho}^\top}{N(1+\|\boldsymbol{\rho}\|_N^2)^{5/2}} + \varphi'(\|\boldsymbol{\rho}\|_N^2)\boldsymbol{I}_{N-1} + \varphi''(\|\boldsymbol{\rho}\|_N^2) \cdot \frac{2\boldsymbol{\rho}\boldsymbol{\rho}^\top}{N}.$$

From this it is clear that $-C_{\max}\boldsymbol{I}_{N-1} \preceq \nabla^2 \widetilde{H}^{\mathsf{proj}}_{N,\boldsymbol{y}_T}(\boldsymbol{\rho})$ for suitable $C_{\max}$. For the other direction, note that $\boldsymbol{M}(\boldsymbol{\rho})$ has eigenvalue $\frac{1}{(1+\|\boldsymbol{\rho}\|_N^2)^{3/2}} + \varphi'(\|\boldsymbol{\rho}\|_N^2)$ in all directions orthogonal to $\boldsymbol{\rho}$, and

$$\frac{1-2\|\boldsymbol{\rho}\|_N^2}{(1+\|\boldsymbol{\rho}\|_N^2)^{5/2}} + \varphi'(\|\boldsymbol{\rho}\|_N^2) + 2\|\boldsymbol{\rho}\|_N^2\varphi''(\|\boldsymbol{\rho}\|_N^2)$$

in the direction of $\boldsymbol{\rho}$. By (6.19), $\boldsymbol{M}(\boldsymbol{\rho}) \succeq \varepsilon_0\boldsymbol{I}_{N-1}$, and thus $\nabla^2\widetilde{H}^{\mathsf{proj}}_{N,\boldsymbol{y}_T}(\boldsymbol{\rho}) \preceq -C_{\min}\boldsymbol{I}_{N-1}$ for $C_{\min} = T\varepsilon_0/2$. □

Finally, we verify that φ satisfying (6.19) exists.

Fact 6.9.9. *For suitable $C > 0$, the function*

$$\varphi(x) = C\mathbf{1}\{x > \varepsilon_0\}\left(x - \frac{\varepsilon_0^2}{x} - 2\varepsilon_0\log\frac{x}{\varepsilon_0}\right)$$

is nonnegative, twice continuously differentiable, and satisfies (6.19).

Proof. Note that for $x > \varepsilon_0$,

$$\varphi'(x) = C\left(1 - \frac{\varepsilon_0}{x}\right)^2, \qquad \varphi''(x) = \frac{2C\varepsilon}{x^2}\left(1 - \frac{\varepsilon_0}{x}\right).$$

Thus $\lim_{x\downarrow\varepsilon_0}\varphi''(x) = 0$, so φ is twice continuously differentiable. Note that $\varphi' \geq 0$, so integrating shows $\varphi \geq 0$. Let

$$C_0 = \min_{x\geq 0}\frac{1-2x}{(1+x)^{5/2}}$$

and set C so that $C_0 + \varphi'(2\varepsilon_0) \geq \varepsilon_0$. Note $\varphi'' \geq 0$, and thus φ' is increasing; thus (6.19) holds for all $x \geq 2\varepsilon_0$. For all $x \in [0, 2\varepsilon_0]$, we verify that

$$\frac{1}{(1+x)^{3/2}} \geq \frac{1-2x}{(1+x)^{5/2}} \geq \frac{1-4\varepsilon_0}{(1+2\varepsilon_0)^{5/2}} \geq \varepsilon_0,$$

so (6.19) holds. □

Proof of Proposition 6.9.2. By Proposition 6.9.8, $\widetilde{\nu}^{\mathsf{proj}}_{H_N,\boldsymbol{y}_T}$ is $O(1)$-smooth and strongly log-concave. By [CLA+21, Theorem 3], MALA run for time $\chi_{\text{log-conc}} = \mathrm{poly}(N)$ outputs $\boldsymbol{\rho}^{\mathsf{MALA}} \sim \nu^{\mathsf{MALA}}$, where $\mathrm{TV}(\nu^{\mathsf{MALA}}, \widetilde{\nu}^{\mathsf{proj}}_{H_N,\boldsymbol{y}_T}) \leq 1/N$. Combined with Corollary 6.9.7, we find that (with probability $1 - o_N(1)$), $\mathrm{TV}(\nu^{\mathsf{proj}}_{H_N,\boldsymbol{y}_T}, \nu^{\mathsf{MALA}}) = o_N(1)$. Lemma 6.9.5 completes the proof. □

6.10 Failure of stochastic localization in complementary regime

In this section, we prove Theorem 6.2.3. Similarly to Subsection 6.3.2, we may analyze the process (6.27) by passing to a planted model. For any $T > 0$, let $\check{\mathbb{P}}, \check{\mathbb{Q}} \in \mathcal{P}(S_N \times \mathscr{H}_N \times C([0,T], \mathbb{R}^N \times \cdots \times (\mathbb{R}^N)^{\otimes J}))$ be the laws of $(\boldsymbol{\sigma}, H_N, (\vec{\boldsymbol{y}}_t)_{t\in[0,T]})$, generated as follows.

- Under $\check{\mathbb{Q}}$,

$$H_N \sim \mu_{\mathsf{null}}, \qquad \boldsymbol{\sigma} \sim \mu_{H_N}, \qquad \boldsymbol{y}_t^j = \tau_j(t)\boldsymbol{\sigma}^{\otimes j} + \boldsymbol{B}^j_{\tau_j(t)}, \quad \forall j = 1, \ldots, J,$$

  for $(\boldsymbol{B}^1_t, \ldots, \boldsymbol{B}^J_t)_{t\geq 0}$ independent of $(H_N, \boldsymbol{\sigma})$. Equivalently, $H_N \sim \mu_{\mathsf{null}}$, $(\vec{\boldsymbol{y}}_t)_{t\geq 0}$ is given by the SDE (6.27), and for any odd j such that $\lim_{t\to\infty}\tau_j(t) = \infty$, $\boldsymbol{\sigma}$ is the unique solution to $\boldsymbol{\sigma}^{\otimes j} = \lim_{t\to\infty}\boldsymbol{y}_t^j/\tau_j(t)$.

- Under $\check{\mathbb{P}}$,
$$(H_N, \boldsymbol{\sigma}) \sim \mu_{\mathsf{pl}}, \qquad \boldsymbol{y}_t^j = \tau_j(t)\boldsymbol{\sigma}^{\otimes j} + \boldsymbol{B}^j_{\tau_j(t)}, \quad \forall j = 1, \dots, J,$$
for $(\boldsymbol{B}_t^1, \dots, \boldsymbol{B}_t^J)_{t\geq 0}$ independent of $(H_N, \boldsymbol{\sigma})$. Equivalently, we can generate first H_N, then $(\vec{\boldsymbol{y}}_t)_{t\geq 0}$ by (6.27), and finally $\boldsymbol{\sigma}$ as above. Furthermore, the law of $(H_N, \boldsymbol{\sigma}) \sim \mu_{\mathsf{pl}}$ can be described by either (6.34) or (6.35).

Analogously to Proposition 6.3.4, we have
$$\frac{\mathsf{d}\check{\mathbb{P}}}{\mathsf{d}\check{\mathbb{Q}}}(\boldsymbol{\sigma}, H_N, (\vec{\boldsymbol{y}}_t)_{t\in[0,T]}) = \frac{Z(H_N)}{\mathbb{E}Z(H_N)},$$
and this ratio is tight by Lemma 6.3.2. Thus $\check{\mathbb{P}}$ and $\check{\mathbb{Q}}$ are mutually contiguous.

Therefore, it suffices to analyze the AMP iteration (6.30) under $\check{\mathbb{P}}$. Similarly to (6.38), we find that conditional on $\vec{\boldsymbol{y}}_t$, the posterior law of $\boldsymbol{\sigma}$ under $\check{\mathbb{P}}$ is
$$\check{\mu}_t(\mathsf{d}\boldsymbol{\sigma}) = \frac{1}{Z}\exp \check{H}_{N,t}(\boldsymbol{\sigma})\mu_0(\mathsf{d}\boldsymbol{\sigma}),$$
where
$$\check{H}_{N,t}(\boldsymbol{\sigma}) = N\xi(\langle \boldsymbol{x}, \boldsymbol{\sigma}\rangle_N) + \widetilde{H}_N(\boldsymbol{\sigma}) + \sum_{j=1}^J \frac{1}{N^{j-1}}\langle \boldsymbol{y}_t^j, \boldsymbol{\sigma}^{\otimes j}\rangle \stackrel{d}{=} N\check{\xi}_t(\langle \boldsymbol{x}, \boldsymbol{\sigma}\rangle_N) + \widetilde{H}_{N,t}(\boldsymbol{\sigma}),$$
for $\widetilde{H}_{N,t}$ a spin glass with mixture
$$\check{\xi}_t(q) = \xi(q) + \sum_{j=1}^J \tau_j(t)^2 q^j\ .$$
Let $q_{\mathsf{AMP}} = q_{\mathsf{AMP}}(t)$ be the smallest solution to $\check{\xi}_t'(q) = \frac{q}{1-q}$. Note that a solution exists because $\check{\xi}_t'(0) \geq 0$ and $\lim_{q\uparrow 1} \frac{q}{1-q} = +\infty$.

Proposition 6.10.1. *We have*
$$\lim_{k\to\infty} \operatorname*{p-lim}_{N\to\infty} \langle \boldsymbol{x}, \check{\boldsymbol{m}}^k\rangle_N = \lim_{k\to\infty} \operatorname*{p-lim}_{N\to\infty} \langle \check{\boldsymbol{m}}^k, \check{\boldsymbol{m}}^k\rangle_N = q_{\mathsf{AMP}}.$$
Consequently, for all $1 \leq j \leq J$,
$$\lim_{k\to\infty} \lim_{N\to\infty} \mathbb{E}\frac{1}{N^j}\left\|\boldsymbol{x}^{\otimes j} - (\check{\boldsymbol{m}}^k)^{\otimes j}\right\|_2^2 = 1 - q_{\mathsf{AMP}}^j.$$

Proof. Since $q \mapsto \frac{\check{\xi}_t'(q)}{1+\check{\xi}_t'(q)}$ is increasing, the sequence $(\check{q}_k)_{k\geq 0}$ defined in (6.29) is increasing. Furthermore, if $\check{q}_k \leq q_{\mathsf{AMP}}$, then
$$\check{q}_{k+1} = \frac{\check{\xi}_t'(\check{q}_k)}{1 + \check{\xi}_t'(\check{q}_k)} \leq \frac{\check{\xi}_t'(q_{\mathsf{AMP}})}{1 + \check{\xi}_t'(q_{\mathsf{AMP}})} = q_{\mathsf{AMP}},$$
and therefore by induction $(\check{q}_k)_{k\geq 0}$ is bounded above by q_{AMP}. As the limit of $(\check{q}_k)_{k\geq 0}$ must be a fixed point of $q \mapsto \frac{\check{\xi}_t'(q)}{1+\check{\xi}_t'(q)}$, we have $\lim_{k\to\infty} \check{q}_k = q_{\mathsf{AMP}}$. By state evolution, similarly to the proof of Proposition 6.4.3, the first conclusion follows. Since
$$\frac{1}{N^j}\left\|\boldsymbol{x}^{\otimes j} - (\check{\boldsymbol{m}}^k)^{\otimes j}\right\|_2^2 = \langle \boldsymbol{x}, \boldsymbol{x}\rangle_N^j - 2\langle \boldsymbol{x}, \check{\boldsymbol{m}}^k\rangle_N^j + \langle \check{\boldsymbol{m}}^k, \check{\boldsymbol{m}}^k\rangle_N^j,$$
the second conclusion follows from the first. $\square$

Let
$$Q_{\mathsf{bayes}} = Q_{\mathsf{bayes}}(t) = \arg\max_{q\in[0,1)} \left\{\check{\xi}_t(q) + q + \log(1-q)\right\} \subseteq [0,1) \tag{6.263}$$
be the set of all maximizers of this quantity, and let
$$q_{\mathsf{bayes}} = q_{\mathsf{bayes}}(t) = \inf Q_{\mathsf{bayes}}(t).$$

Lemma 6.10.2. *For any t, the equation $\check{\xi}_t'(q) = \frac{q}{1-q}$ has finitely many solutions $q \in [0,1)$. Moreover, $Q_{\mathsf{bayes}}(t)$ is a finite set for all t. If $T_1 \subseteq [0,+\infty)$ is the set of t_1 such that $|Q_{\mathsf{bayes}}(t_1)| > 1$, then for each $t_1 \in T_1$, there exists $\delta > 0$ such that $(t_1-\delta, t_1+\delta) \cap T_1 = \{t_1\}$.*

Proof. Let $f_t(q) = (1-q)\check{\xi}_t'(q) - q$, so any solution to $\check{\xi}_t'(q) = \frac{q}{1-q}$ is a zero of f_t. Note that f_t is not identically zero: if it were, then $\check{\xi}_t'(q) = \frac{q}{1-q}$, contradicting that the coefficients γ_p^2 of ξ satisfy $\sum_{p\ge 2} 2^p \gamma_p^2 < \infty$. Since f_t is complex analytic in the unit disc, its zero set has no limit point, and in particular it has finitely many zeros in $[0,1)$. This shows that there are finitely many solutions to $\check{\xi}_t'(q) = \frac{q}{1-q}$.

Note that $\frac{\mathsf{d}}{\mathsf{d}q}(\check{\xi}_t(q) + q + \log(1-q)) = \check{\xi}_t'(q) - \frac{q}{1-q}$. Any interior maximizer of (6.263) must therefore satisfy the stationarity condition $\check{\xi}_t'(q) = \frac{q}{1-q}$. Since $\check{\xi}_t'(0) \ge 0$, 0 can be a maximizer only if it also solves this equation. Thus $Q_{\mathsf{bayes}}(t)$ is finite.

Consider an arbitrary $t_1 \in T_1$ and let $Q = Q_{\mathsf{bayes}}(t_1)$. For each $\widetilde{q} \in Q$, let $I_{\widetilde{q}} = [\widetilde{q}-\varepsilon, \widetilde{q}+\varepsilon]$, where $\varepsilon > 0$ is small enough that these intervals do not overlap. By continuity, for sufficiently small δ and all $t \in (t_1-\delta, t_1+\delta)$, all maximizers of $\check{\xi}_t(q) + q + \log(1-q)$ lie in $\bigcup_{\widetilde{q}\in Q} I_{\widetilde{q}}$. Let

$$m(t,\widetilde{q}) = \max_{q\in I_{\widetilde{q}}} \left\{\check{\xi}_t(q) + q + \log(1-q)\right\}, \qquad q(t,\widetilde{q}) = \arg\max_{q\in I_{\widetilde{q}}} \left\{\check{\xi}_t(q) + q + \log(1-q)\right\}.$$

Note that $q(t_1,\widetilde{q}) = \widetilde{q}$ for each $\widetilde{q} \in Q$. Since the maximum of $\check{\xi}_{t_1}(q) + q + \log(1-q)$ is attained over $I_{\widetilde{q}}$ uniquely at $\widetilde{q}$, by continuity $\lim_{t\to t_1} q(t,\widetilde{q}) = \widetilde{q}$.

For $\widetilde{q} \in Q$, $t \in (t_1, t_1+\delta)$, we have

$$\frac{m(t,\widetilde{q}) - m(t_1,\widetilde{q})}{t-t_1} \ge \frac{\check{\xi}_t(q(t_1,\widetilde{q})) - \check{\xi}_{t_1}(q(t_1,\widetilde{q}))}{t-t_1} = \sum_{j=1}^J \tau_j'(t_1) q(t_1,\widetilde{q})^j + O(t-t_1),$$
$$\frac{m(t,\widetilde{q}) - m(t_1,\widetilde{q})}{t-t_1} \le \frac{\check{\xi}_t(q(t,\widetilde{q})) - \check{\xi}_{t_1}(q(t,\widetilde{q}))}{t-t_1} = \sum_{j=1}^J \tau_j'(t_1) q(t,\widetilde{q})^j + O(t-t_1).$$

Taking the limit $t \downarrow t_1$ yields

$$\lim_{t\downarrow t_1} \frac{m(t,\widetilde{q}) - m(t_1,\widetilde{q})}{t-t_1} = \sum_{j=1}^J \tau_j'(t_1)\widetilde{q}^j.$$

A similar argument shows the left-derivative is also equal to this. Therefore

$$\frac{\partial}{\partial t} m(t,\widetilde{q})\big|_{t=t_1} = \sum_{j=1}^J \tau_j'(t_1)\widetilde{q}^j.$$

This quantity is distinct for different $\widetilde{q} \in Q$. Therefore, for all $t \in (t_1-\delta, t_1+\delta) \setminus \{t_1\}$, $|Q_{\mathsf{bayes}}(t)| = 1$. $\square$

Proposition 6.10.3. *Suppose $t \notin T_1$ satisfies $q_{\mathsf{bayes}}(t) > 0$. Let $\check{\xi}_t(q) = \sum_{p\ge 1} \beta_p^2 q^p$ (where we suppress the dependence of the β_p on t). For any p such that $\beta_p > 0$, we have (recall the definition of $\boldsymbol{m}_p(\vec{\boldsymbol{y}}_t, t)$ in Eq. (6.28)):*

$$\lim_{N\to\infty} \mathbb{E}\frac{1}{N^p}\left\|\boldsymbol{x}^{\otimes p} - \boldsymbol{m}_p(\vec{\boldsymbol{y}}_t,t)\right\|_2^2 = 1 - q_{\mathsf{bayes}}^p.$$

We first prove a preparatory lemma. In what follows, we let $\widetilde{\beta}_{p'} = \beta_{p'}$ be fixed for all $p' \ne p$ and treat $\widetilde{\beta}_p$ as a variable. Define $\widetilde{\xi}(q) = \sum_{p'\ge 1} \widetilde{\beta}_{p'}^2 q^{p'}$; we sometimes emphasize the dependence on $\widetilde{\beta}_p$ by writing $\widetilde{\xi}^{\widetilde{\beta}_p}(q)$. Let $\mathcal{P}$ denote the Parisi functional for spherical spin glasses, see e.g. [Tal06a, Equation (1.12)]. (In the proof below we will only need the replica-symmetric case of this functional, which is given in Proposition 6.5.14.) Further, for $q \in (-1,1)$, let

$$\widetilde{\xi}_q(s) = \widetilde{\xi}(q^2 + (1-q^2)s) - \widetilde{\xi}(q^2),$$

and define

$$P(\widetilde{\beta}_p) = \sup_{q\in[0,1)} \left\{\widetilde{\xi}(q) + \mathcal{P}(\widetilde{\xi}_q) + \frac{1}{2}\log(1-q^2)\right\}$$

Lemma 6.10.4. *Assume the setting of Proposition 6.10.3. For all $\widetilde{\beta}_p$ in a neighborhood of β_p,*

$$P(\widetilde{\beta}_p) = \frac{1}{2} \sup_{q\in[0,1)} \left\{ \widetilde{\xi}(1) + \widetilde{\xi}(q) + q + \log(1-q) \right\}. \tag{6.264}$$

Furthermore, P is differentiable at β_p, with

$$P'(\beta_p) = \beta_p(1 + q_{\mathsf{bayes}}^p). \tag{6.265}$$

Proof. By Proposition 6.5.14 with $u = \frac{q}{1+q}$, for all $q \in [0,1)$,

$$\mathcal{P}(\widetilde{\xi}_q) \le \frac{1}{2} \left\{ \widetilde{\xi}_q(1) - \widetilde{\xi}_q(u) + \frac{u}{1-u} + \log(1-u) \right\} = \frac{1}{2} \left\{ \widetilde{\xi}(1) - \widetilde{\xi}(q) + q - \log(1+q) \right\}, \tag{6.266}$$

and thus

$$P(\widetilde{\beta}_p) \le \frac{1}{2} \sup_{q\in[0,1)} \left\{ \widetilde{\xi}(1) + \widetilde{\xi}(q) + q + \log(1-q) \right\}.$$

Since $\lim_{q\uparrow 1} \log(1-q) = -\infty$, the supremum is attained. Let $q(\widetilde{\beta}_p)$ denote the maximizer. Arguing identically to the proof of Proposition 6.5.15, (6.266) is an equality at $q = q(\widetilde{\beta}_p)$. This proves (6.264).

Note that $q(\beta_p) = q_{\mathsf{bayes}}$ by definition. Since $t \notin T_1$, the maximum in (6.264) at $\widetilde{\beta}_p = \beta_p$ is attained uniquely at q_{bayes}. By continuity, $\lim_{\widetilde{\beta}_p \to \beta_p} q(\widetilde{\beta}_p) = q_{\mathsf{bayes}}$ as well. Note that for any $\widetilde{\beta}_p > \beta_p$,

$$\begin{aligned}
\frac{P(\widetilde{\beta}_p) - P(\beta_p)}{\widetilde{\beta}_p - \beta_p} &\ge \frac{\widetilde{\xi}^{\widetilde{\beta}_p}(1) + \widetilde{\xi}^{\widetilde{\beta}_p}(q(\beta_p)) - \widetilde{\xi}^{\beta_p}(1) - \widetilde{\xi}^{\beta_p}(q(\beta_p))}{\widetilde{\beta}_p - \beta_p} = 2\beta_p(1 + q(\beta_p)^p) + O(\widetilde{\beta}_p - \beta_p), \\
\frac{P(\widetilde{\beta}_p) - P(\beta_p)}{\widetilde{\beta}_p - \beta_p} &\le \frac{\widetilde{\xi}^{\widetilde{\beta}_p}(1) + \widetilde{\xi}^{\widetilde{\beta}_p}(q(\widetilde{\beta}_p)) - \widetilde{\xi}^{\beta_p}(1) - \widetilde{\xi}^{\beta_p}(q(\widetilde{\beta}_p))}{\widetilde{\beta}_p - \beta_p} = 2\beta_p(1 + q(\widetilde{\beta}_p)^p) + O(\widetilde{\beta}_p - \beta_p).
\end{aligned}$$

Taking the limit $\widetilde{\beta}_p \downarrow \beta_p$ yields

$$\lim_{\widetilde{\beta}_p \downarrow \beta_p} \frac{P(\widetilde{\beta}_p) - P(\beta_p)}{\widetilde{\beta}_p - \beta_p} = 2\beta_p(1 + q_{\mathsf{bayes}}^p).$$

A similar argument shows the left derivative also equals this, proving (6.265). $\square$

Proof of Proposition 6.10.3. Let $\widetilde{H}_N$ be a spin glass Hamiltonian with mixture $\widetilde{\xi}$, and let

$$F_N(\widetilde{\beta}_p) = \frac{1}{N} \mathbb{E} \log \int_{S_N} \exp\left\{ N\widetilde{\xi}(\langle \boldsymbol{x}, \boldsymbol{\sigma} \rangle_N) + \widetilde{H}_N(\boldsymbol{\sigma}) \right\} \, \mathsf{d}\mu_0(\boldsymbol{\sigma}).$$

Since the restriction of $\widetilde{H}_N$ to the band $\langle \boldsymbol{x}, \boldsymbol{\sigma} \rangle_N = q$ is a spin glass with mixture $\widetilde{\xi}_q$, the Parisi formula [Tal06a, Theorem 1.1] implies

$$\lim_{N\to\infty} F_N(\widetilde{\beta}_p) = \sup_{q\in(-1,1)} \left\{ \widetilde{\xi}(q) + \mathcal{P}(\widetilde{\xi}_q) + \frac{1}{2}\log(1-q^2) \right\}.$$

This equals $P(\widetilde{\beta}_p)$ because the supremum over $(-1,0]$ is clearly at most the supremum over $[0,1)$. By Hölder's inequality, $F_N(\widetilde{\beta}_p)$ is convex in $\widetilde{\beta}_p$. So, for any $\delta > 0$,

$$\frac{F_N(\beta_p) - F_N(\beta_p - \delta)}{\delta} \le F_N'(\beta_p) \le \frac{F_N(\beta_p + \delta) - F_N(\beta_p)}{\delta}.$$

Differentiability of P (by Lemma 6.10.4) then implies

$$\lim_{N\to\infty} F_N'(\beta_p) = P'(\beta_p) = \beta_p(1 + q_{\mathsf{bayes}}^p). \tag{6.267}$$

Let $\langle\cdot\rangle$ denote average w.r.t. the Gibbs measure corresponding to Hamiltonian $\check{H}_{N,t}$, which coincides in law with $N\widetilde{\xi}(\langle \boldsymbol{x},\boldsymbol{\sigma}\rangle_N) + \widetilde{H}_N(\boldsymbol{\sigma})$ for $\widetilde{\beta}_p = \beta_p$. Note that $\boldsymbol{m}_p(\vec{\boldsymbol{y}}_t,t) = \langle \boldsymbol{\sigma}^{\otimes p}\rangle$. We calculate that

$$\begin{aligned} F_N'(\beta_p) &= 2\beta_p \mathbb{E}\langle \langle \boldsymbol{x},\boldsymbol{\sigma}\rangle_N^p\rangle + \beta_p\left(1 - \mathbb{E}\langle\langle \boldsymbol{\sigma},\boldsymbol{\sigma}\rangle_N^p\rangle\right) \\ &= \beta_p\left(1 + 2\mathbb{E}\frac{\langle \boldsymbol{x}^{\otimes p}, \boldsymbol{m}_p(\vec{\boldsymbol{y}}_t,t)\rangle}{N^p} - \mathbb{E}\frac{\langle \boldsymbol{m}_p(\vec{\boldsymbol{y}}_t,t), \boldsymbol{m}_p(\vec{\boldsymbol{y}}_t,t)\rangle}{N^p}\right). \end{aligned}$$

Comparing with (6.267) shows

$$\lim_{N\to\infty}\left\{2\mathbb{E}\frac{\langle \boldsymbol{x}^{\otimes p}, \boldsymbol{m}_p(\vec{\boldsymbol{y}}_t,t)\rangle}{N^p} - \mathbb{E}\frac{\langle \boldsymbol{m}_p(\vec{\boldsymbol{y}}_t,t), \boldsymbol{m}_p(\vec{\boldsymbol{y}}_t,t)\rangle}{N^p}\right\} = q_{\mathsf{bayes}}^p.$$

Since

$$\mathbb{E}\frac{1}{N^p}\left\|\boldsymbol{x}^{\otimes p} - \boldsymbol{m}_p(\vec{\boldsymbol{y}}_t,t)\right\|_2^2 = 1 - 2\mathbb{E}\frac{\langle \boldsymbol{x}^{\otimes p}, \boldsymbol{m}_p(\vec{\boldsymbol{y}}_t,t)\rangle}{N^p} + \mathbb{E}\frac{\langle \boldsymbol{m}_p(\vec{\boldsymbol{y}}_t,t), \boldsymbol{m}_p(\vec{\boldsymbol{y}}_t,t)\rangle}{N^p},$$

the result follows. □

Lemma 6.10.5. *If there exists $q \in [0,1)$ such that $\xi''(q) > \frac{1}{(1-q)^2}$, then there exists $t \ge 0$ such that $\check{\xi}_t'(q) = \frac{q}{1-q}$ has more than one solution.*

Proof. Let $g_t(q) = \check{\xi}_t'(q) - \frac{q}{1-q}$, so solutions to $\check{\xi}_t'(q) = \frac{q}{1-q}$ are zeros of g_t. Suppose for contradiction that for all $t \ge 0$, g_t has unique zero $q_{\mathsf{AMP}}(t)$. Then, for all t, $g_t > 0$ on $[0, q_{\mathsf{AMP}}(t))$ (this is vacuous if $q_{\mathsf{AMP}}(t) = 0$) and $g_t < 0$ on $(q_{\mathsf{AMP}}(t), 1)$. Note that for each q, $g_t(q)$ is continuous and increasing in t, and thus $q_{\mathsf{AMP}}(t)$ is also continuous and increasing.

Recall that $\|\tau(t)\|_1 = t$ for all t. For each $q \in (0,1)$,

$$g_t(q) \ge \sum_{j=1}^J j\tau_j(t)q^{j-1} - \frac{q}{1-q} \ge \|\tau(t)\|_1 q^{J-1} - \frac{q}{1-q} = tq^{J-1} - \frac{q}{1-q}. \tag{6.268}$$

It follows that $g_t(q) > 0$ for sufficiently large t. Thus $\lim_{t\to+\infty} g_{\mathsf{AMP}}(t) = 1$, so $q_{\mathsf{AMP}}(t)$ ranges over all of $[0,1)$ as t ranges over $[0,+\infty)$.

Since $\xi''(q) > \frac{1}{(1-q)^2}$ for some $q \in [0,1)$, the function g_0 is not monotonically decreasing. Let $0 \le q_1 < q_2 < 1$ be such that $g_0(q_1) < g_0(q_2)$. Note that

$$g_t(q_1) - g_0(q_1) = \sum_{j=1}^J j\tau_j(t) q_1^{j-1} \le \sum_{j=1}^J j\tau_j(t)q_2^{j-1} = g_t(q_2) - g_0(q_2).$$

Thus $g_t(q_1) < g_t(q_2)$. Set t such that $q_1 = q_{\mathsf{AMP}}(t)$, so that $g_t(q_1) = 0$. This implies that $g_t(q_2) > 0$, and therefore g_t has another zero in $[q_2, 1)$. □

Lemma 6.10.6. *If there exists $t \ge 0$ such that $\check{\xi}_t'(q) = \frac{q}{1-q}$ has more than one solution, then there exists a nontrivial interval $I = [t_-, t_+] \subseteq [0,+\infty)$ such that for all $t' \in I$, $q_{\mathsf{AMP}}(t') \neq q_{\mathsf{bayes}}(t')$.*

Proof. Let g_t be defined as in the proof of Lemma 6.10.5 and $q_1 = q_{\mathsf{AMP}}(t)$, so that q_1 is the smallest zero of g_t. Let $q_2 > q_1$ be the next smallest zero of g_t. Note that by Lemma 6.10.2, either $g_t(q) > 0$ for all $q \in (q_1, q_2)$ or $g_t(q) < 0$ for all $q \in (q_1, q_2)$.

Suppose the former case holds. We will show the conclusion holds with $I = [t, t-\delta]$ for small δ. We first show that we must have $t > 0$, so this is a valid interval. Suppose for contradiction that $t = 0$; then $q_1 = 0$. So, $g_0(q) = \xi'(q) - \frac{q}{1-q}$ is positive on $(0, q_2)$. This implies that for $q \in (0, q_2]$,

$$\xi(q) + q + \log(1-q) = \int_0^q g_0(s)\ \mathsf{d}s > 0,$$

contradicting (6.31). Note that

$$\left(\check{\xi}_t(q_2) + q_2 + \log(1-q_2)\right) - \left(\check{\xi}_t(q_1) + q_1 + \log(1-q_1)\right) = \int_{q_1}^{q_2} g_t(q)\ \mathsf{d}q > 0.$$

We claim that $q_{\mathsf{AMP}}(t')$ is continuous on $t' \in I$, for small enough δ. If $q_1 = 0$, this is clear because $q_{\mathsf{AMP}}(t)$ is increasing. Otherwise, since $g_t(0) \geq 0$ and q_1 is the smallest zero of g_t, we have $g_t(q) > 0$ for $q \in [0, q_1)$. Since the $g_t(q)$ are continuous and increasing in t, the claim follows. It follows that for sufficiently small δ, for all $t' \in I$ and $q_1' = q_{\mathsf{AMP}}(t')$,

$$\left(\check{\xi}_{t'}(q_2) + q_2 + \log(1-q_2)\right) - \left(\check{\xi}_{t'}(q_1') + q_1' + \log(1-q_1')\right) > 0.$$

Thus $q_{\mathsf{AMP}}(t') \neq q_{\mathsf{bayes}}(t')$ for all $t' \in I$.

Finally, we consider the case that $g_t(q) < 0$ for all $q \in (q_1, q_2)$. Then, $g_t > 0$ on $[0, q_1)$ (vacuously if $q_1 = 0$) and $g_t < 0$ on (q_1, q_2). Let t'' be the smallest time such that $\inf_{q \in [q_1, q_2]} g_{t''}(q) \geq 0$; this is finite by the discussion surrounding (6.268). Since $g_t(q)$ is increasing in t, we have $g_{t''} \geq 0$ for $q \in [0, q_2]$, with equality attained at some $q \in [q_1, q_2]$. By definition, $q_{\mathsf{AMP}}(t'')$ is the smallest such q. As $f_{t''}(q_2) > g_t(q_2) = 0$, we have $q_{\mathsf{AMP}}(t'') < q_2$. The result now follows from the first case. □

Proof of Theorem 6.2.3. By the last two lemmas, there exists a nontrivial interval $I = [t_-, t_+] \subseteq [0, +\infty)$ such that $q_{\mathsf{AMP}}(t) \neq q_{\mathsf{bayes}}(t)$ for all $t \in I$. Since $q_{\mathsf{bayes}}(t)$ is a maximizer of (6.263), it satisfies the stationarity condition $\check{\xi}_t'(q) = \frac{q}{1-q}$, and therefore $q_{\mathsf{AMP}}(t) < q_{\mathsf{bayes}}(t)$. It also follows that $q_{\mathsf{bayes}}(t) > 0$.

Let $U(t)$ be the number of nonzero coefficients of $\check{\xi}_t$ of degree at most J. This is an increasing function with at most J discontinuities; let T_0 be the set of these discontinuities.

We will show the theorem holds with $\mathcal{I} = I \setminus (T_0 \cup T_1)$. (Recall the definition of T_1 in Lemma 6.10.2.) This is a positive measure set by Lemma 6.10.2. Consider any $t \in \mathcal{I}$. Since $t \notin T_0$, there exists $1 \leq j \leq J$ such that the q^j coefficient of $\check{\xi}_t$ is positive and $\tau_j'(t) > 0$. By Propositions 6.10.1 and 6.10.3,

$$\begin{aligned}\lim_{k\to\infty}\lim_{N\to\infty} \mathbb{E}\frac{1}{N^j}\left\|\boldsymbol{x}^{\otimes j} - (\check{\boldsymbol{m}}^k)^{\otimes j}\right\|_2^2 &= 1 - q_{\mathsf{AMP}}(t)^j,\\ \lim_{N\to\infty} \mathbb{E}\frac{1}{N^j}\|\boldsymbol{x}^{\otimes j} - \boldsymbol{m}_j(\vec{\boldsymbol{y}}_t, t)\|^2 &\leq 1 - q_{\mathsf{bayes}}(t)^j.\end{aligned}$$

Since $q_{\mathsf{AMP}}(t) < q_{\mathsf{bayes}}(t)$, the conclusion follows. □

Chapter 7

Weak Poincaré inequalities, simulated annealing, and sampling from spherical spin glasses

Abstract – There has been a recent surge of powerful tools to show rapid mixing of Markov chains, via functional inequalities such as *Poincaré inequalities*. In many situations, Markov chains fail to mix rapidly from a worst-case initialization, yet are expected to approximately sample from a random initialization. For example, this occurs if the target distribution has *metastable states*, small clusters accounting for a vanishing fraction of the mass that are essentially disconnected from the bulk of the measure. Under such conditions, a Poincaré inequality cannot hold, necessitating new tools to prove sampling guarantees.

We develop a framework to analyze simulated annealing, based on establishing so-called *weak Poincaré inequalities*. These inequalities imply mixing from a suitably warm start, and simulated annealing provides a way to chain such warm starts together into a sampling algorithm. We further identify a local-to-global principle to prove weak Poincaré inequalities, mirroring the spectral independence and localization schemes frameworks for analyzing mixing times of Markov chains.

As our main application, we prove that simulated annealing samples from the Gibbs measure of a spherical spin glass for inverse temperatures up to a natural threshold, matching recent algorithms based on algorithmic stochastic localization. This provides the first Markov chain sampling guarantee that holds beyond the *uniqueness threshold* for spherical spin glasses, where mixing from a worst-case initialization is provably slow due to the presence of metastable states. As an ingredient in our proof, we prove bounds on the operator norm of the covariance matrix of spherical spin glasses in the full replica-symmetric regime.

Additionally, we resolve a question related to sampling using data-based initializations.

7.1 Introduction

A common task of interest in computer science, probability, and physics is to efficiently sample from Gibbs distributions. For a Hamiltonian energy function $H : \Omega \to \mathbb{R}$ over state space $\Omega \subseteq \mathbb{R}^N$, the associated Gibbs distribution μ_H is defined by $\mathsf{d}\mu_H(x) \propto \exp(H(x))\mathsf{d}x$.

The class of *Markov chain Monte Carlo* (MCMC) algorithms is arguably the most widely used tool for sampling from Gibbs distributions. In this paradigm, one sets up a Markov chain P_H whose stationary distribution is μ_H, and outputs the final state of a $\mathsf{poly}(N)$-time random walk according to P_H. Common choices include the Glauber dynamics, for discrete state spaces such as $\Omega = \{\pm 1\}^N$, and the Langevin diffusion, for continuous state spaces such as $\Omega = \mathbb{R}^N$ or $\sqrt{N} \cdot \mathbb{S}^{N-1}$.

To prove that such an algorithm indeed correctly samples from μ_H, one bounds the *mixing time* of the Markov chain. A common route to prove a bound on the mixing time is to establish functional inequalities, such as *Poincaré* inequalities. There are now powerful frameworks for proving such functional inequalities, such as *spectral independence* [ALO21] and *localization schemes* [CE22]. The development of these frameworks has led to a flurry of activity in analyzing mixing times of Markov chains, including the resolution of several long-standing open problems in the algorithmic theory of counting and sampling [ALOV24, ALO21, AJK+22, EKZ22, CE22].

The implications of these inequalities are quite strong. In particular, they imply that for *any* initial distribution ν, for an appropriate divergence function, a single step of the Markov chain shrinks the distance to the stationary distribution by a significant multiplicative factor:

$$\mathrm{Divergence}(P_H\nu\|\mu_H) \leq \left(1 - \frac{1}{\mathsf{poly}(N)}\right)\mathrm{Divergence}(\nu\|\mu_H)\,.$$

The presence of such a functional inequality typically implies that a Markov chain mixes rapidly from a *worst-case initialization.*

Sampling from random initializations. Many natural Markov chains are expected to produce approximate samples from the Gibbs measure when started at a *random* initialization, but fail to mix rapidly from a worst-case initialization. Often, this is because the Gibbs measure contains pathological clusters (termed *metastable states* in the physics literature) that are essentially disconnected from most of the measure, and account for a vanishing fraction of the total mass. A Markov chain initialized in such a cluster will remain trapped inside it and fail to mix, and therefore methods that show mixing from worst-case initializations cannot give effective guarantees in such settings.

However, one may still hope to show that from a random initialization, the Markov chain samples from the non-pathological part of the Gibbs measure, which is statistically indistinguishable from the true Gibbs measure. In our work, we prove that under suitable conditions, the *simulated annealing* algorithm samples from a distribution close to the Gibbs measure.

Simulated annealing. In the simulated annealing algorithm, one defines a "schedule" of inverse temperatures, i.e. for $i = 0, \ldots, T$, let $\beta_i \coloneqq i/T$. The algorithm initializes at a sample from the uniform distribution $\mu_{\beta_0 H}$. Then, for $i = 1, \ldots, T$, the i-th stage of the algorithm runs the Markov chain $P_{\beta_i H}$ corresponding to $\mu_{\beta_i H}$ for $\mathsf{poly}(N)$ time, initialized at the output of the previous stage.

The underlying idea of this algorithm is that, for T sufficiently large, the Gibbs distribution $\mu_{\beta_{i-1}H}$ is a "warm start" for $\mu_{\beta_i H}$, i.e. an initialization with suitably bounded likelihood ratio with $\mu_{\beta_i H}$. So, if one could show that each of the Markov chains $P_{\beta_i H}$ (approximately) mixes rapidly from a warm start, one may inductively argue that the output of the i-th stage of the algorithm is an approximate sample from $\mu_{\beta_i H}$. In other words, simulated annealing chains a sequence of warm starts together into a sampling algorithm.

This algorithmic idea is widely used empirically, and has also been employed to obtain algorithms for approximating the volumes of convex bodies [DFK91, DF91, LS90, KLS97], approximating the number of perfect matchings in a bipartite graph [JSV04], and sampling from the random field Ising model at sufficiently high temperatures [AEGP23], among others. However, we lack a general theory for why simulated annealing achieves provable guarantees beyond the settings of sampling from log-concave distributions and convex bodies. Indeed, in contrast to the general recipes available to prove mixing from worst-case initialization, proofs of rapid mixing from warm starts often employ ad-hoc arguments.

One of the main contributions of this work is a framework for proving mixing from warm starts, which combined with the above discussion provides general sufficient conditions under which simulated annealing samples from the Gibbs measure. We achieve this by generalizing the frameworks of spectral independence and localization schemes, previously employed to prove mixing from worst-case initialization, to show mixing from warm starts (see Section 7.6 for details). As we discuss just below, our framework gives sampling guarantees for simulated annealing in regimes where mixing from worst-case initializations is provably false.

Main application: spherical spin glasses. In a *spherical mixed p-spin glass*, $H : \sqrt{N} \cdot \mathbb{S}^{N-1} \to \mathbb{R}$ is a random Hamiltonian parameterized by coefficients $\beta, \gamma_2, \ldots, \gamma_{p_*} \geq 0$ where:

$$H(\sigma) = \beta \sum_{p \geq 2} \frac{\gamma_p}{N^{(p-1)/2}} \sum_{i_1,\ldots,i_p} \boldsymbol{g}_{i_1,\ldots,i_p} \sigma_{i_1} \cdots \sigma_{i_p}\,, \tag{7.1}$$

for i.i.d. $\boldsymbol{g}_{i_1,\ldots,i_p} \sim \mathcal{N}(0,1)$. The Gibbs distribution μ_H is very well-studied in probability, statistical physics, and average-case algorithm design, as it simultaneously exhibits rich behavior and is amenable to analytic tools. Notably, this model undergoes numerous sharp phase transitions as one increases β. For small β, the model satisfies a Poincaré inequality [GJ19]. Beyond a *uniqueness* transition β_{uniq}, small isolated clusters in μ_H known as *metastable states* start to appear [BJ24]. In particular, the natural Markov chain *Langevin diffusion* initialized from such states mixes slowly, thereby precluding a Poincaré inequality. However, these states account for a vanishing fraction of the measure under μ_H, and the Langevin diffusion with a random initialization is expected to still mix rapidly over a $1 - o_N(1)$ fraction of μ_H, thereby producing a sample with vanishing total variation distance from μ_H.

The threshold for efficient algorithmic sampling is believed to occur at the *shattering* transition β_{sh}—beyond this transition, the Gibbs measure shatters into an exponential number of poorly-connected clusters with exponentially small mass, and mixing is provably slow [CHS93, AMS25, GJK23]. It is expected that all efficient algorithms fail to sample from the Gibbs measure above β_{sh}, and recently [AMS25] gave rigorous evidence for this picture by showing that all *stable* algorithms fail.

We use our framework to prove that *annealed* Langevin diffusion, where one begins by running Langevin diffusion for $\beta_0 = 0$, and slowly increases the inverse temperature to the target β, samples from the spherical mixed p-spin glass. This leads to the first rigorous guarantee in this problem for a Markov chain beyond the uniqueness threshold.

Theorem 7.1.1 (Informal)**.** *For any choice of $\gamma_2, \ldots, \gamma_{p_*}$, there is a threshold $\beta_{\mathsf{SL}} \le \beta_{\mathsf{sh}}$ such that for any $\beta < \beta_{\mathsf{SL}}$, with probability at least $1 - e^{-\Omega(N^{1/5})}$ over the randomness of H, annealed Langevin diffusion run for $\mathsf{poly}(N)$ time samples from a distribution whose total variation distance to μ_H is at most $e^{-\Omega(N^{1/5})}$.*

The thresholds β_{SL} and β_{sh} are formally defined as the supremal β such that the inequalities (SL) and (Non-shattering) below hold. The recent work [HMP24] produces a different sampling algorithm based on algorithmic stochastic localization, which succeeds to the same threshold β_{SL}; see below for further discussion. This threshold is a fundamental barrier for stochastic localization based approaches, and we explain its physical significance in Remark 7.7.4.

Remark 7.1.2. In many models, we have $\beta_{\mathsf{uniq}} < \beta_{\mathsf{SL}} < \beta_{\mathsf{sh}}$, and β_{SL} is close to β_{sh}. For example, for the pure p-spin models (where $\gamma_p = 1$ and all the other γ_i are equal to 0), $\beta_{\mathsf{uniq}} \asymp (\log p)^{-1/2}$, while $\beta_{\mathsf{SL}}, \beta_{\mathsf{sh}} \asymp 1$ and $\beta_{\mathsf{SL}}/\beta_{\mathsf{sh}}$ is bounded from below by the universal constant $\sqrt{e}/2$. See [HMP24, Remark 1.1, Eq. 1.8].

7.1.1 Weak Poincaré inequalities and localization schemes

The starting point of our work is a relaxation of Poincaré inequalities, known as *weak Poincaré inequalities*, which can be leveraged to prove mixing from warm starts. To simplify the discussion, we restrict here to the setting of discrete Markov chains. Our main application is to a continuous Markov chain, namely the Langevin diffusion for a spherical spin glass, and we outline the differences in Remark 7.1.7 below.

Let P_H be a time-reversible Markov chain with stationary distribution μ_H. For any functions $f, g : \Omega \to \mathbb{R}$, define the *Dirichlet form* as $\mathcal{E}(f,g) \coloneqq \mathbb{E}_{\boldsymbol{x}\sim\mu_H} \mathbb{E}_{\boldsymbol{y}\sim P_H \boldsymbol{x}} (f(\boldsymbol{x}) - f(\boldsymbol{y}))(g(\boldsymbol{x}) - g(\boldsymbol{y}))$. We say P_H satisfies a *C-Poincaré inequality* if for any function $f : \Omega \to \mathbb{R}$:

$$\mathcal{E}(f,f) \ge C \cdot \mathsf{Var}[f]\,,$$

for $C \ge 1/\mathsf{poly}(N)$. A Poincaré inequality has a classic implication for rapid mixing. In particular, for ν_t as the distribution obtained by running P_H for continuous time t starting at a distribution ν_0, we have:

$$\chi^2(\nu_t \| \mu_H) \le \exp(-Ct) \cdot \chi^2(\nu_0 \| \mu_H)\,.$$

We say P_H satisfies a (C,ε)-weak Poincaré inequality if for any function $f : \Omega \to \mathbb{R}$:

$$\mathcal{E}(f,f) \ge C \cdot \mathsf{Var}[f] - \varepsilon \cdot \|f - \mathbb{E} f\|_\infty^2$$

One can derive the following mixing guarantee from a weak Poincaré inequality; see, e.g., [RW01, Theorem 2.1].

$$\chi^2(\nu_t \| \mu_H) \le \exp(-Ct) \cdot \chi^2(\nu_0 \| \mu_H) + \varepsilon \cdot \left\| \frac{\mathsf{d}\nu_0}{\mathsf{d}\mu_H} - 1 \right\|_\infty^2 . \tag{7.2}$$

In particular, if ν_0 is a warm start for μ_H in the sense that $\left\| \frac{d\nu_0}{d\mu_H} - 1 \right\|_\infty$ is suitably small, this implies that the Markov chain's output distribution ν_t approximates μ_H.

Since the target measure in one stage of simulated annealing is a warm start for that of the next stage, such a guarantee allows one to inductively argue that simulated annealing succeeds at sampling. We summarize this implication below.

Theorem 7.1.3 (Informal, see Theorem 7.4.12)**.** *If $P_{\beta H}$ satisfies a weak Poincaré inequality with suitable parameters for every $\beta \in [0, 1]$, then simulated annealing succeeds at sampling from μ_H.*

Localization schemes for weak Poincaré inequalities. We restrict to the following simple setting: μ_H is a distribution on $\{\pm 1\}^N$. Let P_H be the *Glauber dynamics* Markov chain where in a single step from x, we pick a uniformly random coordinate $i \sim [N]$, and toggle x_i with probability:

$$\frac{\mu_H(x^{\oplus i})}{\mu_H(x) + \mu_H(x^{\oplus i})}.$$

A special case of the localization schemes framework is the *spectral independence* framework of Anari, Liu, and Oveis Gharan [ALO21].

Theorem 7.1.4 ([AL20, ALO21])**.** *The following* local-to-global principle *reduces proving a Poincaré inequality to establishing bounds on the spectrum of* influence matrices*. Suppose for **every** $S \subseteq [N]$, and **every** pinning x_S of coordinates in S, the spectral norm of its* influence matrix *Ψ_{S,x_S} is at most α, then the Glauber dynamics chain satisfies a $n^{-O(\alpha)}$-Poincaré inequality. Here, the influence matrix Ψ_{S,x_S} is an $(n - |S|) \times (n - |S|)$ matrix indexed by vertices $v \notin S$, where*

$$\Psi_{S,x_S}[a, b] \coloneqq \Pr[x_a = +1 | x_b = +1] - \Pr[x_a = -1 | x_b = +1].$$

While the above theorem has been influential and useful in proving mixing time bounds for a variety of Markov chains relevant to sampling combinatorial structures, the "for every" requirement in the above theorem is quite punishing in average-case settings. For example, in the presence of metastable states, P_H does not satisfy a Poincaré inequality, but may nevertheless satisfy a weak Poincaré inequality. In such cases, there are choices of S and x_S for which Ψ_{S,x_S} has large spectral norm, and the above statement has no implications for the mixing time of P_H.

We give a general local-to-global principle to prove weak Poincaré inequalities. A one-line summary of this local-to-global principle is:

> *Bounded influence over **all** pinnings implies a Poincaré inequality.*

An analogous summary of the local-to-global principle in the present paper is:

> *Bounded influence over **typical** pinnings implies a **weak** Poincaré inequality.*

To give a more concrete instantiation of our message, our result implies a "softer" version of Theorem 7.1.4, tolerant to some "bad" pinnings, which we state below.

Theorem 7.1.5 (Special case of Lemma 7.6.8)**.** *Let $i_1, \ldots, i_N$ be a random permutation of $[N]$, let $S_t \coloneqq \{i_1, \ldots, i_t\}$, and let $\boldsymbol{x} \sim \mu_H$. Suppose with probability $1 - \varepsilon$ over the randomness of $\boldsymbol{x}$ and the permutation $i_1, \ldots, i_N$, we have that for every $t \in [N]$, the influence matrix $\Psi_{S_t, \boldsymbol{x}_{S_t}}$ has spectral norm bounded by α. Then, P_H satisfies a $(n^{-O(\alpha)}, O(\varepsilon))$-weak Poincaré inequality.*

Remark 7.1.6. The reader should think of the spectral norm of $\Psi_{S_t, \boldsymbol{x}_{S_t}}$ as quantifying how much variance of the distribution $\mu_H | \boldsymbol{x}_{S_t}$ is "explained" by revealing $\boldsymbol{x}_{i_{t+1}}$.

Remark 7.1.7. Theorem 7.1.5 holds at a more general level, for a large family of *localization schemes*; see [CE22] for examples of localization schemes and further discussion. The localization scheme at play in the above local-to-global principle is process of revealing coordinates of a Gibbs sample $\boldsymbol{x}$ in random order.

In our main application of sampling from a spherical spin glass using simulated annealing of Langevin diffusion, we consider a different localization scheme, *stochastic localization*, where the revealed information at time t is $\boldsymbol{y}_t = t\boldsymbol{x} + B_t$ where $(B_t)_{t \geq 0}$ is a standard Brownian motion. Analyzing this localization scheme requires studying *exponential tilts* rather than pinnings of μ_H. The analogous local-to-global principle in this setting is:

Bounded covariance over ***typical*** *exponential tilts implies a* ***weak*** *Poincaré inequality.*

We defer a technical discussion to Section 7.2, and refer to Lemma 7.6.8 for a formal statement.

7.1.2 Sampling from spherical spin glasses

We now state our main results for sampling from spherical spin glasses. We will encode the coefficients $\gamma_2, \ldots, \gamma_{p_*}$ in (7.1) into the *mixture function* $\xi(q) = \sum_{p=2}^{p_*} \gamma_p^2 q^p$. Note that the parameter β in (7.1) can of course be absorbed into the γ_p, so we can state thresholds directly in terms of the function ξ. Physics heuristics [CHS93] suggest that Glauber dynamics and Langevin diffusion, with random initialization, sample from μ_H with vanishing total variation error under the following condition. Note that this and the below conditions take the form of an upper bound on ξ or its derivatives, and therefore demarcate a region of sufficiently high temperature.

$$\xi'(q) < \frac{q}{1-q} \text{ for all } q \in (0,1). \tag{Non-shattering}$$

Recent work by one of the authors, Montanari, and Pham [HMP24] gives an algorithm based on simulating Eldan's stochastic localization process [Eld13, Eld20b] (see below), which samples from μ_H with vanishing total variation error under the following condition.

$$\xi''(q) < \frac{1}{(1-q)^2} \text{ for all } q \in [0,1). \tag{SL}$$

Note that this condition implies (Non-shattering), which can be seen by integrating the inequality. [HMP24] also shows a matching hardness result, that for any strictly replica symmetric model (see (Strict RS) below) *not* satisfying (SL), a generalized family of stochastic localization algorithms fails to sample from μ_H.

Our main result is that simulated annealing samples from μ_H in the same regime.

Theorem 7.1.8 (See Theorem 7.7.2)**.** *Under* (SL)*, with probability at least* $1 - e^{-\Omega(N^{1/5})}$ *over the randomness of* H*, annealed Langevin dynamics produces a sample whose total variation distance to* μ_H *is at most* $e^{-\Omega(N^{1/5})}$*.*

As alluded to in the above discussion, the main input to our framework is a high-probability covariance bound on the *random* exponential tilts of the Gibbs measure encountered along the stochastic localization process. Combined with our weak Poincaré inequality framework, this implies that simulated annealing samples from the Gibbs measure. On the way to proving these covariance bounds, we establish a high-probability covariance bound on all spherical spin glasses in the *(strictly) replica symmetric* phase, a high-temperature phase where the model enjoys a certain notion of correlation decay.

$$\xi''(0) < 1 \text{ and } \xi(q) + q + \log(1-q) < 0 \text{ for all } q \in (0,1). \tag{Strict RS}$$

Theorem 7.1.9 (Informal, see Theorem 7.7.32)**.** *Under* (Strict RS)*, with probability* $1 - e^{-\Omega(N^{1/5})}$ *over the randomness of* H*,* $\|\mathsf{Cov}(\mu_H)\|_{\mathsf{op}} = O(1)$*.*

This is the first covariance bound to cover the entire replica symmetric phase with higher order interactions, and we believe it is interesting in its own right. This result is sharp: in the complement of the replica symmetric regime, arguing as in [AG24, Proposition 4.2] shows that $\mathbb{E}\,\|\mathsf{Cov}(\mu_H)\|_{\mathsf{op}}$ is diverging, of order $\Omega(\sqrt{N})$.

The relation between (SL) and (Strict RS) is as follows. First, (Strict RS) follows from (SL) by integrating twice. Second, (SL) is equivalent to the condition that random exponential tilts of μ_H of any magnitude are typically replica symmetric. This is needed for the algorithmic stochastic localization approach of [HMP24], and arises in the current work (where stochastic localization appears as an analysis tool, rather than as an algorithm) for a similar reason, see Remark 7.7.4.

The connection from Theorem 7.1.9 to high-probability covariance bounds on the tilted measures encountered along the localization process relies on a reduction developed in [HMP24]. This reduction implies that typically, the vast majority of the mass of these tilted measures live near a certain codimension-2 band passing through a *TAP fixed point*, which behaves like a spin glass in two fewer dimensions. The proof of

Theorem 7.1.9 also builds on tools developed in [HMP24], and by one of the authors and Sellke in [HS23b], which together provide high-precision control of partition functions in the replica symmetric regime.

On the other hand, our approach also leads to several improvements over earlier results. First, we obtain a sampler with total variation error $e^{-\Omega(N^{1/5})}$ with probability $1-e^{-\Omega(N^{1/5})}$, whereas [HMP24] obtains total variation error $N^{-\varepsilon}$ with probability $1-N^{-\varepsilon}$, for small constant ε. Our total variation error is close to the best possible, as beyond the uniqueness threshold, at least a $e^{-O(N)}$ fraction of μ_H is typically trapped in metastable states [BJ24], which are hard to reach. Moreover, there is no longer a need to encode a mean estimator for the stochastic localization process (see below) directly in the algorithm; running a natural Markov chain is sufficient.

More conceptually, our work gives the first analysis of a Markov chain for this problem that "sees" the benignness of a random initialization and overcomes the uniqueness threshold.

7.1.3 Weak Poincaré inequalities beyond annealing

The discussion thus far has been focused on proving mixing time bounds for Markov chains initialized at warm starts. In fact, our framework extends beyond this and can be used to prove rapid mixing of a Markov chain initialized at a distribution that "sees" the different components of the target distribution. For instance, consider the simple scenario where the target distribution π is a mixture of two disconnected component distributions, each of which satisfies a (true) Poincaré inequality. The disconnectedness means that the full distribution π does not satisfy a true Poincaré inequality. However, if we initialize at a distribution that splits its mass equally between the two components, we would expect a Markov chain to rapidly mix to the target distribution.

How does one convert this belief to a (generalizable) proof? The key is that while the distribution may not satisfy a Poincaré inequality for *all* functions, a variant of such an inequality does hold for functions encountered along the trajectory of the Markov chain. More concretely, we may prove the following theorem.

Theorem 7.1.10 (Informal, see Theorem 7.4.6)**.** *Consider the trajectory $(\nu_t)_{t\geq 0}$ of a Markov chain with stationary distribution π, initialized at a distribution ν_0. Suppose that for all $s \leq T$,*

$$\mathcal{E}\left(\frac{\mathsf{d}\nu_s}{\mathsf{d}\pi}, \frac{\mathsf{d}\nu_s}{\mathsf{d}\pi}\right) \geq \rho_{\mathrm{PI}}\left(\mathsf{Var}_\pi\left[\frac{\mathsf{d}\nu_s}{\mathsf{d}\pi}\right] - \delta\right).$$

Then,

$$\chi^2(\nu_T\|\pi) \leq e^{-2\rho_{\mathrm{PI}}T}\chi^2(\nu_0\|\pi) + \delta.$$

We remark that our earlier equation (7.2) is a near-immediate consequence of the above. Returning to the above example with two disconnected components, if ν_s placed exactly half its mass on each of the two components, the error δ can be taken to be 0.

For our first application in Section 7.5, we use this picture of how the initialization can capture "symmetries" in the distribution.

Sampling from mixtures of log-concave distributions with advice. An example of a distribution where we can take advantage of "symmetries" is the following. Suppose we have a distribution π which is a mixture of K distributions

$$\pi = \sum_{i=1}^{K} p_i \pi_i,$$

each of which is well-connected (e.g., satisfies a Poincaré inequality). We do not expect a Markov chain to rapidly mix to π from a worst-case initialization. Does the scenario change if we initialize more cleverly? To be concrete, suppose we are given m samples $x_1, \ldots, x_m$ from π, and initialize our Markov chain at the empirical distribution $\sum_{i=1}^{m} \delta_{x_i}$. If the component measures (π_i) are "far apart" and do not interact with each other, we would expect the Markov chain to rapidly mix from this initialization if the fraction of points in each cluster is (approximately) equal to the correct fraction p_i. On the other hand, if the component measures were very close together, we would expect their mixture to also satisfy a Poincaré inequality.

However, it is unclear how to translate this intuition to a proof. In previous work [KV24], sampling guarantees are provided for this algorithm, but the running time has a doubly exponential dependence on the number of components K. Our second illustration of weak Poincaré inequalities provides high-probability sampling guarantees for this problem, by running Langevin diffusion for time that is polynomial in all parameters involved. We refer the reader to Section 7.5 for the details of the theorem statement and its (self-contained) proof.

This problem is studied extensively in an independent work of Koehler, Lee, and Vuong [KLV24]. Motivated by the success of *score matching* methods in modern machine learning, they prove that Langevin dynamics and Glauber dynamics converge to the stationary distribution when initialized from the above empirical distribution under similar conditions to our setting, even if the Markov chain updates come from a slightly perturbed distribution (i.e. if they were learned by a score matching algorithm). They also use their techniques to give an efficient algorithm for learning approximately low-rank Ising models.

7.1.4 Related work

Markov chain mixing and localization schemes. The first use of the local-to-global phenomenon in mixing was in the work of [ALOV24] on establishing rapid mixing of the "basis exchange" walk on bases of a matroid, which used the local-to-global theorem for simplicial complexes from [KO20]. Their approach was later formalized into the framework of *spectral independence* [ALO21], which was widely successful in resolving numerous problems in algorithmic sampling and counting; see [Liu23] for a comprehensive literature survey.

In the world of sampling from continuous distributions, most recent progress on the KLS conjecture on the Poincaré constant of isotropic log-concave distributions (see [LV24] and the recent survey [KL24]) has employed Eldan's stochastic localization [Eld13]. Later, stochastic localization was used in the work of Eldan, Koehler, and Zeitouni [EKZ22] to analyze the Poincaré constant for Glauber dynamics on Ising models. The seemingly unrelated techniques of spectral independence and stochastic localization approaches to analyzing mixing times were unified under the framework of localization schemes [CE22], which, as an application, also simplified the proof of [EKZ22].

Weak Poincaré inequalities. The study of weak Poincaré inequalities was initiated in the work of Aida [Aid98] and Mathieu [Mat06] in the context of proving other functional inequalities. The work of Röckner and Wang [RW01] observed the connection between a Markov chain satisfying a weak Poincaré inequality, and rapid mixing from "sufficiently warm starts". We refer the reader to the monograph of Wang [Wan06, Chapter 4] for a comprehensive treatment of weak Poincaré inequalities and their implications to mixing and concentration.

Weak Poincaré inequalities are also related to the notion of s-conductance, a weakened version of conductance introduced in [LS93] which has been used frequently in the literature on sampling from convex bodies (see [Che23b, Section 7.4.2] for a textbook treatment). This connection is explained in [GMT06]. We also refer the reader to [CGG07], which defines a notion of weak log-Sobolev inequality and uses it to derive a rapid mixing result.

The work [AEGP23] gives a sampling algorithm for the ferromagnetic random-field Ising model on a finite domain $D \subseteq \mathbb{Z}^d$, which follows an approach of chaining warm starts similar to the present work, inspired by convex body sampling literature [LS93]. [AEGP23] shows that in a certain parameter regime, the Glauber dynamics for this model satisfy a weak Poincaré inequality. They then construct an increasing sequence of sub-domains $D_0 \subset D_1 \subset \cdots \subset D_T = D$ and show that a sample from the model on D_i can be converted to a warm start for the model on D_{i+1}. Since the weak Poincaré inequality implies mixing from a warm start, this yields a sampling algorithm based on running the Glauber dynamics on this increasing sequence of models.

The work [AJK$^+$21b] introduces a related notion of restricted modified log-Sobolev inequality, which implies entropy contraction (without an additive error, in contrast to a weak Poincaré inequality) for all warm starts. This is used to derive optimal mixing times for several Markov chains. In the opposite direction, [PS19] introduces a strengthened log-Sobolev inequality where the entropy is bounded by a nonlinear function of the Dirichlet form. This is used to obtain improved hypercontractivity and Fourier coefficient bounds for functions with small support.

Sampling from random initializations. The separation between worst-case mixing times and mixing from a random initialization has been studied in a variety of other settings. [CDL+12, BGZ25] characterize which product measure initializations enjoy rapid mixing in a temperature range where worst-case mixing is exponential for the Curie-Weiss Potts model. Notably, as discussed in [BGZ25, Section 1.3], their analysis also characterizes mixing from initializations constructed by simulated annealing, [LS16, LS17] show that a uniform initialization halves the mixing time for Glauber dynamics for the ferromagnetic Ising model on bounded degree graphs, such as the 1D lattice. [GS22] introduces the notion of weak spatial mixing in a phase, and proves that Glauber dynamics for the ferromagnetic Ising model on the 2D lattice has rapid mixing when initialized uniformly at $\pm\vec{\mathbf{1}}$. [GS24] uses the same notion to study mixing from a similar random initialization for a certain natural Markov chain for the random cluster model. [BNN24] show rapid mixing for Glauber dynamics for the exponential random graph model when initialized from a carefully chosen Erdős–Rényi random graph.

Sampling from spherical spin glasses and algorithmic stochastic localization. There is a long history of work studying Markov chain dynamics on spin glasses. An important line of work [CHS93, CK93, BCKM98, BDG06, BGJ20, CCM21, Sel24b] studies the Langevin dynamics for spherical spin glasses on an N-independent time scale. While the Langevin dynamics do not mix on this time scale, these works capture important statistics of the trajectory such as the energy attained by the Langevin dynamics after a given time, and uncover deep phenomena such as *aging*.

Rapid mixing guarantees at sufficiently high temperature were obtained in [GJ19] for the Langevin dynamics for spherical spin glasses, and in [BB19, EKZ22, AJK+22, ABXY24, AJK+24, AKV24] for the Glauber dynamics for the Sherrington–Kirkpatrick model [SK75] and Ising spin glasses. These approaches show mixing from a worst-case initialization via a functional inequality.

Recently, [AMS22, AMS23b] introduced a new sampling algorithm based on simulating Eldan's stochastic localization scheme [Eld13, Eld20b]. This approach has since been used in applications such as Bayesian posterior sampling [MW23, MW24], and is closely related to the denoising diffusions method in machine learning [SDWMG15, HJA20, SSDK+21] (see [Mon23b] for details). The resulting algorithm samples in a wider range of temperatures, though with the weaker guarantee of vanishing *Wasserstein* rather than total variation error. The recent work [HMP24] improved this guarantee to total variation, and the resulting algorithm succeeds to the same threshold (SL) as in the present work.

Within the algorithmic stochastic localization approach, the main task is to estimate the means of a sequence of exponential tilts of the Gibbs measure, which appear as the drift process of a stochastic differential equation parametrizing the localization process. In [AMS22], this is achieved with an estimator based on approximate message passing (AMP), which is accurate to leading order. [HMP24] develops an improved estimator with a suitable correction term, which improves the algorithm's guarantee from Wasserstein to total variation error.

Covariance bounds for spin glasses. There has been a great deal of recent work on covariance bounds for spin glasses [BXY23, AG24, BSXY24], in part due to the connection between covariance bounds and functional inequalities developed in the localization schemes literature. In particular, [AG24, BSXY24] address the case of the Sherrington–Kirkpatrick (SK) model, and [BXY23] addresses the SK model with external field.

7.1.5 Open problems

We conclude with several open problems.

Non-sampling guarantees for simulated annealing. While we initiate a study of simulated annealing to attain sampling guarantees, one could ask how to analyze simulated annealing beyond sampling. In recent work [LMR+24], three of the authors, Liu, and Raghavendra introduce the framework of *locally stationary distributions* to analyze slow-mixing Markov chains, and leverage it to obtain recovery guarantees for the spiked Wigner and stochastic block model inference problems. We start by reiterating [LMR+24, Problems 1.20 and 1.21]—is simulated annealing computationally optimal for random CSPs with planted solutions?

Further, consider the problem of optimizing the Hamiltonian (7.1) of the mixed p-spin model. Historically, simulated annealing was one of the earliest algorithms developed for this problem [CHKW23]. The works [Mon21, Sub21a, AMS21, Sel24a] develop algorithms that are optimal among suitably Lipschitz algorithms [HS25] and conjecturally among all efficient algorithms. The limiting energy obtained by natural Markov chain dynamics is an long-standing question in its own right [CK93], which was solved for pure models in [Sel24b] but is otherwise open. We ask:

Problem 7.1.11. What energy does simulated annealing obtain when run on the Hamiltonian (7.1)?

We refer the reader to [MRT04, FFRT21] and references therein for relevant discussion. We also ask the following question, which seems instrumental to making progress towards the above.

Problem 7.1.12. How does a non-worst-case initialization (such as one constructed by simulated annealing) affect the locally stationary distribution that is reached by a Markov chain?

Along similar lines, we have the following concrete question about understanding Markov chains from non-worst-case initializations.

Worst-case combinatorial optimization via simulated annealing. The paradigm of solving a semidefinite program and rounding its solution has been extremely successful at achieving optimal approximation guarantees for a wide variety of combinatorial optimization problems, especially constraint satisfaction problems [KKMO07, Rag08].

However, on large families of instances (sparse ones for instance), the solutions produced by these SDPs can be refined locally to improve the approximation ratio, but these improvements do not match the corresponding hardness thresholds. For example, for the problem of Max Cut, the classical SDP algorithm [GW95] gives an α_{GW}-approximation for $\alpha_{GW} \approx 0.878$, and a local refinement [HK23] produces an $\alpha_{\mathrm{GW}} + \Omega\left(\frac{1}{d^2}\right)$-approximation. On the other hand, it is (UG-)hard [Tre01] to approximate the max-cut better than $\alpha_{\mathrm{GW}} + O\left(\frac{1}{\sqrt{d}}\right)$.

Problem 7.1.13. Does a Markov chain initialized at the SDP solution attain a $\alpha_{\mathrm{GW}} + \Omega\left(\frac{1}{\sqrt{d}}\right)$-approximation to the max-cut in a bounded degree graph?

Sampling from spin glasses up to the shattering threshold. It is conjectured that the Langevin diffusion with uniform random initialization samples from spherical p-spin models for inverse temperatures up to the shattering threshold (Non-shattering) [CHS93, CK93]. Similarly, this is conjectured for the Glauber dynamics Markov chain for models over the hypercube $\{\pm 1\}^N$ instead of the sphere S_N, for an analogous shattering threshold. As a start, can we show such guarantees for simulated annealing (as opposed to a fixed-temperature Markov chain from uniform initialization)?

Problem 7.1.14. Does simulated annealing sample from p-spin models up to the shattering threshold?

The failure of algorithmic stochastic localization beyond the (SL) condition [HMP24, Section 10] suggests that ideas beyond our proof strategy are required to prove the above.

Simulated annealing in more general models. For sampling from the spherical p-spin model, our results show that simulated annealing succeeds in the regime (SL) where algorithmic stochastic localization succeeds. At the level of proofs, these methods are also closely related, as both revolve around suitable control of the localization process: in the algorithmic stochastic localization approach, this is used to construct a mean estimator for the localized measures, and in our approach it is used to bound the localized measures' covariances. These tasks are closely linked; see Remark 7.7.4.

One question is whether simulated annealing succeeds in more general models. In particular, samplers based on algorithmic stochastic localization have been developed for the Sherrington–Kirkpatrick model in the replica symmetric regime [AMS22, Cel24], p-spin models over the hypercube [AMS23b], and posteriors of spiked matrix models [MW23]. These samplers are proven to have vanishing Wasserstein error, and sampling with vanishing total variation error remains an open problem in these models. It would be interesting to show that simulated annealing achieves this. More speculatively, we may ask if there is a general reduction from a sampling guarantee for algorithmic stochastic localization to one for simulated annealing.

#BIS. A major open problem in the field of approximate counting is settling the complexity of #BIS: where the algorithmic task is to approximate the number of independent sets in a bipartite graph. So far, algorithmic progress for this problem has been limited to restricted classes of graphs, such as lattices & tori [HPR19], and expander graphs [JKP20]. Numerous interesting approximate counting problems have been shown to be #BIS-hard [CGM12, GJ12, CGG+16, GSVY16]. While vanilla Glauber dynamics fails at the corresponding sampling task, it is plausible that a variant of simulated annealing succeeds.

Problem 7.1.15. Does (a simple variant of) simulated annealing succeed at sampling a uniformly random independent set in a bipartite graph?

Structural guarantees from weak Poincaré inequalities. According to physics heuristics, the Gibbs measure of a spherical mixed p-spin glass between β_{uniq} and β_{sh} consists of one main cluster accounting for nearly all the mass, and metastable states with exponentially small mass that are poorly connected to the main cluster and each other. We do not prove this picture, but the weak Poincaré inequality we obtain (up to β_{SL}) is sufficient to imply a sampling guarantee for simulated annealing. One open direction is to show that the above picture holds, and that the main cluster satisfies a genuine Poincaré inequality. More generally, one may ask:

Problem 7.1.16. If a distribution satisfies a weak Poincaré inequality, is it TV-close to a distribution satisfying a true Poincaré inequality?

We note that Lemmas 7.4.9 and 7.A.5 show a converse of this statement, that if we perturb a distribution satisfying a true Poincaré inequality (for the Langevin diffusion or Glauber dynamics Markov chains), the resulting distribution satisfies a weak Poincaré inequality.

7.1.6 Organization

In Section 7.2, we give a technical overview of how we use weak Poincaré inequalities to analyze simulated annealing for our main application of sampling from spherical p-spin distributions.

In Section 7.3, we cover some basic preliminaries that will be useful. Then, in Section 7.4, we formally define weak functional inequalities and establish some of their basic properties.

In Section 7.5, we demonstrate the effectiveness of this framework by showing how to sample from a mixture of distributions satisfying Poincaré inequalities from data-based initializations.

Our main application to spherical p-spin models spans Sections 7.6 to 7.8, and requires more background in stochastic localization and spin glass theory. In Section 7.6, we review some basic properties of stochastic localization and show how to adapt the framework of localization schemes from [CE22] to prove weak functional inequalities. Then, in Section 7.7, we initiate the discussion of weak Poincaré inequalities for spherical p-spin models. To assist the reader in understanding the proof of a weak Poincaré inequality, we provide a separate technical overview in Section 7.7.1. The rest of Section 7.7 reduces the proof to proving high-probability covariance bounds for strictly replica-symmetric models with small external field, which is then established in Section 7.8.

Acknowledgements

We are extremely grateful to Andrea Montanari and Huy Tuan Pham for early discussions on this problem, and to Ahmed El Alaoui, Sitan Chen, Sinho Chewi, Jason Gaitonde, Kuikui Liu, Francisco Pernice, and Mark Sellke for insightful discussions. We would like to thank Thiago Bergamaschi for pointing out an error in an application to the ferromagnetic Potts model in an earlier version of this paper.

7.2 Technical overview

Let H be a Hamiltonian on state space Ω, and let μ_H be its Gibbs distribution. Our goal in this section is to describe our strategy to prove that simulated annealing succeeds at sampling. In our application, Ω is the scaled sphere $S_N \coloneqq \sqrt{N} \cdot \mathbb{S}^{N-1}$, and μ_H comes with an associated Markov chain known as *Langevin diffusion*, which we denote with P_H. For ease of exposition, we restrict the discussion to this setting, though much of it holds in a more general setting.

Definition 7.2.1 (Simulated annealing, informal)**.** Initialize at the uniform distribution on S_N (which is equal to μ_0), and for each $i \in [m]$, run $P_{\frac{i}{m}H}$ for time T.

For the sequel, we abbreviate $P_{\frac{i}{m}H}$ and $\mu_{\frac{i}{m}H}$ as P_i and μ_i, and we use $P_{i,t}$ to denote running P_i for time t. The strategy to prove that simulated annealing succeeds at sampling is to establish a weak Poincaré inequality for P_i for all i.

Let L be the infinitesimal generator of P_i. For functions $f, g : \Omega \to \mathbb{R}$, we define the *Dirichlet form* $\mathcal{E}(f, g)$ as $\mathbb{E}_{\mu_i}[f \,\mathrm{L}\, g]$.

Remark 7.2.2. In the case of Langevin diffusion for a distribution π, the Dirichlet form can be evaluated as

$$\mathcal{E}(f, g) = \mathop{\mathbb{E}}_{\mu_i} \langle \nabla f, \nabla g \rangle \,,$$

where ∇ denotes the Euclidean gradient if π is supported on $\mathbb{R}^N$, and the Riemannian gradient on S_N if π is supported on S_N.

As discussed in Section 7.1, we say a Markov chain satisfies a *weak Poincaré inequality* with parameters (C, ε) if

$$\mathcal{E}(f, f) \geq C \cdot \mathsf{Var}[f] - \varepsilon \cdot \|f - \mathbb{E}\, f\|_\infty^2 - \varepsilon \cdot \sup_{x \in \Omega} \|\nabla f(x)\|^2 \,,$$

which implies the following mixing result Theorem 7.4.6 for the chi-squared divergence; see also [RW01, Theorem 2.1]. Defining ν_t as the distribution after running the Markov chain for time t from initial distribution ν_0, we have

$$\chi^2(\nu_t \| \mu_i) \leq \exp(-Ct) \cdot \chi^2(\nu_0 \| \mu_i) + \varepsilon \cdot \left\| \frac{\mathsf{d}\nu_0}{\mathsf{d}\mu_i} - 1 \right\|_\infty^2 + \varepsilon \cdot \sup_{x \in \Omega} \left\| \nabla \frac{\mathsf{d}\nu_0}{\mathsf{d}\mu_i}(x) \right\|^2 \,.$$

Analyzing simulated annealing with weak Poincaré inequalities. To see why the above statement plays well with simulated annealing, imagine plugging in initialization $\nu_0 = \mu_{i-1}$. By selecting the number of annealing steps $m = \mathsf{poly}(N)$, we can ensure $\left\| \frac{\mathsf{d}\nu_0}{\mathsf{d}\mu_i} - 1 \right\|_\infty$ and $\sup_{x \in \Omega} \left\| \nabla \frac{\mathsf{d}\nu_0}{\mathsf{d}\mu_i}(x) \right\|$ are $O(1)$. The guarantee after running the Markov chain for some sufficiently large polynomial time T is then

$$\chi^2(P_{i,T}\mu_{i-1} \| \mu_i) \leq O(\varepsilon) \,,$$

which in particular implies

$$\mathrm{TV}(P_{i,T}\mu_{i-1}, \mu_i) \leq O(\sqrt{\varepsilon}) \,.$$

When we combine the above with the data processing inequality, we then get the following guarantee for $\nu_{m,T} \coloneqq P_{m,T} \cdots P_{2,T} P_{1,T} \nu_0$, the distribution that simulated annealing samples from.

$$\begin{aligned} \mathrm{TV}(P_{m,T} \cdots P_{1,T}\mu_0, \mu_m) &\leq \mathrm{TV}(P_{m,T} \cdots P_{1,T}\mu_0, P_{m,T}\mu_{m-1}) + \mathrm{TV}(P_{m,T}\mu_{m-1}, \mu_m) \\ &\leq \mathrm{TV}(P_{m-1,T} \cdots \mu_0, \mu_{m-1}) + O(\sqrt{\varepsilon}) \,. \end{aligned}$$

Applying the above inequality m times tells us that $\mathrm{TV}(\nu_{m,T}, \mu_m) \leq O(\sqrt{\varepsilon} \cdot m)$.

We now turn our attention to the proof technique for showing a weak Poincaré inequality.

How to prove weak Poincaré inequalities. Suppose our goal is to prove a weak Poincaré inequality for a measure π. The high-level strategy in the localization schemes approach for proving a weak Poincaré inequality is to design a *measure decomposition* of π: for some mixture distribution ρ, express π as $\mathbb{E}_{\boldsymbol{z} \sim \rho}\, \pi_{\boldsymbol{z}}$. Refer to Section 7.3.1 for a brief review of measure decompositions. Once we have a measure decomposition in hand, establishing the following simple set of inequalities forms the crux of the argument. Let f be a function such that $\mathbb{E}_\pi\, f = 1$.

1. *Conservation of Dirichlet form.*

$$\mathcal{E}_\pi(f, f) \geq \mathop{\mathbb{E}}_{\boldsymbol{z} \sim \rho} \mathcal{E}_{\pi_{\boldsymbol{z}}}(f, f) \,.$$

In the case of Langevin diffusion, this is an equality, and in the case of Glauber dynamics, the inequality is true by a generic concavity argument; see, e.g. [AJK+22, Page 19] or [LMRW24, Page 5].

2. *Weak Poincaré inequality for good component measures.*

$$\mathop{\mathbb{E}}_{\boldsymbol{z}\sim\rho} \mathcal{E}_{\pi_{\boldsymbol{z}}}(f,f) \geq C \cdot \mathop{\mathbb{E}}_{\boldsymbol{z}\sim\rho} \mathsf{Var}_{\pi_{\boldsymbol{z}}}[f] \cdot \mathbf{1}[\boldsymbol{z} \text{ “good”}] - \varepsilon\|f-1\|_\infty^2 - \varepsilon\|\nabla f\|_\infty^2 \,,$$

   where the “good” $\pi_{\boldsymbol{z}}$ are those which satisfy a (c,ε)-weak Poincaré inequality. This inequality follows from the nonnegativity of norms and Dirichlet forms.

3. *Approximate conservation of variance.*

$$\mathop{\mathbb{E}}_{\boldsymbol{z}\sim\rho} \mathsf{Var}_{\pi_{\boldsymbol{z}}}[f] \geq \alpha \cdot \mathsf{Var}_{\pi}[f] \,.$$

   This is one of the parts that depends nontrivially on π and the decomposition ρ, and we discuss the general proof strategy for this portion based on localization schemes.

4. *High-probability goodness of component measures.*

$$\Pr_{\boldsymbol{z}\sim\rho}[\boldsymbol{z} \text{ “good”}] \geq 1-\varepsilon \,.$$

   This part also requires analyzing the measure decomposition we design. Ideally, the measure decomposition presents us with “simpler” measures than π itself.

Once we have the above inequalities at hand, we get a $(c\alpha, 2\varepsilon)$-weak Poincaré inequality; see Lemma 7.4.10 for details.

How to construct a good measure decomposition. Henceforth, we restrict our attention to the case where $\pi = \mu_H$, the Gibbs distribution for a spherical mixed p-spin glass model. In the discussion below, we fix H as a typical Hamiltonian, and drop the phrase “with high probability” for events that occur with high probability over the randomness of H.

To construct our measure decomposition, we rely on Eldan's stochastic localization [Eld13]. Our inspiration is the use of stochastic localization as a tool for measure decomposition for proving Poincaré inequalities in the work of Chen and Eldan [CE22]. Stochastic localization is a measure-valued random process $(\mu_t)_{t\geq 0}$ described by:

$$\mathsf{d}\mu_t(x) \propto \exp(\langle \boldsymbol{y}_t, x\rangle - \frac{t}{2}\|x\|^2)\mathsf{d}\mu_H(x) \,,$$

where $\boldsymbol{y}_t = \boldsymbol{\sigma} + B_t$ where $\boldsymbol{\sigma} \sim \mu_H$ and $(B_t)_{t\geq 0}$ is a standard Brownian motion; see [AM22, Theorem 2] for a proof of why the above description of stochastic localization is equivalent to the more traditional definition via a stochastic differential equation that μ_t obeys.

We run stochastic localization up to a stopping time τ, defined as

$$\tau \coloneqq \min\{t : 0 \leq t \leq T,\, \|\mathsf{Cov}(\mu_t)\| \geq K \text{ or } t = T\} \,,$$

where T is chosen as a sufficiently large constant, independent of N. We impose the constraint on the covariance matrix as it is relevant to satisfying approximate conservation of variance: [CE22, Eq. (20)] proves that a measure decomposition based on stochastic localization run for time at most T satisfies approximate conservation of variance with parameter $\alpha = \exp(-KT)$ if $\|\mathsf{Cov}(\mu_t)\|_{\mathsf{op}}$ is bounded by K almost surely. Hence, by construction, we automatically ensure that our measure decomposition satisfies approximate conservation of variance.

For the measure decomposition to ultimately be useful, we also need to argue that the component measures satisfy a weak Poincaré inequality with high probability. Building on technical results in Huang, Montanari, and Pham [HMP24, Section 9.2], we show that the stochastic localization process run up to time T starting at μ_H gives a distribution satisfying an $(\Omega(1), \exp(-\Omega(n)))$-weak Poincaré inequality with probability $1-\exp(-\Omega(n))$ over the randomness of the stochastic localization path; see Lemma 7.7.8 for details. Unfortunately, in the situation where the stochastic localization process stops before T, we do not have a simple way to show a weak Poincaré inequality, and for our analysis, treat $\boldsymbol{z}$ arising from early stopping as “bad”.

Thus, we have: $\Pr_{\boldsymbol{z}\sim\rho}[\boldsymbol{z} \text{ “good”}] \geq 1 - \exp(-\Omega(n)) - \Pr[\tau < T]$. To bound $\Pr[\tau < T]$, it is sufficient to prove a high-probability covariance norm bound on the entire stochastic localization path for $0 \leq t \leq T$. Most of the technical work in this paper is devoted to proving this covariance norm bound.

Theorem 7.2.3 (Informal version of Lemma 7.7.7)**.** *For a typical H, with probability $1 - e^{-\Omega(n^{1/5})}$ over the randomness of the stochastic localization path, we have $\|\mathsf{Cov}(\mu_t)\|_{\mathsf{op}} < K$.*

The proof of the covariance bound spans Sections 7.7 and 7.8; we give a detailed technical overview of how it is proved in Section 7.7.1.

7.3 Preliminaries

Notation

- We use S_N to denote the scaled $(N-1)$-sphere, $\sqrt{N} \cdot \mathbb{S}^{N-1}$.
- We use ρ to denote the uniform measure over S_N.
- Given $\sigma_1, \sigma_2 \in S_N$, we use $R(\cdot, \cdot)$ to denote the normalized inner product (i.e. the *overlap*)

$$R(\sigma_1, \sigma_2) \coloneqq \frac{\langle \sigma_1, \sigma_2 \rangle}{N}.$$

- For an interval $I \subseteq [-1, 1]$ and $\boldsymbol{x} \in S_N$, define $\mathsf{Band}(\boldsymbol{x}, I) \coloneqq \{\sigma \in S_N : R(\sigma, \boldsymbol{x}) \in I\}$.
- We use c to denote small constants whose values may change from line to line, and C to denote similarly fickle large constants.
- Let $f : \Omega \to \mathbb{R}$ be any function. We define $\operatorname{osc}(f) \coloneqq \sup f - \inf f$.
- Let $f : \Omega \to \mathbb{R}$ be a smooth function. If $\Omega \subseteq \mathbb{R}^N$, then ∇f denotes its Euclidean gradient. If $\Omega \subseteq S_N$, then $\nabla_{\mathsf{sp}} f$ denotes the Riemannian gradient on S_N. When the correct notion of gradient is clear from context, by an abuse of notation we will suppress this distinction and simply write ∇f.

7.3.1 Measure decompositions

Our framework for proving weak functional inequalities relies on the notion of a measure decomposition.

Definition 7.3.1 (Measure decomposition)**.** Let π be a distribution on $\mathbb{R}^N$. Let ρ be a mixture distribution, also on $\mathbb{R}^N$, which indexes into a family of mixture components $\{\pi_z\}_{z \in \mathbb{R}^N}$. We say that (ρ, π_z) is a measure decomposition for π if

$$\pi = \mathop{\mathbb{E}}_{\boldsymbol{z} \sim \rho} \pi_{\boldsymbol{z}} \,.$$

One reason measure decompositions are useful is that they compose nicely with worst-case functional inequalities, as shown in the following lemma.

Lemma 7.3.2 ([BB19, AJK+22, CE22])**.** *Let π be a distribution over $\Omega \subseteq \mathbb{R}^N$, and $\pi = \mathbb{E}_{\boldsymbol{z} \sim \rho} \pi_{\boldsymbol{z}}$ a measure decomposition of π such that*

- *for all functions f, $\mathbb{E}_{\boldsymbol{z} \sim \rho} \mathsf{Var}_{\pi_{\boldsymbol{z}}}[f] \geq C_{\mathrm{Var}} \mathsf{Var}_{\pi}[f]$, and*
- *Every $\pi_{\boldsymbol{z}}$ satisfies a ρ_{PI}-Poincaré inequality with respect to Langevin diffusion.*

Then, π satisfies a $\rho_{\mathrm{PI}} C_{\mathrm{Var}}$-Poincaré inequality.

In Lemma 7.4.10, we will show an average-case relaxation of the above result, that π satisfies a weak Poincaré inequality if *most* measures in the decomposition satisfy weak Poincaré inequalities. Then, in Section 7.6, we construct explicit measure decompositions using the localization schemes framework introduced in [CE22]. This will show weak Poincaré inequalities for our measures of interest.

Besides proving functional inequalities, measure decompositions have also been directly used for sampling and inference (see, e.g., [MW24, LMR+24]).

7.3.2 Langevin diffusion

In this paper, we study Langevin diffusion on $\mathbb{R}^N$ and the scaled sphere S_N. These definitions can be directly generalized to the setting of Riemannian manifolds, but we do not comment further on this.

Definition 7.3.3 (Langevin diffusion on $\mathbb{R}^N$)**.** Let π be a distribution on $\mathbb{R}^N$ with density at x proportional to $e^{-V(x)}$ for some function V. The *Langevin diffusion* process with stationary distribution π is the solution to the stochastic differential equation

$$\mathsf{d}Z_t = -\nabla V(Z_t)\mathsf{d}t + \sqrt{2}\mathsf{d}B_t,$$

where $(B_t)_{t\geq 0}$ is a standard Brownian motion.

Definition 7.3.4 (Langevin diffusion on S_N)**.** Let π be a distribution on S_N with $\mathsf{d}\pi(x) \propto e^{-V(x)}\mathsf{d}\rho(x)$, where $V : S_N \to \mathbb{R}$. The Langevin diffusion process with stationary distribution π is the solution to the stochastic differential equation

$$\mathsf{d}Z_t = -\nabla_{\mathsf{sp}} V(Z_t)\mathsf{d}t + \sqrt{2}\mathsf{d}B_t,$$

where $(B_t)_{t\geq 0}$ is a standard spherical Brownian motion. (For a textbook introduction to spherical Brownian motion, see [Hsu02].)

Fact 7.3.5 ([Che23b, Example 1.2.17])**.** *The Langevin diffusion SDE with stationary distribution π is reversible with respect to π. In particular, the ergodicity of the process implies that $\mathsf{KL}\mathcal{L}(Z_t)\pi \xrightarrow{t\to\infty} 0$.*

Furthermore, it is well-known that Langevin diffusion on $\mathbb{R}^N$ with respect to a strongly log-concave stationary distribution converges rapidly.

Definition 7.3.6. Let π be a distribution over $\mathbb{R}^N$ with density proportional to e^{-V}. π is said to be *α-strongly log-concave* if V is α-strongly convex, that is, $\nabla^2 V \succeq \alpha I$.

Fact 7.3.7 ([Che23b, Theorem 1.2.24])**.** *Let π be a distribution satisfying a log-Sobolev inequality with constant ρ_{LS}, in that for any differentiable function $f : \mathbb{R}^N \to \mathbb{R}_{>0}$,*

$$\mathop{\mathbb{E}}_{\pi} \|\nabla\sqrt{f}\|^2 \geq \rho_{\mathrm{LS}} \mathsf{Ent}_\pi[f].$$

Then, if π_t is the distribution at time t of Langevin diffusion,

$$\mathsf{KL}\pi_t\pi \leq \mathsf{KL}\pi_0\pi e^{-\rho_{\mathrm{LS}}\cdot t}.$$

Furthermore, α-strongly log-concave distributions π satisfy a log-Sobolev inequality with constant α.

7.4 Weak functional inequalities

In this paper, we study continuous-time Markov chains.

Definition 7.4.1 (Markov semigroup)**.** Let $(X_t)_{t\geq 0}$ denote a continuous-time Markov process on *state space* Ω. Let $(P_t)_{t\geq 0}$ be the associated *Markov semigroup* operator; P_t acts on functions $f : \Omega \to \mathbb{R}$ via $P_t f(x) = \mathbb{E}[f(X_t)|X_0 = x]$. Throughout, we assume that the semigroup is *reversible* with respect to stationary distribution π. Furthermore, let L denote the *infinitesimal generator* of P_t, i.e., $P_t = e^{-t\mathrm{L}}$. For functions $f, g : \Omega \to \mathbb{R}$, we define the *Dirichlet form* as $\mathcal{E}(f, g) = \mathbb{E}_\pi[f\mathrm{L}g]$.

See e.g. [Che23b, Section 1.2] for a textbook treatment. Of particular interest to us are the two settings where the semigroup corresponds to a discrete-time Markov chain or the Langevin diffusion defined in Section 7.3.2. In these cases, the Dirichlet form satisfies the following explicit identities.

Fact 7.4.2 (Dirichlet form from discrete-time Markov chain)**.** *Let P be the transition matrix of a reversible discrete-time Markov chain with stationary distribution π. We can define an associated continuous-time semigroup operator $(P_t)_{t\geq 0}$ by setting $\mathrm{L} = I - P$. The Dirichlet form for the continuous-time dynamics satisfies*

$$\mathcal{E}(f, g) := \mathop{\mathbb{E}}_{\boldsymbol{x}\sim\pi} \mathop{\mathbb{E}}_{\boldsymbol{y}\sim_P \boldsymbol{x}} (f(\boldsymbol{x}) - f(\boldsymbol{y}))(g(\boldsymbol{x}) - g(\boldsymbol{y}))\,.$$

Here, for a probability distribution μ, we say $\boldsymbol{x} \sim \mu$ to denote a sample $\boldsymbol{x}$ from μ, and we use $\boldsymbol{y} \sim_P \boldsymbol{x}$ for a single transition from $\boldsymbol{x}$ according to P.

Fact 7.4.3 (Dirichlet form for Langevin diffusion)**.** *We will need the following explicit identities for the Dirichlet form for Langevin diffusion.*

(1) *When $(P_t)_{t\geq 0}$ corresponds to Langevin diffusion on $\mathbb{R}^N$ with stationary distribution π, the Dirichlet form is $\mathcal{E}(f,g) = \mathbb{E}_\pi[\langle \nabla f, \nabla g\rangle]$.*

(2) *When $(P_t)_{t\geq 0}$ corresponds to Langevin diffusion on S_N with stationary distribution π, the Dirichlet form is $\mathcal{E}(f,g) = \mathbb{E}_\pi[\langle \nabla_{\mathsf{sp}} f, \nabla_{\mathsf{sp}} g\rangle]$.*

Definition 7.4.4. We say π satisfies a *weak Poincaré inequality* if for some error functional $\mathsf{Error} : \mathbb{R}^{\Omega}_{>0} \to \mathbb{R}_{\geq 0}$ and $\rho_{\mathrm{PI}} > 0$,

$$\mathsf{Var}_\pi[f] \leq \frac{1}{\rho_{\mathrm{PI}}} \cdot \mathcal{E}(f,f) + \mathsf{Error}(f)\,.$$

Similarly, we say π satisfies a *weak modified log-Sobolev inequality* if for some error functional $\mathsf{Error} : \mathbb{R}^{\Omega}_{>0} \to \mathbb{R}_{\geq 0}$ and $\rho_{\mathrm{LS}} \geq 0$,

$$\mathsf{Ent}_\pi[f] \leq \frac{1}{\rho_{\mathrm{LS}}} \cdot \mathcal{E}(f, \log f) + \mathsf{Error}(f).$$

Theorem 7.4.5. *Consider the trajectory $(\nu_t)_{t\geq 0}$ of a reversible continuous-time Markov chain with stationary distribution π, initialized at the distribution ν_0, and suppose that π satisfies a weak MLSI with parameters* Error *and ρ_{LS}. Fix $T > 0$, and set Λ_T to be the distribution on $[0,T]$ with density $\Lambda_T(s) = \frac{e^{\rho_{\mathrm{LS}}}}{e^{\rho_{\mathrm{LS}}T}-1} \cdot e^{\rho_{\mathrm{LS}}s}$. Then,*

$$\mathsf{KL}\nu_T\pi \leq e^{-\rho_{\mathrm{LS}}T}\mathsf{KL}\nu_0\pi + \mathop{\mathbb{E}}_{\boldsymbol{s}\sim\Lambda_T}[\mathsf{Error}(\tfrac{\mathsf{d}\nu_{\boldsymbol{s}}}{\mathsf{d}\pi})].$$

Proof. Let $f_0 = \frac{\mathsf{d}\nu_0}{\mathsf{d}pi}$, and let $f_t = P_t f_0 = \frac{\mathsf{d}\nu_t}{\mathsf{d}\pi}$ (this last equality holds due to reversibility). For ease of notation, set $\mathsf{Error}_t = \mathsf{Error}(f_t)$ for $t \geq 0$. Recalling that $\mathcal{E}(f_t, \log f_t) = -\frac{\mathsf{d}}{\mathsf{d}t}\mathsf{KL}\nu_t\pi$, the weak MLSI says that

$$-\mathcal{E}(f_t, \log f_t) + \rho_{\mathrm{LS}} \cdot \mathsf{KL}\nu_t\pi - \rho_{\mathrm{LS}} \cdot \mathsf{Error}_t \leq 0,$$

so

$$\frac{\mathsf{d}}{\mathsf{d}t}\left(e^{\rho_{\mathrm{LS}}t} \cdot \mathsf{KL}\nu_t\pi - \rho_{\mathrm{LS}}\int_0^t e^{\rho_{\mathrm{LS}}s}\mathsf{Error}_s \mathsf{d}s\right) \leq 0.$$

Therefore,

$$e^{\rho_{\mathrm{LS}}T} \cdot \mathsf{KL}\nu_T\pi - \rho_{\mathrm{LS}}\int_0^T e^{\rho_{\mathrm{LS}}s}\mathsf{Error}_s \mathsf{d}s \leq \mathsf{KL}\nu_0\pi,$$

and

$$\mathsf{KL}\nu_T\pi \leq e^{-\rho_{\mathrm{LS}}T}\mathsf{KL}\nu_0\pi + \rho_{\mathrm{LS}}\int_0^T e^{\rho_{\mathrm{LS}}(s-T)}\mathsf{Error}_s.$$

Noting that $\Lambda_T(s) = \frac{\rho_{\mathrm{LS}}}{e^{\rho_{\mathrm{LS}}T}-1} \cdot e^{\rho_{\mathrm{LS}}s} \geq \rho_{\mathrm{LS}}e^{\rho_{\mathrm{LS}}(s-T)}$, the above implies that

$$\mathsf{KL}\nu_T\pi \leq e^{-\rho_{\mathrm{LS}}T}\mathsf{KL}\nu_0\pi + \mathop{\mathbb{E}}_{s\sim\Lambda_T}[\mathsf{Error}_s]\,,$$

as desired. $\square$

By essentially the same proof, we obtain the analogous result for weak Poincaré inequalities.

Theorem 7.4.6. *Consider the trajectory $(\nu_t)_{t\geq 0}$ of a (continuous-time) Markov chain with stationary distribution π, initialized at the distribution ν_0, and suppose that π satisfies a weak Poincaré inequality with parameters* Error *and ρ_{PI}. Fix $T > 0$, and set Λ_T to be the distribution on $[0,T]$ with density $\Lambda_T(s) = \frac{e^{2\rho_{\mathrm{PI}}}}{e^{2\rho_{\mathrm{PI}}T}-1} \cdot e^{2\rho_{\mathrm{PI}}s}$. Then,*

$$\chi^2(\nu_T\|\pi) \leq e^{-2\rho_{\mathrm{PI}}T}\chi^2(\nu_0\|\pi) + \mathop{\mathbb{E}}_{\boldsymbol{s}\sim\Lambda_T}[\mathsf{Error}(\tfrac{\mathsf{d}\nu_{\boldsymbol{s}}}{\mathsf{d}\pi})].$$

For the analysis of the annealed Langevin dynamics, we will also require the following definition. For $f : \Omega \to \mathbb{R}$, let $\operatorname{osc}(f) \coloneqq \sup(f) - \inf(f)$, and let ∇f denote the Riemannian gradient.

Definition 7.4.7 (Weak functional inequalities for Langevin)**.** We say a distribution π on $\Omega \subseteq \mathbb{R}^N$ or $\Omega \subseteq S_N$ satisfies a $(\rho_{\mathrm{PI}}, \varepsilon)$-weak Poincaré inequality if for all differentiable functions f,

$$\mathsf{Var}_\pi[f] \leq \frac{1}{\rho_{\mathrm{PI}}} \cdot \mathcal{E}(f,f) + \varepsilon \cdot (\mathrm{osc}(f)^2 + \sup_{x \in \Omega} \|\nabla f\|^2).$$

Similarly, we say π satisfies a $(\rho_{\mathrm{LS}}, \varepsilon)$-weak modified log-Sobolev inequality if for all differentiable functions f,

$$\mathsf{Ent}_\pi[f] \leq \frac{1}{\rho_{\mathrm{LS}}} \cdot \mathcal{E}(f, \log f) + \varepsilon \cdot (\mathrm{osc}(\sqrt{f})^2 + \sup_{x \in \Omega} \|\nabla f\|^2).$$

Remark 7.4.8. As mentioned in the beginning of this section, by replacing the Riemannian gradient with the discrete gradient, an analogous theory can be developed for annealed Glauber dynamics; see Definition 7.A.1.

We shall typically use weak Poincaré inequalities with functions f that have expectation 1, where we bound $\mathrm{osc}(f) \leq 2\|f-1\|_\infty$.

7.4.1 Properties of weak functional inequalities

In this section, we state some crucial properties of weak functional inequalities for Langevin diffusion on $\mathbb{R}^N$ or S_N. With minor modifications, the same results hold for Glauber dynamics on finite state spaces; see Section 7.A for formal details.

Lemma 7.4.9. *Let π be a distribution on $\mathbb{R}^N$ or S_N satisfying a ρ_{PI}-Poincaré inequality for Langevin diffusion, and π' a distribution such that $\mathrm{TV}(\pi, \pi') \leq \delta$. Then, π' satisfies a $\big(\rho_{\mathrm{PI}}, \delta \max(\rho_{\mathrm{PI}}^{-1}, 1)\big)$-weak Poincaré inequality for Langevin diffusion.*

Proof. There exists a coupling $\mathcal{C}$ of (π, π') such that for $(x, x') \sim \mathcal{C}$, $\Pr[x \neq x'] \leq \delta$. Thus,

$$\begin{aligned}
\mathcal{E}_{\pi'}(f,f) &= \mathop{\mathbb{E}}_{\pi'} \|\nabla f\|^2 \\
&\geq \mathop{\mathbb{E}}_{\pi} \|\nabla f\|^2 - \delta \sup \|\nabla f\|^2 \\
&\geq \rho_{\mathrm{PI}} \mathsf{Var}_\pi[f] - \delta \sup \|\nabla f\|^2.
\end{aligned}$$

Let $I = [\inf f, \sup f]$. Note that $\mathsf{Var}_\pi[f] = \inf_{a \in I} \mathbb{E}_\pi[(f-a)^2]$. For each $a \in I$,

$$\mathop{\mathbb{E}}_{\pi}[(f-a)^2] \geq \mathop{\mathbb{E}}_{\pi'}[(f-a)^2] - \delta \cdot \mathrm{osc}(f)^2,$$

and therefore

$$\mathsf{Var}_\pi[f] \geq \mathsf{Var}_{\pi'}[f] - \delta \cdot \mathrm{osc}(f)^2. \tag{7.3}$$

Combining with the above shows

$$\mathcal{E}_{\pi'}(f,f) \geq \rho_{\mathrm{PI}} \mathsf{Var}_{\pi'}[f] - \delta \left(\rho_{\mathrm{PI}} \cdot \mathrm{osc}(f)^2 + \sup \|\nabla f\|^2 \right). \qquad \square$$

As foreshadowed previously, measure decompositions compose well with weak functional inequalities. Indeed, the following lemma can be viewed as a relaxation of the setup to prove genuine functional inequalities (cf. Lemma 7.3.2).

Lemma 7.4.10. *Let π be a distribution over $\mathbb{R}^N$ or S_N, and $\pi = \mathbb{E}_{\boldsymbol{z} \sim \rho} \pi_{\boldsymbol{z}}$ a measure decomposition of π such that*

- *for all functions f, $\mathbb{E}_{\boldsymbol{z} \sim \rho} \mathsf{Var}_{\pi_{\boldsymbol{z}}}[f] \geq C_{\mathrm{Var}} \mathsf{Var}_\pi[f]$, and*
- *with probability $1-\eta$ over $\boldsymbol{z} \sim \rho$, $\pi_{\boldsymbol{z}}$ satisfies a $(\rho_{\mathrm{PI}}, \delta)$-weak Poincaré inequality with respect to Langevin diffusion.*

Then, π satisfies a $\left(\rho_{\mathrm{PI}} C_{\mathrm{Var}}, \frac{\delta+\eta}{C_{\mathrm{Var}}}\right)$-weak Poincaré inequality.

Proof. Let us say that $\boldsymbol{z}$ is *good* if $\pi_{\boldsymbol{z}}$ satisfies a weak Poincaré inequality, and f be a function. Then,

$$\begin{aligned}
\mathcal{E}_\pi(f,f) &= \mathop{\mathbb{E}}_{\boldsymbol{z}\sim\rho} \mathcal{E}_{\pi_{\boldsymbol{z}}}(f,f) \\
&\geq \mathop{\mathbb{E}}_{\boldsymbol{z}\sim\rho} \mathcal{E}_{\pi_{\boldsymbol{z}}}(f,f)\mathbf{1}_{\boldsymbol{z}\text{ is good}} \\
&\geq \mathop{\mathbb{E}}_{\boldsymbol{z}\sim\rho} \rho_{\mathrm{PI}}\mathsf{Var}_{\pi_{\boldsymbol{z}}}[f]\mathbf{1}_{\boldsymbol{z}\text{ is good}} - \delta\rho_{\mathrm{PI}}\cdot(\operatorname{osc}(f)^2+\sup\|\nabla f\|^2) \\
&= \mathop{\mathbb{E}}_{\boldsymbol{z}\sim\rho} \rho_{\mathrm{PI}}\mathsf{Var}_{\pi_{\boldsymbol{z}}}[f] - \delta\rho_{\mathrm{PI}}\cdot(\operatorname{osc}(f)^2+\sup\|\nabla f\|^2) - \mathop{\mathbb{E}}_{\boldsymbol{z}\sim\rho} \rho_{\mathrm{PI}}\mathsf{Var}_{\pi_{\boldsymbol{z}}}[f]\mathbf{1}_{\boldsymbol{z}\text{ is not good}} \\
&\geq \rho_{\mathrm{PI}} \mathop{\mathbb{E}}_{\boldsymbol{z}\sim\rho} \mathsf{Var}_{\pi_{\boldsymbol{z}}}[f] - (\delta\rho_{\mathrm{PI}}+\eta\rho_{\mathrm{PI}})\cdot(\operatorname{osc}(f)^2+\sup\|\nabla f\|^2) \\
&\geq C_{\mathrm{Var}}\rho_{\mathrm{PI}}\mathsf{Var}_\pi[f] - (\delta\rho_{\mathrm{PI}}+\eta\rho_{\mathrm{PI}})\cdot(\operatorname{osc}(f)^2+\sup\|\nabla f\|^2)\,.
\end{aligned}$$

The desired follows. □

7.4.2 Weak Poincaré inequalities and annealed Markov chains

The notion of weak functional inequalities defined in Definition 7.4.7 can be naturally applied in the context of simulated annealing, which we now define.

Definition 7.4.11 (Annealing scheme)**.** Let H be a Hamiltonian over Ω, and $(\mu_\beta)_{\beta\geq 0}$ the class of distributions over Ω with $\mu_\beta(\sigma) \propto e^{\beta H(\sigma)}$. For each $\beta \leq \beta_0$, let $P = P_\beta$ be a (reversible and ergodic) Markov chain with stationary distribution μ_β.

An *(inverse) temperature schedule* is any function $\beta : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$. An *annealing scheme* $\mathcal{A}$ is the time-inhomogeneous Markov chain such that at time t, one applies the Markov chain $P_{\beta(t)}$.

Of interest is the temperature schedule of the form $t \mapsto \delta\cdot\lfloor\frac{t}{T}\rfloor$, with the chain being run for time $T\cdot\left(\frac{\beta_0}{\delta}+1\right)$.

Theorem 7.4.12. *Let $T,\delta > 0$ such that $k_0 \coloneqq \frac{\beta_0}{\delta}$ is an integer. Suppose that for each $\beta = k\delta$ for $0 \leq k \leq k_0$, μ_β satisfies a $(\rho_{\mathrm{PI}},\varepsilon)$-weak Poincaré inequality for P_β. Consider the annealing scheme given by schedule $t\mapsto\delta\cdot\lfloor\frac{t}{T}\rfloor$, run for total time $T\cdot\left(\frac{\beta_0}{\delta}+1\right)$. Let ν be the output distribution of this annealing scheme. Then,*

$$\mathrm{TV}(\nu,\mu_{\beta_0}) \leq \frac{\beta_0}{\delta}\cdot\left[(1+\delta\sup\|\nabla H\|)e^{2\delta\|H\|_\infty}-1\right]\cdot O\left(e^{-2\rho_{\mathrm{PI}}T}+\varepsilon\right)^{1/2}.$$

Remark 7.4.13. Setting $\varepsilon = 0$ and $\delta = \beta_0$ matches the guarantees of [CE22] (after applying Pinsker's inequality).

Proof. We shall prove the above using a simple inductive argument – our goal will be to show that initialized at μ_β, the $P_{\beta+\delta}$ Markov chain run for time T yields a distribution sufficiently close (in total variation distance) to $\mu_{\beta+\delta}$. The total variation distance between the distribution that the annealed Markov chain outputs and the true distribution μ_{β_0} is then upper bounded by the sum of these total variation errors.

Let $\nu^{(r,k)}$ be the distribution obtained by running the annealed Markov chain initialized with $\mu_{r\delta}$ until inverse temperature $k\delta$. In particular, $\nu^{(r,k)}$ corresponds to the result of running our annealed Markov chain for $T(k-r)$ time, and $\nu^{(k.k)} = \mu_{k\delta}$. We are interested in bounding $\mathrm{TV}\big(\nu^{(0,k_0)},\mu_{\beta_0}\big)$. We have

$$\begin{aligned}
\mathrm{TV}\Big(\nu^{(0,k_0)},\mu_{\beta_0}\Big) &= \mathrm{TV}\Big(\nu^{(k_0-1,k_0)},\mu_{\beta_0}\Big) + \sum_{1\leq r\leq k_0-1}\Big(\mathrm{TV}\Big(\nu^{(r-1,k_0)},\mu_{\beta_0}\Big)-\mathrm{TV}\Big(\nu^{(r,k_0)},\mu_{\beta_0}\Big)\Big) \\
&\leq \mathrm{TV}\Big(\nu^{(k_0-1,k_0)},\mu_{\beta_0}\Big) + \sum_{1\leq r\leq k_0-1}\mathrm{TV}\Big(\nu^{(r-1,k_0)},\nu^{(r,k_0)}\Big) && \text{(Triangle inequality)} \\
&\leq \mathrm{TV}\Big(\nu^{(k_0-1,k_0)},\mu_{\beta_0}\Big) + \sum_{1\leq r\leq k_0-1}\mathrm{TV}\Big(\nu^{(r-1,r)},\nu^{(r,r)}\Big) && \text{(Data processing)} \\
&= \sum_{1\leq r\leq k_0}\mathrm{TV}\Big(\nu^{(r-1,r)},\mu_{r\delta}\Big).
\end{aligned}$$

We now turn to controlling the error functional $\operatorname{osc}(f)^2 + \sup \|\nabla f\|^2$. Fix an arbitrary β, and set f to be the likelihood ratio $\frac{\mathrm{d}\mu_\beta}{\mathrm{d}\mu_{\beta+\delta}}$. Then,

$$\begin{aligned}\|f-1\|_\infty &\leq \left\| \frac{e^{-\delta H}}{\mathbb{E}_{\mu_{\beta+\delta}} e^{-\delta H}} - 1 \right\|_\infty \\ &\leq \left\| \frac{e^{-\delta H}-1}{\mathbb{E}_{\mu_{\beta+\delta}} e^{-\delta H}} \right\|_\infty + \left| \frac{1}{\mathbb{E}_{\mu_{\beta+\delta}} e^{-\delta H}} - 1 \right| \\ &\leq \frac{e^{\delta\|H\|_\infty}-1}{e^{-\delta\|H\|_\infty}} + \frac{e^{\delta\|H\|_\infty}-1}{e^{-\delta\|H\|_\infty}} \leq 2\cdot(e^{2\delta\|H\|_\infty}-1).\end{aligned}$$

Hence, $\operatorname{osc}(f) \leq 4\cdot(e^{2\delta\|H\|_\infty}-1)$. Next, a simple computation yields

$$\begin{aligned}\|\nabla f\| &= \frac{\delta e^{-\delta H}}{\mathbb{E}_{\mu_{\beta+\delta}} e^{-\delta H}} \|\nabla H\| \\ &\leq 2\delta \cdot e^{2\delta\|H\|_\infty} \|\nabla H\|,\end{aligned}$$

so we have $\sup\|\nabla f\| \leq 2\delta\cdot e^{2\delta\|H\|_\infty} \sup\|\nabla H\|$.

Since each $\mu_{r\delta}$ satisfies a $(\rho_{\mathrm{PI}},\varepsilon)$-weak Poincaré inequality, Theorem 7.4.6 with the above calculation implies that

$$\begin{aligned}\mathrm{TV}\Big(\nu^{(r-1,r)},\mu_{r\delta}\Big)^2 &\leq \chi^2\Big(\nu^{(r-1,r)}\|\mu_{r\delta}\Big) \\ &\leq e^{-2\rho_{\mathrm{PI}}T}\cdot\chi^2\big(\mu_{(r-1)\delta}\|\mu_{r\delta}\big) + \varepsilon\cdot(16(e^{2\delta\|H\|_\infty}-1)^2 + 4(\delta e^{2\delta\|H\|_\infty}\sup\|\nabla H\|)^2) \\ &\leq (16(e^{2\delta\|H\|_\infty}-1)^2 + 4(\delta e^{2\delta\|H\|_\infty}\sup\|\nabla H\|)^2)\left(e^{-2\rho_{\mathrm{PI}}T}+\varepsilon\right).\end{aligned}$$

Plugging this back into the earlier sequence of equations completes the proof. □

Remark 7.4.14. While the proof above has been stated for the annealing scheme where at time t the Hamiltonian is of the form $\sigma \mapsto \beta(t)\cdot H(\sigma)$, the proof immediately extends to essentially any annealing scheme that changes the Hamiltonian "slowly", in that if H_t is the Hamiltonian at time t, $\|H_{t+T}-H_t\|_\infty \leq \delta$ for all t. A concrete example of such a scheme that might work better than the vanilla annealing is that which at time t has as Hamiltonian $\sigma \mapsto H(\beta(t)\cdot\sigma)$.

7.5 Vignette: sampling from mixture models with advice

We are interested in the following question.

> Let π be a distribution over $\mathbb{R}^N$ with density proportional to e^{-V}. Given oracle access to the gradient ∇V, when is it possible to efficiently produce samples that are close (in total variation distance) to π?

We begin with an overview of existing results towards the above question. Recall from Fact 7.3.7 that for distributions satisfying a Poincaré inequality, such as strongly log-concave distributions, Langevin diffusion enjoys rapid mixing. Beyond this setting, however, very little is known. [BCE+22, CWZZ24] prove certain "local mixing" guarantees for Langevin diffusion on non-log-concave distributions, but these do immediately not translate to sampling guarantees. The works [GLR18, LRG18, GTC24] use Langevin diffusion-based algorithms to sample from mixtures of log-concave distributions. Furthermore, the first of these papers proves that it is hard to sample from a mixture of two Gaussian distributions with distinct covariance matrices given access to just the gradient ∇V.

In [KV24], the first theoretical guarantees are provided for a new model designed to circumvent this issue, where in addition to being given access to the gradient ∇V, we are also given "advice" in the form of m samples from the distribution (also see [NHH+20] and [Hin10, GLZ+18, XLZW16] for related discussion). In particular, they show that when the stationary distribution is a mixture of constantly many strongly log-concave distributions, Langevin diffusion initialized at the empirical measure on the advice gets close to the stationary distribution. However, their dependence on the number of components K is doubly exponential. The main result in this section improves the doubly exponential dependence to a polynomial one for any mixture of distributions satisfying Poincaré inequalities. Similar results are obtained by Koehler, Lee, & Vuong [KLV24].

Theorem 7.5.1. *Let $\varepsilon, \delta \in (0,1)$, and let π a mixture*

$$\pi = \sum_{i=1}^{K} p_i \pi_i$$

of distributions $(\pi_i)_{i=1}^K$, where each π_i satisfies a Poincaré inequality with constant (at least) ρ_{PI}. Further assume that $p_i \geq p_*$ for all i. Let ν_0 be a random distribution over $\mathbb{R}^N$ such that $\mathbb{E}\,\nu_0 = \pi$, in that for any measurable subset A of $\mathbb{R}^N$, $\mathbb{E}\,\nu_0(A) = \pi(A)$. Set

$$m = \Omega\left(\frac{\log(1/\delta)}{p_* \varepsilon^2}\right).$$

Let $\nu_1, \ldots, \nu_m$ be iid draws from ν_0, and ν the uniform mixture over the $(\nu_i)_{i=1}^m$. Further suppose that with probability at least $1-\delta$, $\chi^2(\nu_i \| \pi) \leq M$. Denoting by μ_T the distribution attained by running Langevin diffusion for time T initialized ν, it holds that

$$\Pr\left[\chi^2(\mu_T \| \pi) \leq \varepsilon\right] \geq 1 - O(\delta),$$

for $T = \Omega\left(\frac{1}{\rho_{\mathrm{PI}}} \log\left(\frac{M}{\varepsilon}\right)\right)$, where the probability is over the draws of ν_i.

Remark 7.5.2. One should think of ν_0 as being the point mass distribution supported on a (random) sample drawn from π. Alternatively, one can think of ν_0 as being the distribution obtained by drawing a sample x_0 according to π, then running Langevin diffusion for a short amount of time — doing this would make the χ^2-divergence $\chi^2(\nu_0 \| \pi)$ finite. We also remark that a version of this proof goes through if we have that each π_i satisfies a log-Sobolev inequality instead of a Poincaré inequality, working with KL divergences instead.

Proof of Theorem 7.5.1. The idea of the proof will be to show that up to some additive error depending on the samples, π does satisfy a Poincaré inequality with respect to the distributions along the path of Langevin diffusion initialized at the empirical distribution. This error corresponds to how imbalanced the samples are in terms of the mixture weights — a straightforward concentration argument using Bernstein's inequality then shows that this error is small, so the χ^2 divergence essentially decays exponentially fast, as if π satisfied a true Poincaré inequality.

Let f_t be the Radon-Nikodym derivative of μ_t (obtained by running Langevin diffusion initialized at ν) with respect to π. By definition, we have

$$\begin{aligned}
\chi^2(\mu_t \| \pi) &= \mathbb{E}_{\pi}[f_t^2] - 1 \\
&= \sum_{i=1}^{K} p_i \left(\mathbb{E}_{\pi_i}[f_t^2] - 1\right) \\
&= \sum_{i=1}^{K} p_i \mathsf{Var}_{\pi_i}[f_t] + \sum_{i=1}^{K} p_i \left(\mathbb{E}_{\pi_i}[f_t]^2 - 1\right).
\end{aligned}$$

Because each π_i satisfies a Poincaré inequality, the first term is bounded as

$$\sum_{i=1}^{K} p_i \mathsf{Var}_{\pi_i}[f_t] \leq \frac{1}{\rho_{\mathrm{PI}}} \sum_{i=1} p_i \mathbb{E}_{\pi_i} \|\nabla f_t\|^2 = \frac{1}{\rho_{\mathrm{PI}}} \mathbb{E}_{\pi} \|\nabla f_t\|^2.$$

Consequently,

$$\chi^2(\mu_t \| \pi) \leq \frac{1}{\rho_{\mathrm{PI}}} \cdot \mathbb{E}_{\pi} \|\nabla f_t\|^2 + \sum_{i=1}^{K} p_i \left(\mathbb{E}_{\pi_i}[f_t]^2 - 1\right). \tag{7.4}$$

Theorem 7.4.6 then yields that

$$\chi^2(\mu_T \| \pi) \leq \chi^2(\mu_0 \| \pi) \cdot e^{-\rho_{\mathrm{PI}} T} + \mathop{\mathbb{E}}_{s \sim \Lambda_T}\left[\sum_{i=1}^{K} p_i \left(\mathop{\mathbb{E}}_{\pi_i}[f_t]^2 - 1\right)\right]$$
$$\leq M e^{-\rho_{\mathrm{LS}} T} + \mathop{\mathbb{E}}_{s \sim \Lambda_T}\left[\sum_{i=1}^{K} p_i \left(\mathop{\mathbb{E}}_{\pi_i}[f_t]^2 - 1\right)\right].$$

Above, we use that because the KL divergence to π of each of the ν_i is at most M, so is that of the mixture $\mu_0 = \nu$.

To conclude, we shall establish tail bounds on

$$\mathop{\mathbb{E}}_{s \sim \Lambda_T}\left[\sum_{i=1}^{K} p_i \left(\mathop{\mathbb{E}}_{\pi_i}[f_s]^2 - 1\right)\right].$$

For $1 \leq j \leq m$, let $f_s^{(j)}$ be the Radon-Nikodym derivative of $\mu_s^{(j)}$ with respect to π, where $\mu_s^{(j)}$ is the distribution obtained by running Langevin diffusion for time s initialized at ν_j. It is not difficult to see that $f_s = \frac{1}{m}\sum_{j=1}^{m} f_s^{(j)}$.

First, for fixed s and j, we use the fact that the $(\mathbb{E}_{\pi_i}[f_s^{(j)}])_j$ are independent mean 1 random variables, with Hoeffding's inequality, to get tail bounds for $\mathbb{E}_{\pi_i}[f_s]^2 - 1$. We may use this to bound a certain Orlicz norm of this random variable — this bound on the norm also transfers to $\mathbb{E}_{s \sim \Lambda_T}\left[\sum_{i=1}^{K} p_i \left(\mathbb{E}_{\pi_i}[f_s]^2 - 1\right)\right]$ as it is a convex combination of random variables with bounded Orlicz norm. This immediately yields the desired tail bound.

Fix s and i. To start, we have the almost sure bounds

$$\frac{1}{p_*} = \frac{1}{p_*}\mathop{\mathbb{E}}_{\pi}[f_s^{(j)}] = \frac{1}{p_*}\sum_{r=1}^{K} p_r \mathop{\mathbb{E}}_{\pi_r}[f_s^{(j)}] \geq \mathop{\mathbb{E}}_{\pi_i}[f_s^{(j)}] \geq 0.$$

Note that because the expected ν_j is equal to π, $\mathbb{E}_{\nu_j}\mathbb{E}_{\pi_i}[f_s^{(j)}] = 1$ for any j. Furthermore, because $\mathbb{E}_{\pi_i}\left[f_s^{(j)}\right]$ is a mean 1 random variable which is bounded in $\left[0, \frac{1}{p_*}\right]$, its variance is at most $\frac{1}{p_*}$ (see e.g. [BD00]). Bernstein's inequality implies that

$$\Pr\left[\left|\mathop{\mathbb{E}}_{\pi_i} f_s - 1\right| > t\right] = \Pr\left[\left|\frac{1}{m}\sum_{i=1}^{m}\mathop{\mathbb{E}}_{\pi_i}\left[f_s^{(j)}\right] - 1\right| > t\right] \leq 2\exp\left(-\frac{mp_*}{2} \cdot \frac{t^2}{1+t}\right).$$

Thus, for any $t > 0$,

$$\Pr\left[\left|\mathop{\mathbb{E}}_{\pi_i}[f_s]^2 - 1\right| > t\right] \leq \Pr\left[\left|\mathop{\mathbb{E}}_{\pi_i} f_s - 1\right| > \frac{t}{2(1+\sqrt{t})}\right]$$
$$\leq 2\exp\left(-\frac{mp_*}{8} \cdot \frac{\left(\frac{t}{1+\sqrt{t}}\right)^2}{1 + \frac{t}{1+\sqrt{t}}}\right).$$

Now, consider the Orlicz norm $\|\cdot\|_\psi$ associated to the above family of tail bounds. As mentioned earlier, standard machinery may be used to go from the above tail bounds to a bound on the norm $\left\|\mathbb{E}_{\pi_i}[f_s]^2 - 1\right\|_\psi$. Convexity of the norm yields the same bound on $\left\|\mathbb{E}_{s \sim \Lambda_T}\sum_{i=1}^{K} p_i \left(\mathbb{E}_{\pi_i}[f_s]^2 - 1\right)\right\|_\psi$. Translating this back to a tail bound, we get that

$$\Pr\left[\left|\mathop{\mathbb{E}}_{s \sim \Lambda_T}\sum_{i=1}^{K} p_i \left(\mathop{\mathbb{E}}_{\pi_i}[f_s]^2 - 1\right)\right| > \frac{\varepsilon}{2}\right] \leq 2\exp\left(-\frac{mp_*\varepsilon^2}{10}\right) \leq \delta.$$

Conditioning on the above event not happening, we get that

$$\chi^2(\mu_T\|\pi) \le \chi^2(\mu_0\|\pi) \cdot e^{-\rho_{\mathrm{PI}}\cdot T} + \frac{\varepsilon}{2} \le \varepsilon,$$

as desired. □

7.6 Stopped localization schemes

7.6.1 Localization schemes

We review some basic notions for the localization schemes framework introduced in [CE22].

Definition 7.6.1 (Linear-tilt localization scheme)**.** Let $\mu = \mu_0$ be a probability measure, $(\mu_t)_{t\in\mathbb{Z}_{\ge 0}}$ be a localization process. A linear-tilt localization scheme is one where μ_t is defined by

$$\mu_{t+1}(x) = \mu_t(x)\,(1 + \langle x - \boldsymbol{m}(\mu_t), Z_t\rangle)$$

where Z_t is a random variable with $\mathbb{E}[Z_t|\mu_t] = 0$ and $\boldsymbol{m}(\mu_t)$ denotes the mean of μ_t.

For our main application to p-spin models, we will focus on a continuous-time version of linear-tilt localization known as stochastic localization [Eld13].

Definition 7.6.2 (Stochastic localization)**.** Let μ be a probability measure on $\Omega \subseteq \mathbb{R}^N$, $(B_t)_{t\ge 0}$ be a standard Brownian motion on $\mathbb{R}^N$. The stochastic localization process with driving matrix $(C_t)_{t\ge 0}$ is a localization process $(\mu_t)_{t\ge 0}$ with $\mu_0 = \mu$ and

$$\mu_t(x) \propto \mu_0(x)\exp(-\tfrac{1}{2}\langle x, \Sigma_t x\rangle + \langle y_t, x\rangle),$$

where $\Sigma_t = \int_0^t C_s^2 \mathsf{d}s$ and $y_t = \int_0^t C_s^2 \boldsymbol{m}(\mu_s)\mathsf{d}s + C_s \mathsf{d}B_s$.

A crucial property of these localization schemes is that establishing (approximate) conservation of variance reduces to bounding the covariance matrices of the intermediate distributions μ_t.

Lemma 7.6.3 (Conservation of variance for linear-tilt [CE22, Claim 22])**.** *Let $(\mu_t)_{t\in\mathbb{Z}_{\ge 0}}$ be a linear-tilt localization process. Suppose that for all $t \le T$ we have*

$$\left\|\mathsf{Cov}(Z_t|\mu_t)^{1/2} \cdot \mathsf{Cov}(\mu_t) \cdot \mathsf{Cov}(Z_t|\mu_t)^{1/2}\right\|_{\mathsf{op}} \le K_t,$$

where $K_t \in [0,1]$. Then for any function φ,

$$\frac{\mathbb{E}\,\mathsf{Var}_{\mu_T}[\varphi]}{\mathsf{Var}_{\mu}[\varphi]} \ge \prod_{t=0}^{T-1}(1 - K_t).$$

Lemma 7.6.4 (Conservation of variance for stochastic localization)**.** *Let $(\mu_t)_{t\ge 0}$ be a stochastic localization process with driving matrix $(C_t)_{t\ge 0}$. Suppose that for all $t \le T$ we have*

$$\left\|C_t^{1/2} \cdot \mathsf{Cov}(\mu_t) \cdot C_t^{1/2}\right\|_{\mathsf{op}} \le K_t$$

where $K_t \in [0,1]$. Then for any function φ,

$$\frac{\mathbb{E}\,\mathsf{Var}_{\mu_T}[\varphi]}{\mathsf{Var}_{\mu}[\varphi]} \ge e^{-\int_0^T K_t \mathsf{d}t}.$$

7.6.2 Proving weak Poincaré inequalities using stopped localization schemes

To apply Theorem 7.4.12, we required weak Poincaré inequalities for the measures of interest. To show these, we next introduce a generic tool to prove these using Lemma 7.4.10, building on the localization schemes framework introduced in Section 7.6.1. Let μ be a distribution. Using a localization scheme, we would like to design a measure decomposition $\mu = \mathbb{E}_{\boldsymbol{z}\sim\rho}\, \mu_{\boldsymbol{z}}$ such that

- for all functions f, $\mathsf{Var}_{\pi}[f] \leq C_{\mathrm{Var}}\, \mathbb{E}_{\boldsymbol{z}\sim\rho} \mathsf{Var}_{\pi_{\boldsymbol{z}}}[f]$, and
- with probability $1-\eta$ over $\boldsymbol{z} \sim \rho$, $\pi_{\boldsymbol{z}}$ satisfies a $(\rho_{\mathrm{PI}}, \delta)$-weak Poincaré inequality.

One way to ensure the first condition — approximate conservation of variance — is to simply stop the localization scheme whenever it fails to hold. Indeed, the following lemma immediately follows from Lemma 7.6.3.

Lemma 7.6.5. *Let $\mu = \mu_0$ be a measure, and let $(\mu_t)_{t\in\mathbb{Z}_{\geq 0}}$ be a linear-tilt localization process defined by*

$$\mu_{t+1}(x) = \mu_t(x)\,(1 + \langle x - \boldsymbol{m}(\mu_t), Z_t\rangle)$$

for some random variable Z_t with $\mathbb{E}[Z_t|\mu_t] = 0$. Let $T > 0$ be an arbitrary stopping time and $0 \leq K_t < 1$ for each $t \geq 0$, and consider the stopping time

$$\tau = T \wedge \inf_{t\geq 0} \left\{ \left\| \mathsf{Cov}(Z_t|\mu_t)^{1/2} \cdot \mathsf{Cov}(\mu_t) \cdot \mathsf{Cov}(Z_t|\mu_t)^{1/2} \right\|_{\mathsf{op}} \geq K_t \right\}.$$

Then, for any function φ,

$$\frac{\mathbb{E}\,\mathsf{Var}_{\mu_\tau}[\varphi]}{\mathsf{Var}_{\mu}[\varphi]} \geq \prod_{t\geq 0}(1 - K_t).$$

Similarly, we have the following lemma for stochastic localization, which follows from Lemma 7.6.4.

Lemma 7.6.6. *Let $\mu = \mu_0$ be a measure, and $(\mu_t)_{t\geq 0}$ be a stochastic localization process with driving matrix $(C_t)_{t\geq 0}$. Let $T, K > 0$ be constant parameters, and consider the stopping time*

$$\tau = T \wedge \inf_{t\geq 0} \left\{ \left\| C_t^{1/2} \cdot \mathsf{Cov}(\mu_t) \cdot C_t^{1/2} \right\|_{\mathsf{op}} \geq K \right\}.$$

Then,

$$\frac{\mathbb{E}\,\mathsf{Var}_{\mu_\tau}[\varphi]}{\mathsf{Var}_{\mu}[\varphi]} \geq e^{-TK}.$$

Remark 7.6.7. The localization process in the above lemmas can depend on φ, and need not be a linear-tilt localization. The more general requirement is that

$$\frac{\mathbb{E}\left[\mathsf{Var}_{\mu_{t+1}}[\varphi]|\mu_t\right]}{\mathsf{Var}_{\mu_t}[\varphi]} \geq K_t \quad \text{or} \quad \frac{1}{\mathsf{Var}_{\mu_t}[\varphi]} \cdot \frac{\mathsf{d}}{\mathsf{d}s}\, \mathbb{E}\left[\mathsf{Var}_{\mu_s}[\varphi] \mid \mu_t\right]\Big|_{s=t} \geq K.$$

This can always be achieved by stopping the localization process whenever these conditions fail to hold.

With these elements in hand, we now show how to prove a weak Poincaré inequality using stopped localization schemes.

Lemma 7.6.8. *Let $\mu = \mu_0$ be a measure, and $(\mu_t)_{t\geq 0}$ be a stochastic localization process with driving matrix $(C_t)_{t\geq 0}$. Let $T, K > 0$ be constant parameters. Suppose that with probability $1 - \eta_1$, it holds that $\left\| C_t^{1/2} \cdot \mathsf{Cov}(\mu_t) \cdot C_t^{1/2} \right\|_{\mathsf{op}} < K$ for all $t \in [0, T]$. Further suppose that with probability $1 - \eta_2$, μ_T satisfies a $(\rho_{\mathrm{PI}}, \delta)$-weak Poincaré inequality. Then, μ satisfies a $\left(\rho_{\mathrm{PI}} e^{-TK}, e^{TK}\left(\delta + \eta_1 + \eta_2\right)\right)$-weak Poincaré inequality.*

Proof. As in Lemma 7.6.6, define the stopping time

$$\tau = T \wedge \inf_{t\geq 0}\left\{\left\|C_t^{1/2}\cdot \mathsf{Cov}(\mu_t)\cdot C_t^{1/2}\right\|_{\mathsf{op}} \geq K\right\}.$$

Consider the measure decomposition $\mu = \mathbb{E}\,\mu_\tau$. By Lemma 7.6.6, this decomposition is variance-conserving with parameter e^{-K}. By the hypothesis of the lemma, $\tau = T$ with probability $1-\eta_1$, and μ_1 satisfies a weak Poincaré inequality with probability $1-\eta_2$. Consequently, μ_τ satisfies a weak Poincaré inequality with probability at least $1-\eta_1-\eta_2$. Lemma 7.4.10 completes the proof. □

Remark 7.6.9. An analogous lemma to the above holds if $(\mu_t)_{t\in\mathbb{Z}_{\geq 0}}$ is any linear-tilt localization process.

While it will not be used in this paper, we note that a similar method proves a weak Poincaré inequality for a natural Markov chain associated to a localization scheme. This includes for example the restricted Gaussian dynamics; see [CE22] for several other examples.

Lemma 7.6.10. *Let $\mu = \mu_0$ be a measure, and $(\mu_t)_{t\geq 0}$ be a stochastic localization process with driving matrix $(C_t)_{t\geq 0}$. Let $T, K > 0$ be constant parameters. Consider the Markov chain P given by $P_{x\to y} = \mathbb{E}\left[\frac{\mu_T(x)\mu_T(y)}{\mu_0(x)}\right]$. Define the stopping time*

$$\tau = T \wedge \inf_{t\geq 0}\left\{\left\|C_t^{1/2}\cdot \mathsf{Cov}(\mu_t)\cdot C_t^{1/2}\right\|_{\mathsf{op}} \geq K\right\}.$$

If $\tau = T$ with probability at least $1-\delta$, then P satisfies a $(e^{-TK}, \delta e^{TK})$-weak Poincaré inequality.

Proof. For the Markov chain P, the Dirichlet form is given by $\mathcal{E}_P(f,f) = \mathbb{E}\,\mathsf{Var}_{\mu_T}[f]$ (see, e.g., [CE22, Proposition 19]). We then have the chain of inequalities

$$\begin{aligned}\mathbb{E}\,\mathsf{Var}_{\mu_T}[f] &\geq \mathbb{E}\,\mathsf{Var}_{\mu_\tau}[f] - \delta \operatorname{osc}(f)^2\\ &\geq e^{-TK}\mathsf{Var}_{\mu_0}[f] - \delta\operatorname{osc}(f)^2.\end{aligned}$$

The first inequality here is immediate since

$$\mathbb{E}\,\mathsf{Var}_{\mu_\tau}[f] - \mathbb{E}\,\mathsf{Var}_{\mu_T}[f] = \mathbb{E}\,\mathsf{Var}_{\mu_\tau}[f]\mathbf{1}_{\tau\neq T} \leq \operatorname{osc}(f)^2\Pr[\tau\neq T].$$

The second inequality follows from Lemma 7.6.4. □

Remark 7.6.11. As in Remark 7.6.7, the above lemma can be generalized to localization schemes other than stochastic localization.

7.7 Sampling from spherical p-spin models

In this section, we prove that simulated annealing samples from spherical spin glass models for models satisfying (SL). Recall that $S_N = \sqrt{N}\cdot\mathbb{S}^{N-1}$. For $\gamma_2,\gamma_3,\ldots,\gamma_{p_*}\geq 0$, the mixed p-spin Hamiltonian $H_N : S_N \to \mathbb{R}$ is defined by

$$H_N(\sigma) := \sum_{p\geq 2}\frac{\gamma_p}{N^{(p-1)/2}}\sum_{i_1,\ldots,i_p=1}^{N}\boldsymbol{g}_{i_1,\ldots,i_p}\sigma_{i_1}\cdots\sigma_{i_p}, \tag{7.5}$$

for i.i.d. samples $\boldsymbol{g}_{i_1,\ldots,i_p}$ from $\mathcal{N}(0,1)$. This is the gaussian process on $\mathbb{R}^N$ with covariance

$$\mathbb{E}\,H_N(\sigma^1)\cdot H_N(\sigma^2) = N\cdot\xi\left(R(\sigma^1,\sigma^2)\right),$$

where we recall the mixture function ξ is defined by $\xi(s) = \sum_{p=2}^{p_*}\gamma_p^2 s^p$. The algorithm we will study is the following simple annealing scheme for Langevin diffusion.

Definition 7.7.1 (Annealed Langevin diffusion)**.** Let $\delta_N, T_N > 0$ be parameters possibly depending on N. For any $\beta \geq 0$, let $\mu_\beta \coloneqq \mu_{\beta H_N}$, where H_N is the p-spin Hamiltonian. Annealed Langevin diffusion is the annealing scheme $\mathcal{A}$ where $\beta(t) = \delta_N \lfloor t/T_N \rfloor$ and P_β is the Langevin semigroup operator for Gibbs distribution μ_β. In words, $\mathcal{A}$ keeps β constant for time T_N and then increments β by δ_N.

Theorem 7.7.2. *Let H_N be a mixed p-spin Hamiltonian whose mixture function ξ satisfies* (SL)*, which we recall below:*

$$\xi''(q) < \frac{1}{(1-q)^2} \text{ for all } q \in [0,1).$$

Let μ be the associated Gibbs measure over the scaled sphere S_N, with

$$\mathsf{d}\mu(\sigma) \propto \exp(H_N(\sigma))\mathsf{d}\rho(\sigma).$$

With probability $1 - e^{-cN^{1/5}}$ over the randomness of H_N, the following holds. For some parameters $\delta_N = O(N^{-4/5})$, $T_N = \Omega(N^{1/5})$, the output measure ν of the the annealed Langevin diffusion scheme with these parameters satisfies

$$\mathrm{TV}(\nu, \mu) \leq e^{-cN^{1/5}}.$$

Remark 7.7.3. We expect that the error $e^{-cN^{1/5}}$ can be improved to e^{-cN}, matching the fact that a $e^{-O(N)}$ fraction of the Gibbs measure is typically trapped in metastable states between the uniqueness and shattering thresholds [BJ24]. However, we will not pursue this improvement in this paper.

Remark 7.7.4. The condition (SL) is a fundamental barrier for stochastic localization, both as an algorithm and a proof technique. As was essentially shown in [HMP24, Section 10], for models satisfying (Strict RS) but not (SL), the means $\boldsymbol{m}(\mu_t)$ along the localization process do not move stably, in the sense that there exist time intervals of width $o(1)$ in which $\boldsymbol{m}(\mu_t)$ moves by $\Omega(N^{1/2})$. (The condition (Strict RS) is an artifact of the proof, and it is expected that the mean continues to move non-stably beyond the regime (Strict RS)). In the setting of [HMP24], this implies that their algorithmic simulation of the localization process fails, because approximate message passing will not estimate the mean at some times. In our setting, this implies that the covariance $\mathsf{Cov}(\mu_t)$, which arises as the derivative of $\boldsymbol{m}(\mu_t)$, is genuinely not bounded in operator norm at some times, and thus the main input to our framework does not hold.

Remark 7.7.5. While the result above is stated for the continuous time Langevin *diffusion*, the results therein can be adapted to the discretized Langevin Monte Carlo algorithm using standard tools, à la [Che23b, Part II], to obtain a polynomial time sampling algorithm.

To prove the above, we shall use Theorem 7.4.12 in conjunction with Lemma 7.6.8. For the remainder of this section, let $(\gamma_p)_{p\geq 2}$ be a sequence of weights such that the associated mixture function ξ satisfies the condition (SL).

Notation 7.7.6 (Measure decomposition for p-spin models)**.** Let $\mu_{H_N} = \mu_0$. For a large constant time T, let $(\mu_t)_{0\leq t\leq T}$ be the stochastic localization process with driving matrix Id (see Definition 7.6.2).

Lemma 7.7.7 (Covariance bound on stochastic localization path)**.** *There exist constants c, K, depending only on ξ, such that for any constant $T > 0$ the following holds with probability at least $1 - e^{-cN^{1/5}}$ over the randomness of H_N. If $(\mu_t)_{0\leq t\leq T}$ is the (random) trajectory of stochastic localization initialized at $\mu_0 = \mu_{H_N}$, with probability $1 - e^{-cN^{1/5}}$, $\|\mathsf{Cov}(\mu_t)\|_{\mathsf{op}} < K$ for all $0 \leq t \leq T$. In other words,*

$$\Pr_{H_N}\left[\Pr_{(\mu_t)|H_N}\left[\|\mathsf{Cov}(\mu_t)\|_{\mathsf{op}} < K \textit{ for all } 0 \leq t \leq T\right] \geq 1 - e^{-cN^{1/5}}\right] \geq 1 - e^{-cN^{1/5}}.$$

Lemma 7.7.8 (Weak Poincaré inequality for endpoint distributions)**.** *There exists a constant T depending only on ξ such that the following holds with probability at least $1 - e^{-cN}$ over the randomness of H_N. With probability at least $1 - e^{-cN}$, the (random) measure μ_T satisfies a (c, e^{-cN})-weak Poincaré inequality. In other words,*

$$\Pr_{H_N}\left[\Pr_{\mu_T|H_N}\left[\mu_T \textit{ satisfies a } (c, e^{-cN})\textit{-weak Poincaré inequality}\right] \geq 1 - e^{-cN}\right] \geq 1 - e^{-cN}.$$

Let us first see how these two lemmas imply the main theorem.

Proof of Theorem 7.7.2. Fix some $0 \le \beta \le 1$. Note that the Hamiltonian βH_N has mixture function $\xi_\beta(s) = \xi(\beta^2 s)$, and if ξ satisfies (SL) then ξ_β does as well. Plugging in Lemmas 7.7.7 and 7.7.8 into Lemma 7.6.8 implies that with probability at least $1 - e^{-cN^{1/5}}$, $\mu_{\beta H_N}$ satisfies a $(c, e^{-cN^{1/5}})$-weak Poincaré inequality. A union bound implies that with probability $1-e^{-cN^{1/5}}$, for all β encountered along the annealing schedule, $\mu_{\beta H_N}$ satisfies a $(c, e^{-cN^{1/5}})$-weak Poincaré inequality.

By [HS25, Proposition 2.3], with probability $1 - e^{-cN}$, $\|\nabla H_N\|_\infty = O(\sqrt{N})$. The same argument implies that with probability $1 - e^{-cN}$, $\|H_N\|_\infty = O(N)$. With probability $1 - e^{-cN^{1/5}}$, all three of these events occur, and Theorem 7.4.12 completes the proof. $\square$

We conclude this subsection by proving Lemma 7.7.8.

Proof of Lemma 7.7.8. Let $\mathrm{d}\mu_t(\sigma) \propto e^{H_{N,T}(\sigma)}\,\mathrm{d}\sigma$ for $H_{N,T}(\sigma) = H_N(\sigma) + \langle \boldsymbol{y}, \sigma\rangle$. Let

$$S_N(\boldsymbol{y}) = \{\sigma \in S_N : R(\boldsymbol{y}, \sigma) > 0\}.$$

Let $\boldsymbol{U} \in \mathbb{R}^{N\times(N-1)}$ be a matrix whose columns are an orthonormal basis of the orthogonal complement of $\boldsymbol{y}$. Let $\widehat{\boldsymbol{y}} = \sqrt{N}\boldsymbol{y}/\|\boldsymbol{y}\|$ be $\boldsymbol{y}$ (which is a.s. nonzero) scaled to length $\sqrt{N}$, and define the map $\sigma_{\boldsymbol{y}}(\rho) : \mathbb{R}^{N-1} \to S_N(\boldsymbol{y})$ by

$$\sigma_{\boldsymbol{y}}(\rho) = \frac{\widehat{\boldsymbol{y}} + \boldsymbol{U}\rho}{\sqrt{1 + R(\rho,\rho)}}.$$

This is the inverse of the map that first stereographically projects $S_N(\boldsymbol{y})$ from the origin to $\widehat{\boldsymbol{y}} + \boldsymbol{U}\mathbb{R}^{N-1}$, the plane tangent to S_N at $\widehat{\boldsymbol{y}}$, and then maps the resulting point to coordinates given by $\boldsymbol{U}$. Let $\varepsilon_0 = 0.1$, and $A = \{\rho \in \mathbb{R}^{N-1} : \|\rho\|^2 \le \varepsilon_0 N\}$, and note that

$$\sigma_{\boldsymbol{y}}(A) = \{\sigma \in S_N : R(\sigma, \widehat{\boldsymbol{y}}) \ge (1+\varepsilon_0)^{-1/2}\}$$

is a spherical cap around $\widehat{\boldsymbol{y}}$. Let $A' \coloneqq \sigma_{\boldsymbol{y}}(A)$. By arguments in [HMP24, Subsection 9.2], there exists a measure ν (denoted $\tilde{\nu}^{\mathsf{proj}}_{H_N,\boldsymbol{y}}$, see Eq. 2.10 therein) such that the following holds with probability $1 - e^{-cN}$.

- The push-forward of $\nu_{|A}$ through $\sigma_{\boldsymbol{y}}$ coincides with $(\mu_T)_{|A'}$. (Lemma 9.5 therein.)
- $\mu_T(A') = 1 - e^{-cN}$. (Lemma 9.6 therein states this with $1 - o_N(1)$ in place of $1 - e^{-cN}$, but the proof implies bound $1 - e^{-cN}$, as this is the bound given by Proposition 5.12 used therein.)
- $\nu(A) = 1 - e^{-cN}$. (Corollary 9.7 therein, modulo the same issue of $1 - o_N(1)$ versus $1 - e^{-cN}$, which is addressed the same way.)
- ν is $\Omega(1)$-strongly log-concave. (Proposition 9.8 therein.)

By the well-known Bakry-Émery condition (see, e.g., [Che23b, Section 1.2.3]), on this event ν satisfies a ρ_{PI}-Poincaré inequality for some $\rho_{\mathrm{PI}} = \Omega(1)$. We will transfer this inequality to a $(\rho_{\mathrm{PI}}, e^{-cN})$-weak Poincaré inequality for μ_T. Consider a smooth test function $f : S_N \to \mathbb{R}$ and let $\widetilde{f} : \mathbb{R}^{N-1} \to \mathbb{R}$ be defined by $\widetilde{f} = f \circ \sigma_{\boldsymbol{y}}$. Since $\mathrm{TV}\big(\mu_T, (\mu_T)_{|A'}\big) = e^{-cN}$ and $\mathrm{TV}\big(\nu, \nu_{|A}\big) = e^{-cN}$, and $\mathrm{osc}(f') \le \mathrm{osc}(f)$, arguing as in (7.3) shows

$$\begin{aligned}
\mathsf{Var}_{\mu_T}(f) &\le \mathsf{Var}_{(\mu_T)_{|A'}}(f) + e^{-cN}\mathrm{osc}(f) \\
&= \mathsf{Var}_{\nu_{|A}}(\widetilde{f}) + e^{-cN}\mathrm{osc}(f) \\
&\le \mathsf{Var}_{\nu}(\widetilde{f}) + 2e^{-cN}\mathrm{osc}(f).
\end{aligned}$$

By the Poincaré inequality for ν and the definition of the Dirichlet form for Langevin diffusion,

$$\mathsf{Var}_\nu(\widetilde{f}) \le \frac{1}{\rho_{\mathrm{PI}}} \cdot \mathcal{E}_\nu(\widetilde{f}, \widetilde{f}) = \frac{1}{\rho_{\mathrm{PI}}} \cdot \mathop{\mathbb{E}}_{\nu}[\|\nabla \widetilde{f}\|^2]$$

By [HMP24, Proof of Lemma 9.5], the map $\sigma_{\boldsymbol{y}}$ has Jacobian $J_{\sigma_{\boldsymbol{y}}}$ satisfying $\|J_{\sigma_{\boldsymbol{y}}}\|_{\mathsf{op}} \le 1$, and thus for all $\rho \in \mathbb{R}^{N-1}$,

$$\|\nabla \widetilde{f}(\rho)\| = \|\nabla (f \circ \sigma_{\boldsymbol{y}})(\rho)\| \le \|\nabla f(\sigma_{\boldsymbol{y}}(\rho))\|.$$

It follows that

$$\begin{aligned}\mathop{\mathbb{E}}_{\nu}[\|\nabla \widetilde{f}\|^2] &\le \mathop{\mathbb{E}}_{\nu_{|A}}[\|\nabla \widetilde{f}\|^2] + e^{-cN} \sup \|\nabla \widetilde{f}\|^2 \\ &\le \mathop{\mathbb{E}}_{(\mu_T)_{|A'}}[\|\nabla f\|^2] + e^{-cN} \sup \|\nabla f\|^2 \\ &\le \mathop{\mathbb{E}}_{\mu_T}[\|\nabla f\|^2] + 2e^{-cN} \sup \|\nabla f\|^2\end{aligned}$$

Combining the above shows

$$\mathsf{Var}_{\mu_T}(f) \le \frac{1}{\rho_{\mathrm{PI}}} \mathcal{E}_{\mu_T}(f,f) + 2e^{-cN} \left(\operatorname{osc}(f) + \frac{1}{\rho_{\mathrm{PI}}} \sup \|\nabla f\|^2 \right).$$

The result follows by adjusting c. □

7.7.1 Technical overview for covariance bounds

The proof of the main theorem has boiled down to Lemma 7.7.7 — we now give a high-level overview of our proof strategy for this. We wish to show that with very high probability $(1 - e^{-\Omega(N^{1/5})})$, the covariance is bounded along the entire path $(\mu_t)_{0 \le t \le T}$ of stochastic localization. By performing a union bound over time and a standard perturbation argument, it suffices to show that for a fixed time $t \in [0,T]$, μ_t has bounded covariance with very high probability.

To do this, we recall an alternate view of stochastic localization [AM22]. The measure at time t of stochastic localization (with the identity driving matrix) is given as follows. First, draw $\boldsymbol{\sigma} \sim \mu_{H_N}$, and independently $\boldsymbol{g} \sim \mathcal{N}(0, I_N)$. Then, μ_t has the same law as $\mu_{H_N, t\boldsymbol{\sigma} + \sqrt{t}\boldsymbol{g}}$, in that

$$\mu_{H_N, t\boldsymbol{\sigma}+\sqrt{t}\boldsymbol{g}}(\widetilde{\sigma}) \propto \exp\left(H_N(\widetilde{\sigma}) + \langle t\boldsymbol{\sigma} + \sqrt{t}\boldsymbol{g}, \widetilde{\sigma}\rangle\right).$$

As written, the covariance of this distribution is difficult to analyze — the sample $\boldsymbol{\sigma}$ has very complicated correlations with the disorder of the Hamiltonian H_N, making it intractable.

The planting trick. To deal with this, we will use the *planting trick* introduced by Achlioptas and Coja-Oghlan [AC08]. The application of this method in the context of stochastic localization is by now standard [AMS22, AMS25, HMP24], and we review the main ideas for the reader's convenience.

Definition 7.7.9 (Planted p-spin model)**.** The planted measure μ_{pl} is a joint law over a Hamiltonian H_N and a *spike* $\boldsymbol{x} \in S_N$ given by

$$\mathsf{d}\mu_{\mathsf{pl}}(H_N, x) \propto \exp\left(H_N(x)\right) \cdot \mathsf{d}\rho(x) \cdot \mathsf{d}\mu_{\mathsf{null}}(H_N),$$

where ρ is the uniform measure over S_N and μ_{null} is the law over p-spin Hamiltonians with mixture function ξ. We frequently abuse notation to let $\mu_{\mathsf{pl}}(H_N)$ denote the marginal of μ_{pl} on H_N.

To provide further intuition for the above definition, consider the following alternate sampling interpretation of the planted model, which describes the distribution of $\boldsymbol{x}$ conditioned on H_N.

Fact 7.7.10. *Consider the following inference problem. We start by sampling the spike $\boldsymbol{x} \sim S_N$, sample $G^{(p)}$ as a rank-p tensor with iid $\mathcal{N}(0,1)$ entries for $p \ge 2$, and for each p let $M^{(p)} = -G^{(p)} + \frac{\gamma_p}{N^{(p-1)/2}} \boldsymbol{x}^{\otimes p}$.*

Then, the posterior on $\boldsymbol{x}$ after observing the tensors $(M^{(p)})_{p \ge 2}$ is of the form $\mu(\boldsymbol{x} = \sigma \mid (M^{(p)})) \propto \exp(H_N(\sigma))$, where

$$H_N(\sigma) = \sum_{p \ge 2} \frac{\gamma_p}{N^{(p-1)/2}} \langle M^{(p)}, \sigma^{\otimes p}\rangle.$$

Then, the joint law of $(H_N, \boldsymbol{x})$ is μ_{pl}.

The above says that conditioned on H_N, the distribution of $\boldsymbol{x}$ (according to μ_{pl}) is simply distributed as a sample according to μ_{H_N}. That is, the spike $\boldsymbol{x}$ resulting in a Hamiltonian $H_N \sim \mu_{\mathsf{pl}}$ is exchangeable with a sample from μ_{H_N}.

The latter of these interpretations will be very useful for us. When dealing with the measure at time t of stochastic localization applied to the p-spin model, the primary issue was that it was unclear how to deal with the sample $\boldsymbol{\sigma}$ drawn from the Gibbs distribution. However, if we could work with the planted p-spin model, this issue would be absent. Indeed, the exchangability of the spike and a sample implies that the law of μ_t applied to the planted model is given by

$$\mu_{H_N, t\boldsymbol{x}+\sqrt{t}\boldsymbol{g}}(\widetilde{\sigma}) \propto \exp\left(H_N(\widetilde{\sigma}) + \langle t\boldsymbol{x} + \sqrt{t}\boldsymbol{g}, \widetilde{\sigma}\rangle\right),$$

where $\boldsymbol{x}$ is the spike hidden in H_N. This decouples the randomness of the external field $t\boldsymbol{x} + \sqrt{t}\boldsymbol{g}$ and the disorder of the Hamiltonian H_N that arises from the Gaussians $(G^{(p)})_{p\geq 2}$.

As was shown in [HMP24, Corollary 3.5] and recalled just below, the planted and null models are mutually contiguous. Thus high-probability statements from one model transfer to the other, and it suffices to study the planted model.

> For all models satisfying (Strict RS), the measures $\mu_{\mathsf{null}}(H_N)$ and $\mu_{\mathsf{pl}}(H_N)$ from Definition 7.7.9 are mutually contiguous, i.e., for any sequence of events $\mathcal{E}_N$, $\mu_{\mathsf{null}}(\mathcal{E}_N) \to 0$ whenever $\mu_{\mathsf{pl}}(\mathcal{E}_N) \to 0$.

The transfer from the p-spin model to the planted model may then be carried out by setting

$$\mathcal{E}_N = \left\{ H_N : \Pr_{\substack{\boldsymbol{\sigma}\sim\mu_{H_N} \\ \boldsymbol{g}\sim\mathcal{N}(0,I_N)}} \left[\left\| \mathsf{Cov}\left(\mu_{H_N, t\boldsymbol{\sigma}+\sqrt{t}\boldsymbol{g}}\right) \right\| > K \right] < e^{-cN^{1/5}} \right\}.$$

This event is very complicated in the null model, but exchangeability makes it tractable in the planted model. In the actual proof, we will require a stronger (quantitative) version of mutual contiguity; see Proposition 7.7.17 for details.

Now, we must understand what the Hamiltonian in the planted model looks like conditioned on the spike.

Fact 7.7.11. *Consider the following process: sample $\boldsymbol{x} \sim S_N$, $\widetilde{H}_N \sim \mu_{\mathsf{null}}$, and define H_N by $H_N(\sigma) = \widetilde{H}_N(\sigma) + N \cdot \xi(R(\boldsymbol{x}, \sigma))$. Then, the joint law of $(H_N, \boldsymbol{x})$ is μ_{pl}.*

Consequently, our goal is to bound the covariance of the distribution

$$\mu_t(\sigma) \propto \exp\left(\widetilde{H}_N(\sigma) + N \cdot \xi\left(R\left(\boldsymbol{x}, \sigma\right)\right) + \langle t\boldsymbol{x} + \sqrt{t}\boldsymbol{g}, \sigma\rangle\right).$$

for $\widetilde{H}_N \sim \mu_{\mathsf{null}}$ with mixture function ξ. Now, define ξ_t by $\xi_t(s) = \xi(s) + ts$, and extend the definition of the p-spin model (7.5) to allow a random linear term. Then,

$$\mu_t(\sigma) \propto \exp\left(\underbrace{\widetilde{H}_{N,t}(\sigma) + N \cdot \xi_t(R(\boldsymbol{x}, \sigma))}_{H_{N,t}(\sigma)}\right),$$

where $\widetilde{H}_{N,t} \sim \mu_{\mathsf{null}}$ with mixture function ξ_t.

The TAP planted model. We now turn to controlling the covariance matrix of these models. As we will see below, it is relatively easier to bound the covariance matrix (in fact, the second moment matrix) of a model with zero or small external field. However, for any time $t > 0$, $H_{N,t}$ has an external field. We will use a method developed in [HMP24] to reduce to the case of a model with zero or small external field.

Let $\boldsymbol{m}^{\mathsf{true}} = \boldsymbol{m}(\mu_t)$. The main intuition of this reduction is that the Gibbs measure concentrates near a codimension-2 band passing through $\boldsymbol{m}^{\mathsf{true}}$ and orthogonal to $\boldsymbol{m}^{\mathsf{true}}$ and $\boldsymbol{x}$, and furthermore *the model on this band is essentially a replica symmetric model with no external field.* Moreover, one expects that both $R(\boldsymbol{m}^{\mathsf{true}}, \boldsymbol{m}^{\mathsf{true}})$ and $R(\boldsymbol{m}^{\mathsf{true}}, \boldsymbol{x})$ concentrate near a value $q_* = q_*(t)$ defined by $\xi_t'(q_*) = \frac{q_*}{1-q_*}$.

However, $\boldsymbol{m}^{\mathsf{true}}$ is a complicated function of $H_{N,t}$, so it is a priori difficult to reason about the joint distribution of $(\boldsymbol{m}^{\mathsf{true}}, H_{N,t})$. Thus, this reduction is formally carried out by conditioning on a *TAP fixed point* $\boldsymbol{m}^{\mathsf{TAP}}$, which will serve as a proxy for $\boldsymbol{m}^{\mathsf{true}}$. Define the TAP free energy

$$\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m}) = H_{N,t}(\boldsymbol{m}) + \frac{N}{2} \cdot \theta(R(\boldsymbol{m},\boldsymbol{m})) + \frac{N}{2} \log(1 - R(\boldsymbol{m},\boldsymbol{m})),$$

where

$$\theta(s) = \xi(1) - \xi(s) - (1-s)\xi'(s).$$

As shown in [HMP24], for sufficiently small constant $\iota > 0$, with probability $1 - e^{-cN}$ $\mathcal{F}_{\mathsf{TAP}}$ has a unique critical point $\boldsymbol{m}^{\mathsf{TAP}}$ in the region $\mathcal{S}_\iota$ defined by $R(\boldsymbol{m},\boldsymbol{m}), R(\boldsymbol{m},\boldsymbol{x}) \in [q_* - \iota, q_* + \iota]$. Due to the existence and uniqueness of $\boldsymbol{m}^{\mathsf{TAP}}$, it becomes possible to relate $H_{N,t}$ to a "TAP-planted model" where one samples $\boldsymbol{m}^{\mathsf{TAP}}$ *first*, and then samples $H_{N,t}$ conditional on $\nabla \mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m}) = 0$:

Lemma 7.7.12 (See Lemma 7.7.23; essentially due to [HMP24])**.** *For any small constant $\iota > 0$, the following holds. For any $H_{N,t}$-measurable event $\mathcal{E}$, if*

$$\sup_{\boldsymbol{m}^{\mathsf{TAP}} \in \mathcal{S}_\iota} \Pr(\mathcal{E} | \nabla \mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m}^{\mathsf{TAP}}) = 0) \to 0,$$

then $\Pr(\mathcal{E}) \to 0$.

Crucially, the conditional law of $H_{N,t}$ in the TAP-planted model is very tractable, as (for a fixed $\boldsymbol{m}^{\mathsf{TAP}}$) $\nabla \mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m}^{\mathsf{TAP}}) = 0$ amounts to a linear constraint on the Gaussian process $H_{N,t}$. The resulting explicit conditional law of $H_{N,t}$ is described in Lemma 7.7.26.

Remark 7.7.13. While it will not be relevant to our purposes, [AMS22, AMS23b, HMP24] have shown that $\boldsymbol{m}^{\mathsf{TAP}}$ typically approximates $\boldsymbol{m}^{\mathsf{true}}$ well, in the sense that $\|\boldsymbol{m}^{\mathsf{true}} - \boldsymbol{m}^{\mathsf{TAP}}\|^2 = O(1)$, thereby justifying the heuristic that $\boldsymbol{m}^{\mathsf{TAP}}$ is a proxy for $\boldsymbol{m}^{\mathsf{true}}$.

Remark 7.7.14. The idea of reducing to a TAP-planted model has also been used beyond the setting of sampling from spherical spin glasses. In the recent work [Hua24], an analogous reduction is used to obtain the capacity of the Ising perceptron. In this application, passage to the TAP-planted model is used to tightly control a partition function rather than to bound a covariance matrix.

Consequently, we can now work within the TAP-planted model. Let H_{TAP} denote the Hamiltonian $H_{N,t}$ after conditioning on $\boldsymbol{x}$ and $\nabla \mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m}^{\mathsf{TAP}}) = 0$. As

$$\mathsf{Cov}(\mu_{H_{\mathsf{TAP}}}) \preceq \mathop{\mathbb{E}}_{\boldsymbol{\sigma} \sim \mu_{H_{\mathsf{TAP}}}} (\boldsymbol{\sigma} - \mathbf{v})(\boldsymbol{\sigma} - \mathbf{v})^\top$$

for any $\mathbf{v} \in \mathbb{R}^N$ (with equality at $\mathbf{v} = \boldsymbol{m}^{\mathsf{true}}$), it suffices to control the operator norm of

$$\mathop{\mathbb{E}}_{\boldsymbol{\sigma} \sim \mu_{H_{\mathsf{TAP}}}} (\boldsymbol{\sigma} - \boldsymbol{m}^{\mathsf{TAP}})(\boldsymbol{\sigma} - \boldsymbol{m}^{\mathsf{TAP}})^\top.$$

Reduction to slices of the sphere. Next, to control the covariance, let us decompose the sphere into codimension-2 slices

$$S(a,b) := \left\{ \sigma \in S_N : R(\sigma, \boldsymbol{m}) = \left(1 + \frac{a}{\sqrt{N}}\right) R(\boldsymbol{m},\boldsymbol{m}),\ R(\sigma, \boldsymbol{x}) = \left(1 + \frac{b}{\sqrt{N}}\right) R(\boldsymbol{x},\boldsymbol{m}) \right\},$$

with the central slice centered at $\boldsymbol{m} = \boldsymbol{m}^{\mathsf{TAP}}$. Let $\mu_{a,b}$ be the measure $\mu_{H_{\mathsf{TAP}}}$ conditioned to lie in the codimension-2 slice $S(a,b)$.

The concentration of the Gibbs measure described in the previous section implies that, viewed as random variables of a sample $\sigma \sim \mu_{H_{\mathsf{TAP}}}$, a and b are well-concentrated around 0. Let $v_{a,b}$ be the center of $S(a,b)$.

The covariance of the distribution may be bounded as

$$
\begin{aligned}
\mathsf{Cov}(\mu) &\preceq \mathbb{E}(\boldsymbol{\sigma}-\boldsymbol{m})(\boldsymbol{\sigma}-\boldsymbol{m})^\top \\
&= \mathop{\mathbb{E}}_{(a,b)} \mathop{\mathbb{E}}_{\boldsymbol{\sigma}\sim\mu_{a,b}} (\boldsymbol{\sigma}-v_{a,b}+v_{a,b}-\boldsymbol{m})(\boldsymbol{\sigma}-v_{a,b}+v_{a,b}-\boldsymbol{m})^\top \\
&\preceq 2\mathop{\mathbb{E}}_{(a,b)} \mathop{\mathbb{E}}_{\boldsymbol{\sigma}\sim\mu_{a,b}} (\boldsymbol{\sigma}-v_{a,b})(\boldsymbol{\sigma}-v_{a,b})^\top + 2\mathop{\mathbb{E}}_{(a,b)} (v_{a,b}-\boldsymbol{m})(v_{a,b}-\boldsymbol{m})^\top \\
&\preceq 2\mathop{\mathbb{E}}_{(a,b)} \mathop{\mathbb{E}}_{\boldsymbol{\sigma}\sim\mu_{a,b}} (\boldsymbol{\sigma}-v_{a,b})(\boldsymbol{\sigma}-v_{a,b})^\top + 2\mathop{\mathbb{E}}_{(a,b)} O\left(a^2+b^2\right). \qquad (7.6)
\end{aligned}
$$

One can interpret $v_{a,b}$ as explaining the variation within the slice originating from the $\boldsymbol{m}$ and $\boldsymbol{x}$ directions. Hence, as alluded to in the previous discussion about the TAP planted model, the key fact is that under $\mu_{a,b}$, the recentered sample $\boldsymbol{\sigma}-v_{a,b}$ is a sample from a spherical spin glass in two lower dimensions, as can be shown by calculating the covariance of the (conditioned) Gaussian process H_{TAP} restricted to this slice. This verification is carried out in Corollary 7.7.29.

These codimension-2 models have the crucial property that the spherical spin glass on the slice $a=b=0$ is a model satisfying (Strict RS) with no external field (i.e. degree-1 term), while nearby slices have a small (random) external field of magnitude $\sqrt{a^2+b^2}$. In particular, the first term of (7.6) requires bounding the second moment of a Gibbs sample from a strictly RS model with small (random) external field. As a result, (7.6) would be bounded if we proved the following.

1. Let H_N be the Hamiltonian of a slightly generalized mixed p-spin model, where we allow the mixture function ξ to have a small linear term $\gamma_1 q$ (in our proofs we allow $\gamma_1^2 \le N^{-4/5}$), such that the non-degree-1 part $\xi_{\sim 1}$ of ξ satisfies (Strict RS). Then, with high probability,

$$
\left\| \mathop{\mathbb{E}}_{\mu_{H_N}} \sigma\sigma^\top \right\|_{\mathsf{op}} = O\left(1+\gamma_1^2 N\right).
$$

   Much of Section 7.8 is dedicated to showing this.

2. The second moments of a and b are $O(1)$. In fact, we will show in Lemma 7.7.40 that they are essentially $O(1)$-subgaussian.

Let us start by explaining how to show subgaussianity.

Subgaussianity of a,b. The distribution ν of (a,b) is given by

$$
\nu(a,b) \propto \exp\left(\log\widehat{Z}_{a,b} + \frac{N-4}{2}\log\left(1-\frac{\|v_{a,b}\|^2}{N}\right) + H_{\mathsf{TAP}}(v_{a,b})\right).
$$

Here, the first term $\log\widehat{Z}_{a,b}$ is the free energy of the $(N-2)$-dimensional p-spin model $\mu_{a,b}$, obtained by restricting $\mu_{H_{\mathsf{TAP}}}$ to the slice $S(a,b)$ and rescaling the distribution to lie on S_{N-2}. The second term is an effective decrement in the free energy caused by the radius of the sphere $S(a,b)$ shrinking for larger values of a and b. The third term is an effective increment in the free energy coming from the energy of H_{TAP} at the center of the slice $S(a,b)$.

For a fixed (a,b), the only random quantities in the definition of ν are the first and third terms. In Theorem 7.7.32, proved in Section 7.8, we show that the first term may essentially be approximated by a deterministic function of the mixture function of $\mu_{a,b}$, at the cost of incurring a small $O(1)$ error. We do not elaborate on the details of this proof in the technical overview; it is similar to that used to bound the covariance (which we explain shortly). The third term is similar, and is a deterministic function plus a small Gaussian, whose variance is $O(a^2+b^2)$.

Given these bounds, we may show that the distribution ν is strongly log-concave at 0 with high probability over the randomness of H_{TAP}. A simple perturbation argument then implies that ν is strongly log-concave in a macroscopic neighborhood of 0, implying subgaussianity.

Covariance bound for strictly RS models with small external fields. The covariance bound has now boiled down to bounding $\|M\|_{\mathsf{op}}$, for $M = \mathbb{E}_{\mu_{H_N}}[\sigma\sigma^\top]$ the second moment matrix of model satisfying (Strict RS) with small external field. Note that M is a H_N-measurable random variable.

The proof proceeds in two high level steps, which we carry out in Section 7.8.

1. We show using the second moment method that with positive probability over H_N, $\|M\|_{\mathsf{op}}$ is bounded.

2. Using a much simpler argument, we can show that $\|M\|_{\mathsf{op}}$ is essentially $O(N^{-1/10})$-Lipschitz in the disorder. Hence, by gaussian concentration, it concentrates very well around its expectation (which is $O(1)$ by the positive probability bound).

Let us elaborate a bit more on the proof of the first point above. It turns out that, under the condition (Strict RS) with small external field, the leading order contribution to M comes from the degree-2 part of the Hamiltonian $H_{N,2}(\sigma) = \frac{\gamma_2}{N^{1/2}} \sum_{i,j} \boldsymbol{g}_{i,j}\sigma_i\sigma_j$. We will ultimately reduce the study of the covariance matrix of μ_{H_N} to that of $\mu_{H_{N,2}}$, and then show boundedness of $\mathsf{Cov}(\mu_{H_{N,2}})$ using random matrix theory. A similar strategy of isolating the degree-2 component of H_N was used to study the partition function and magnetization of strictly RS models in [HMP24].

Degree-2 behavior. Let us discuss the typical behavior of the covariance of $\mu_{H_{N,2}}$. Define the degree-2 Gibbs measure

$$\mathsf{d}\mu_{H_{N,2}}(\sigma) \propto \exp(H_{N,2}(\sigma))\mathsf{d}\rho(\sigma),$$

with corresponding partition function $Z_{N,2} = \int \exp(H_{N,2}(\sigma))\mathsf{d}\rho(\sigma)$. This is the spherical Sherrington-Kirkpatrick model with interaction matrix $A = \nabla^2 H_N(0)$; note that A is a scaled GOE matrix. Observe that if we shift A by a constant multiple of the identity $\gamma \mathrm{Id}_N$, the measure does not change, as it is supported on S_N. The crucial observation is the following:

> For a careful choice of γ, the measure $\mathsf{d}\mu_{H_{N,2}}(\sigma) \propto \exp(-\frac{1}{2}\langle\sigma, (\gamma\mathrm{Id}_N - A)\sigma\rangle)\mathsf{d}\rho(\sigma)$ looks like a Gaussian with covariance $(\gamma\mathrm{Id}_N - A)^{-1}$.

In fact, we will see that it suffices to pick $\gamma = 1 + \xi''(0)$. The typical value of $\|x\|_2^2$, where $x \sim \mathcal{N}(0, \gamma\mathrm{Id}_N - A)^{-1}$, is equal to $\mathsf{Tr}(\gamma\mathrm{Id}_N - A)^{-1}$. By approximating this trace using the semicircle law for the eigenvalues of A and the explicit choice of γ, we see that $\|x\|_2^2 \approx N$, which justifies the heuristic that this Gaussian approximates the spherical distribution $\mu_{H_{N,2}}$.

For the above discussion to be well-defined, we require that $\gamma\mathrm{Id}_N - A$ is positive definite, which can only occur if the maximum eigenvalue of A is bounded above by $\gamma = 1 + \xi''(0)$. By standard concentration inequalities about the maximum eigenvalue of a GOE matrix, this holds with a constant margin with exponentially good probability. Thus, at least for typical realizations of $H_{N,2}$, the covariance will have bounded operator norm. To make this rigorous, we will use the Laplace transform to precisely control the moments of the overlaps, as was previously done in [BL16, HMP24].

Reduction to degree-2. Below, we give some justification for why one should expect to be able to reduce to the degree-2 behavior. We will heuristically argue this by showing that the partition function Z_N is essentially controlled by the degree-2 portion.

To simplify the discussion, let us assume that we are in a $2 + p$ spin model, so that $H_N(\sigma) = H_{N,2}(\sigma) + H_{N,p}(\sigma)$, where $H_{N,p}(\sigma) = \frac{\gamma_p}{N^{(p-1)/2}} \sum_{i_1,\ldots,i_p=1}^N \boldsymbol{g}_{i_1,\ldots,i_p}\sigma_{i_1}\cdots\sigma_{i_p}$. The corresponding mixture function decomposes as $\xi(q) = \gamma_2^2 q^2 + \xi_{\sim 2}(q)$, so that $\xi_{\sim 2}(q) = \gamma_p^2 q^p$ corresponds to the non degree-2 part of the mixture function. It turns out that, once we condition on $H_{N,2}$ (and hence the value of $Z_{N,2}$), the full partition function Z_N is essentially deterministic. Indeed, we will show in Proposition 7.8.2 that with very high probability,

$$Z_N \approx Z_{N,2} e^{N\xi_{\sim 2}(1)/2}$$

To see why this is reasonable, let us consider the first two moments of Z_N conditioned on the degree-2 Hamiltonian $H_{N,2}$. Indeed, let $\mathbb{E}_{\sim 2}$ denote expectation with respect to $H_{N,p}$ conditioned on $H_{N,2}$. Standard gaussian MGF calculations yield $\mathbb{E}_{\sim 2} Z_N = e^{N\xi_{\sim 2}(1)/2} Z_{N,2}$ and

$$\begin{aligned}\mathop{\mathbb{E}}_{\sim 2}[Z_N^2] &= Z_{N,2}^2 e^{N\xi_{\sim 2}(1)} \int \exp(N\xi_{\sim 2}(R(\sigma^1,\sigma^2)))\mathsf{d}\rho^{\otimes 2}(\sigma^1,\sigma^2) \\ &= (\mathop{\mathbb{E}}_{\sim 2} Z_N)^2 \int \exp(N\gamma_p^2 R(\sigma^1,\sigma^2)^p)\mathsf{d}\rho^{\otimes 2}(\sigma^1,\sigma^2).\end{aligned}$$

At sufficiently high temperatures, the typical overlap behavior $R(\sigma^1,\sigma^2) \asymp N^{-1/2}$, where σ^1,σ^2 are iid draws from μ_{H_N}. This matches the overlap behavior at infinite temperature, where the Gibbs distribution is uniform. Then, pretending that $R(\sigma^1,\sigma^2) = cN^{-1/2}$ for all σ^1,σ^2, we obtain that

$$\mathop{\mathbb{E}}_{\sim 2}[Z_N^2] \approx (\mathop{\mathbb{E}}_{\sim 2} Z_N)^2 \exp(c^p \gamma_p^2 N^{1-p/2}).$$

Since $p \geq 3$, it follows that, conditional on $H_{N,2}$, the conditional variance of Z_N is tiny compared to its conditional expectation. In summary, we see that the higher degree portions of the partition function have negligible contributions to the fluctuations of Z_N, so that the typical behavior of Z_N is controlled by $Z_{N,2}$.

Turning now to the covariance bound, we will control the (i,j)th covariance entry $M_{i,j} := \int \sigma_i\sigma_j e^{H_N(\sigma)}\mathsf{d}\rho(\sigma)$. A crucial fact is that, by rotational invariance of the sphere and gaussians, we can rotate to the eigenbasis of $A = \nabla^2 H_N(0)$ so that A becomes diagonal. When A is diagonal, one can in fact show that

$$\mathop{\mathbb{E}}_{\sim 2}[M_{i,j}^2] \lesssim \frac{1}{N^2}(\mathop{\mathbb{E}}_{\sim 2} M_{i,j})^2,$$

where $\mathbb{E}_{\sim 2} M_{i,j}$ can be interpreted as (up to normalization) the predicted (i,j)th covariance entry by just looking at the degree-2 randomness; see Propositions 7.8.17 and 7.8.18 for details. It follows that the Frobenius norm error of the true covariance compared to the degree-2 covariance is $O(1)$. Combined with the typical behavior of the degree-2 covariance being essentially the diagonal matrix $((1+\xi''(0))\mathrm{Id}_N - A)^{-1}$, we conclude an $O(1)$ covariance bound for μ_{H_N}.

Although this direct moment approach can be made rigorous at sufficiently high temperature, it will not cover the entire regime (Strict RS) of our main theorem. To deal with this, we will use the *free energy typical truncation* recently introduced by [HS23b]. The main idea is that, while pairs σ^1,σ^2 with overlap $R(\sigma^1,\sigma^2) \asymp N^{-1/2}$ do not necessarily dominate the second moment $\mathbb{E}[Z_N^2]$ throughout the regime (Strict RS), there is a truncation $\widetilde{Z}_N$ accounting for nearly all of Z_N, whose second moment is dominated by such pairs. We defer the details to the following sections.

7.7.2 Null models, planted models, and contiguity

As described in the technical overview, we will need a quantitative strengthening of contiguity between the null and planted models. For convenience, let us restate the definition of the planted model.

Definition 7.7.9 (Planted p-spin model)**.** *The planted measure μ_{pl} is a joint law over a Hamiltonian H_N and a* spike *$\boldsymbol{x} \in S_N$ given by*

$$\mathsf{d}\mu_{\mathsf{pl}}(H_N, x) \propto \exp\left(H_N(x)\right) \cdot \mathsf{d}\rho(x) \cdot \mathsf{d}\mu_{\mathsf{null}}(H_N),$$

where ρ is the uniform measure over S_N and μ_{null} is the law over p-spin Hamiltonians with mixture function ξ. We frequently abuse notation to let $\mu_{\mathsf{pl}}(H_N)$ denote the marginal of μ_{pl} on H_N.

Remark 7.7.15 (Interpretation of planted model)**.** Equivalently, the planted measure μ_{pl} can be described as follows.

- Sample $\boldsymbol{x} \sim S_N$.
- Sample $\widetilde{H}_N \sim \mu_{\mathsf{null}}$.
- Define H_N by $H_N(\sigma) = \widetilde{H}_N(\sigma) + N \cdot \xi(R(\boldsymbol{x},\sigma))$.

The following Bayesian interpretation of μ_{pl} will make the planted model amenable to explicit calculation. For $(\boldsymbol{x}, H_N)$ sampled from μ_{pl}, the posterior distribution $\boldsymbol{x}|H_N$ is described by the density:

$$\mathsf{d}\mu_{\boldsymbol{x}|H_N}(\sigma) \propto \exp(H_N(\sigma))\mathsf{d}\rho(\sigma)\,.$$

Therefore, the distribution of (H_N, σ) for $\sigma \sim \mu_{H_N}$ is identical to that of $(H_N, \boldsymbol{x})$.

In order to show the probability bound of $1 - e^{-cN^{1/5}}$ in Lemma 7.7.7, we will prove the following quantitative strengthening of mutual contiguity, under the following quantitative strict RS condition. Note that, since the proof of Theorem 7.7.2 union bounds over $\mathsf{poly}(N)$ many values of β, quantitative control of the error in Lemma 7.7.7 is needed to carry out the proof.

Condition 7.7.16 (ε-strict replica symmetry)**.** We say ξ is ε-strictly replica symmetric if for all $q \in (0, 1)$,

$$\frac{1}{q^2} \cdot (\xi(q) + q + \log(1 - q)) \leq -\varepsilon/2. \tag{7.7}$$

Under this assumption, we prove the following quantitative contiguity result in Section 7.8.

Proposition 7.7.17 (Quantitative contiguity)**.** *Under Condition 7.7.16, there exists $c = c(\varepsilon) > 0$ such that for any event $\mathcal{E}$, if $\mu_{\mathsf{pl}}(\mathcal{E}) = p$, then $\mu_{\mathsf{null}}(\mathcal{E}) \leq e^{-cN^{1/5}} + e^{\frac{1}{c}\sqrt{\log\frac{2}{p}}}p$.*

Thus, from now on, we work under the planted model. One reason the planted model is easier to work with is because of the following lemma, which provides a simple description of the distribution of μ_t (by describing the distribution of the external field at time t) in the planted model.

Lemma 7.7.18. *Let μ_t be the distribution after running stochastic localization with the* Id *driving matrix for time t initialized at μ_{H_N}. Then μ_t arises as the Gibbs distribution of the Hamiltonian $H_{N,t}(\sigma)$:*

$$H_{N,t}(\sigma) = H_N(\sigma) + \langle \boldsymbol{y}_t, \sigma\rangle\,,$$

where

$$(H_N, \boldsymbol{y}_t) \overset{d}{=} (H_N, t\boldsymbol{x} + \sqrt{t}\boldsymbol{g})\,,$$

where $\boldsymbol{x} \sim S_N$, $H_N \sim \mu_{\mathsf{pl}}(\cdot|\boldsymbol{x})$, and $\boldsymbol{g} \sim \mathcal{N}(0, \mathrm{Id}_N)$.

Notation 7.7.19 ($\mu_{\mathsf{pl},t}$, $\xi_t(q)$, $\gamma(q)$)**.** We will use $\mu_{\mathsf{pl},t}$ to denote the distribution of the pair $(H_{N,t}, \boldsymbol{x})$, $\xi_t(q) = \xi(q) + tq$ to refer to the mixture function of $H_{N,t}$, and $\gamma(q)$ to refer to the function $q\xi_t'(q)$.

In the subsequent sections, we will prove a high probability covariance bound for μ_t at a *fixed* time t under the planted model.

Lemma 7.7.20. *There exist universal constants c, T, K, such that for any $t \in [0, T]$, with probability at least $1 - e^{-cN^{1/5}}$ over the randomness of H_N drawn from $\mu_{\mathsf{pl},t}$, we have $\|\mathsf{Cov}(\mu_t)\| \leq K$.*

We now have all the necessary ingredients to prove the covariance bound along the entire localization path for the null model.

Proof of Lemma 7.7.7. Define $\mathcal{T}$ as the discrete set $\{iT/\delta : 1 \leq i \leq 1/\delta, i \in \mathbb{Z}\}$ for $\delta = N^{-100}$. We will prove:

$$\Pr_{H_N}\left[\Pr_{(\mu_t)|H_N}\left[\|\mathsf{Cov}(\mu_t)\|_{\mathsf{op}} < K \text{ for all } t \in \mathcal{T}\right] \geq 1 - e^{-cN^{1/5}}\right] \geq 1 - e^{-cN^{1/5}}\,.$$

A simple continuity argument can be used to derive the desired statement from the above. By taking a union bound over all elements of $\mathcal{T}$, along with Proposition 7.7.17 and Lemma 7.7.20, we can conclude:

$$\mathop{\mathbb{E}}_{H_N}\Pr_{(\mu_t)|H_N}\left[\|\mathsf{Cov}(\mu_t)\|_{\mathsf{op}} > K \text{ for some } t \in \mathcal{T}\right] \leq e^{-2cN^{-1/5}}.$$

The resulting statement then follows from Markov's inequality on the random variable

$$\Pr_{(\mu_t)|H_N}\left[\|\mathsf{Cov}(\mu_t)\|_{\mathsf{op}} > K \text{ for some } t \in \mathcal{T}\right]. \qquad \square$$

7.7.3 TAP planted models

In this section, we formally introduce the TAP planted model and relate it to the planted model from the previous section.

Definition 7.7.21. Let H_N be a planted Hamiltonian with mixture function ξ, and define

$$\theta(s) = \xi(1) - \xi(s) - (1-s)\xi'(s).$$

The associated *TAP free energy* is defined by

$$\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m}) = H_N(\boldsymbol{m}) + \frac{N}{2} \cdot \theta\left(R(\boldsymbol{m},\boldsymbol{m})\right) + \frac{N}{2} \cdot \log\left(1 - R(\boldsymbol{m},\boldsymbol{m})\right).$$

While the TAP free energy is interesting for a multitude of reasons, we will be interested in it because its fixed points provide a good proxy for the mean. Furthermore, the linearity of the TAP free energy in the Gaussian coefficients of the Hamiltonian provides certain desirable properties (that using the true mean would not allow).

Fact 7.7.22 ([HMP24, Fact 4.2])**.** *Let ξ be a mixture function satisfying the condition* (SL)*. For any $t \in [0,\infty)$, let $\xi_t(q) = \xi(q) + tq$. Then there is a unique solution in $[0,1)$, which we denote $q_* = q_*(t)$, to*

$$\xi_t'(q_*) = \frac{q_*}{1-q_*}.$$

Lemma 7.7.23. *For any $K > 0$, sufficiently small (constant) $\iota > 0$ and $\boldsymbol{x} \in S_N$:*

$$\begin{aligned}
&\Pr_{\mu_{\mathsf{pl},t}}\left[\left\|\mathsf{Cov}(\mu_{H_{N,t}})\right\|_{\mathsf{op}} \geq K\right] \\
&\qquad \leq C \cdot \sup_{\boldsymbol{m}\in\mathcal{S}_\iota} \Pr_{\mu_{\mathsf{pl},t}|\boldsymbol{x}}\left[\left\|\mathsf{Cov}(\mu_{H_{N,t}})\right\|_{\mathsf{op}} \geq K \wedge \mathcal{E}_\iota \mid \nabla\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m}) = 0\right]^{1/2} + 2e^{-cN},
\end{aligned}$$

where

$$\mathcal{S}_\iota = \mathcal{S}_\iota(\boldsymbol{x}) \coloneqq \left\{\boldsymbol{m} \in \mathbb{R}^N : |R(\boldsymbol{m},\boldsymbol{m}) - q_*|, |R(\boldsymbol{m},\boldsymbol{x}) - q_*| < \iota\right\},$$

and $\mathcal{E}_\iota$ is the event that $\mathcal{F}_{\mathsf{TAP}}$ has a unique critical point $\boldsymbol{m}^{\mathsf{TAP}}$ in $\mathcal{S}_\iota$, and that

$$\Pr_{\boldsymbol{\sigma}\sim\mu_{H_{N,t}}}\left[R(\boldsymbol{\sigma},\boldsymbol{m}^{\mathsf{TAP}}), R(\boldsymbol{\sigma},\boldsymbol{x}) \in [q_* - \iota, q_* + \iota]\right] \geq 1 - e^{-cN}.$$

Proof. The above statement is effectively due to [HMP24, Propositions 4.4(d) and 4.5(a)]. For the reader's convenience, we include the steps to arriving at the above statement. For any event $\mathcal{E}$ (and in particular, for the event $\mathcal{E}$ defined in [HMP24, Proposition 4.4]), we have:

$$\Pr\left[\left\|\mathsf{Cov}(\mu_{H_{N,t}})\right\|_{\mathsf{op}} \geq K\right] \leq \Pr\left[\left\|\mathsf{Cov}(\mu_{H_{N,t}})\right\|_{\mathsf{op}} \geq K \wedge \mathcal{E}_\iota \wedge \mathcal{E}\right] + \Pr\left[\overline{\mathcal{E}}\right] + \Pr\left[\overline{\mathcal{E}_\iota}\right].$$

The desired statement follows by observing that $\Pr\left[\overline{\mathcal{E}_\iota}\right] \leq e^{-cN}$ by [HMP24, Proposition 4.5(a)], $\Pr\left[\overline{\mathcal{E}}\right] \leq e^{-cN}$ by [HMP24, Proposition 4.4], and applying [HMP24, Proposition 4.4(d)] with $X = \mathbf{1}\left[\left\|\mathsf{Cov}(\mu_{H_{N,t}})\right\|_{\mathsf{op}} \geq K \wedge \mathcal{E}_\iota\right]$. $\square$

Lemma 7.7.23 reduces our task to studying the covariance matrix in a *conditional* planted model.

Notation 7.7.24 (μ_{TAP}, H_{TAP}, $q_{\boldsymbol{m}}$, $q_{\boldsymbol{x}}$)**.** For $\boldsymbol{x} \sim S_N$ and $\boldsymbol{m} \in \mathbb{R}^N$, we consider the distribution $\mu_{\mathsf{TAP},\boldsymbol{x},\boldsymbol{m}}$ of H_{TAP} for $H_{\mathsf{TAP}} \sim (\mu_{\mathsf{pl},t}|\boldsymbol{x}, \nabla\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m}) = 0)$. We use $q_{\boldsymbol{m}}$ and $q_{\boldsymbol{x}}$ to refer to $R(\boldsymbol{m},\boldsymbol{m})$ and $R(\boldsymbol{m},\boldsymbol{x})$ respectively.

Lemma 7.7.25. *Let $\boldsymbol{x} \in S_N$, let $\mathcal{S}_\iota$ be as in Lemma 7.7.23, and let $\boldsymbol{m} \in \mathcal{S}_\iota$. Then for an absolute constant $K > 0$,*

$$\Pr_{H_{\mathsf{TAP}}\sim\mu_{\mathsf{TAP},\boldsymbol{x},\boldsymbol{m}}}\left[\|\mathsf{Cov}(\mu_{H_{\mathsf{TAP}}})\| \geq K \wedge \mathcal{E}_\iota\right] \leq e^{-cN^{1/5}}.$$

Proof of Lemma 7.7.20. The statement is immediate from Lemmas 7.7.23 and 7.7.25. □

The rest of this section is dedicated to proving Lemma 7.7.25. As a first step, we determine the law of the typical Hamiltonian sampled from μ_{TAP}. We prove the following lemma in Appendix 7.B — it follows by routine calculations, using the form of the law of a Gaussian process conditioned on the value of a linear function of it. Recall that $\xi_t(q) = \xi(q) + tq$.

Lemma 7.7.26. *The law of Hamiltonian $H_{\mathsf{TAP}} \sim \mu_{\mathsf{TAP},\boldsymbol{x},\boldsymbol{m}}$ is described by a Gaussian process $(H_{\mathsf{TAP}}(\sigma))_{\sigma\in S_N}$ defined by*

$$\mathbb{E}\ H_{\mathsf{TAP}}(\sigma) = N\xi_t(R(x,\sigma)) - \langle \boldsymbol{x}, v(\sigma)\rangle \cdot \xi_t'(q_{\boldsymbol{x}}) - \frac{\xi_t'(R(\boldsymbol{m},\sigma))}{\gamma'(q_{\boldsymbol{m}})} \cdot \langle \boldsymbol{m}, \sigma\rangle \cdot \left(\theta'(q_{\boldsymbol{m}}) - \frac{1}{1-q_{\boldsymbol{m}}}\right)$$
$$\begin{aligned}\frac{1}{N}\mathsf{Cov}(H_{\mathsf{TAP}}(\sigma), H_{\mathsf{TAP}}(\sigma')) = \xi_t(R(\sigma,\sigma')) - R(\sigma,\sigma')\frac{\xi_t'(R(\boldsymbol{m},\sigma))\xi_t'(R(\boldsymbol{m},\sigma'))}{\xi_t'(q_{\boldsymbol{m}})}& \\ + \frac{\xi_t''(q_{\boldsymbol{m}})}{\gamma'(q_{\boldsymbol{m}})\xi_t'(q_{\boldsymbol{m}})}\gamma(R(\boldsymbol{m},\sigma))\gamma(R(\boldsymbol{m},\sigma')),&\end{aligned}$$

where

$$\begin{aligned} v(\sigma) &\coloneqq \frac{\xi_t'(R(\boldsymbol{m},\sigma))}{\xi_t'(q_{\boldsymbol{m}})}\left[I - \frac{\xi_t''(q_{\boldsymbol{m}})}{\gamma'(q_{\boldsymbol{m}})} \cdot \frac{\boldsymbol{m}\boldsymbol{m}^\top}{N}\right]\sigma \\ \gamma(q) &\coloneqq q \cdot \xi_t'(q)\,. \end{aligned}$$

For the proofs below, it will also be helpful to consider Hamiltonians with a linear term representing an external field. For a sequence $\gamma_1, \gamma_2, \ldots, \gamma_{p_*}$, consider the following generalization of H_N from (7.5):

$$H_N(\sigma) \coloneqq \sum_{p\geq 1} \frac{\gamma_p}{N^{(p-1)/2}} \sum_{i_1,\ldots,i_p=1}^{N} \boldsymbol{g}_{i_1,\ldots,i_p}\sigma_{i_1}\cdots\sigma_{i_p}. \tag{7.8}$$

This has mixture function

$$\xi(s) = \sum_{p\geq 1}\gamma_p^2 q^p.$$

We will write $\xi_{\sim 1}(s) = \sum_{p\geq 2}\gamma_p^2 q^p$ for the part of ξ with degree at least 2, and extend Condition 7.7.16 to such ξ as follows.

Condition 7.7.27 (ε-strict replica symmetry)**.** We say ξ is ε-strictly replica symmetric if $\gamma_1^2 \leq N^{-4/5}$ and $\xi_{\sim 1}$ satisfies Condition 7.7.16.

7.7.4 Slices in TAP planted models

For succinctness, we shall fix $\boldsymbol{x} \in S_N$ and $\boldsymbol{m} \in \mathcal{S}_\iota(\boldsymbol{x})$, and use μ_{TAP} to refer to the distribution $\mu_{\mathsf{TAP},\boldsymbol{x},\boldsymbol{m}}$. For $H_{\mathsf{TAP}} \sim \mu_{\mathsf{TAP}}$, we are interested in bounding the covariance of $\mu_{H_{\mathsf{TAP}}}$. To reason about $\mu_{H_{\mathsf{TAP}}}$, we write it as a mixture of distributions over $(N-2)$-dimensional slices of S_N. For $a, b \in \mathbb{R}$, we define

$$S(a,b) \coloneqq \left\{\sigma \in S_N : R(\sigma,\boldsymbol{m}) = \left(1 + \frac{a}{\sqrt{N}}\right)q_{\boldsymbol{m}},\ R(\sigma,\boldsymbol{x}) = \left(1+\frac{b}{\sqrt{N}}\right)q_{\boldsymbol{x}}\right\}.$$

Let $r_{a,b}$ refer to the radius of this slice, which is equal to

$$r_{a,b} = \sqrt{1 - q_{\boldsymbol{m}}\left(1+\frac{a}{\sqrt{N}}\right)^2 - \frac{q_{\boldsymbol{m}}q_{\boldsymbol{x}}^2}{q_{\boldsymbol{m}} - q_{\boldsymbol{x}}^2}\left(\frac{a-b}{\sqrt{N}}\right)^2}.$$

Note in particular that

$$q_{\boldsymbol{m}}\left(1+\frac{a}{\sqrt{N}}\right)^2 + \frac{q_{\boldsymbol{m}}q_{\boldsymbol{x}}^2}{q_{\boldsymbol{m}} - q_{\boldsymbol{x}}^2}\left(\frac{a-b}{\sqrt{N}}\right)^2 \geq q_{\boldsymbol{m}}\left(1+\frac{a}{\sqrt{N}}\right)^2. \tag{7.9}$$

We refer to the uniform distribution on this slice as $\rho_{a,b}$, and the partition function on the slice as

$$Z_{a,b} := \mathop{\mathbb{E}}_{\sigma \sim \rho_{a,b}} \exp(H_{\mathsf{TAP}}(\sigma)) \,.$$

With this definition, the partition function of the original Hamiltonian is given by

$$Z = \Lambda_N \int Z_{a,b} r_{a,b}^{N-4} \mathsf{d}(a,b)$$

for some fixed number Λ_N depending only on N.

Remark 7.7.28. To see why we scale by $r_{a,b}^{N-4}$, observe that when H_{TAP} is the constant-0 Hamiltonian, the resulting distribution on the sphere should be uniform. The distribution restricted to each slice must also be uniform. However, not all slices are weighted equally — slice that have smaller radii must be downweighted accordingly, with this weighting proportional to $r_{a,b}^{N-4}$ for $S(a,b)$.[1]

Use ν to refer to the distribution over (a,b) where $\mathsf{d}\nu(a,b) \propto Z_{a,b} r_{a,b}^{N-4} \mathsf{d}(a,b)$, and $\mu_{a,b}$ to refer to the distribution $\mu_{H_{\mathsf{TAP}}}$ restricted to $S(a,b)$. Now, we can write $\mu_{H_{\mathsf{TAP}}}$ as the following mixture:

$$\mu_{H_{\mathsf{TAP}}} = \mathop{\mathbb{E}}_{(a,b)\sim\nu} \mu_{a,b} \,.$$

We will need coarse understanding of the tails of ν, and fine understanding of the distribution $\mu_{a,b}$ for small a and b.

Now, let us probe the distribution $\mu_{a,b}$.

$$\frac{\mathsf{d}\mu_{a,b}}{\mathsf{d}\rho_{a,b}}(\sigma) = \frac{\exp(H_{\mathsf{TAP}}(\sigma))}{\mathbb{E}_{\sigma\sim\rho_{a,b}} \exp(H_{\mathsf{TAP}}(\sigma))}$$

Since $S(a,b)$ can be naturally identified with S_{N-2}, the first step to understanding $\mu_{a,b}$ is to express it as a p-spin model on S_{N-2}. To do so, we will verify that *some* Hamiltonian that gives rise to $\mu_{a,b}$ has a mixture function that is given by a polynomial in the overlap. We can write any $\sigma \in S(a,b)$ as:

$$\sigma = \sqrt{N} \cdot v(a,b) + \underbrace{\sqrt{1 - \|v(a,b)\|^2}}_{r_{a,b}} \sigma_\perp$$

for $\sigma_\perp \in S_N$ orthogonal to $\boldsymbol{m}$ and $\boldsymbol{x}$, and for $v(a,b)$ in the span of $\boldsymbol{m}$ and $\boldsymbol{x}$. Let Q be an isometric linear transformation that maps S_{N-2} to $S_N \perp \{\boldsymbol{m}, \boldsymbol{x}\}$. We can write $H_{\mathsf{TAP}}(\sigma) = H_{\mathsf{TAP}}(v(a,b) + r_{a,b} Q\tau)$ (where $\tau \in S_{N-2}$). The following is a consequence of Lemma 7.7.26, and is proved in Appendix 7.B.

Corollary 7.7.29. *For a fixed choice of a and b, the Gaussian process $(H_{\mathsf{TAP}}(v(a,b) + r_{a,b}Q\tau))_{\tau \in S_{N-2}}$ is described by the following law.*

- *Let $H_{a,b}$ be a spherical p-spin Hamiltonian with mixture function $\xi_{a,b}$ given by:*

$$\xi_{a,b}(s) := \xi_t\Big(\|v(a,b)\|^2 + r_{a,b}^2 s\Big) - \xi_t\Big(\|v(a,b)\|^2\Big) - s \cdot \frac{r_{a,b}^2 \xi_t'\Big(q_{\boldsymbol{m}} \cdot \Big(1 + \frac{a}{\sqrt{N}}\Big)\Big)^2}{\xi_t'(q_{\boldsymbol{m}})} \,.$$

- *Let* $V(a,b) := \xi_t\Big(\|v(a,b)\|^2\Big) - \|v(a,b)\|^2 \cdot \frac{\xi_t'\left(\left(1+\frac{a}{\sqrt{N}}\right)q_{\boldsymbol{m}}\right)^2}{\xi_t'(q_{\boldsymbol{m}})} + \frac{\xi_t''(q_{\boldsymbol{m}})}{\gamma'(q_{\boldsymbol{m}})\xi_t'(q_{\boldsymbol{m}})} \cdot \gamma\left(\left(1 + \frac{a}{\sqrt{N}}\right) q_{\boldsymbol{m}}\right)^2.$

The law of $H_{\mathsf{TAP}}(v(a,b) + r_{a,b}Q\tau)$ is the same as that of $H_{a,b}(\tau) + \sqrt{N} \cdot g_{a,b} + \mathbb{E}_{\mu_{\mathsf{TAP}}} H_{\mathsf{TAP}}(v(a,b) + r_{a,b}Q\tau)$ where $g_{a,b}$ is a centered Gaussian of variance $V(a,b)$ independent of $H_{a,b}$.

Now, H_{TAP} is described by the collection $(H_{a,b}, g_{a,b})_{a,b}$. This is *not* an independent collection of random variables. The only structural properties of this collection we will use are:

[1] The constant of proportionality here is something depending only on N.

- For each $a, b \in \mathbb{R}$, we have $H_{a,b}$ and $g_{a,b}$ are independent.
- For any a, b, we have $g_{a,b} = g_{0,0} + \widehat{g}_{a,b}$, where $\widehat{g}_{a,b}$ is a centered Gaussian of variance $O\left(\frac{a^4+b^4}{N^2}\right)$.

We will first give an explicit form for $v(a,b)$.

Fact 7.7.30. *We have*

$$\sqrt{N} \cdot v(a,b) = \boldsymbol{m} \cdot \left(1 + \frac{aq_{\boldsymbol{m}} - bq_{\boldsymbol{x}}^2}{\sqrt{N}(q_{\boldsymbol{m}} - q_{\boldsymbol{x}}^2)}\right) + \frac{q_{\boldsymbol{m}} q_{\boldsymbol{x}}}{q_{\boldsymbol{m}} - q_{\boldsymbol{x}}^2} \boldsymbol{x} \cdot \left(\frac{b-a}{\sqrt{N}}\right).$$

Lemma 7.7.31. *There exists $\varepsilon = \varepsilon(\xi) > 0$ such that for all $t \geq 0$ and all $|a|, |b| \leq \varepsilon N^{1/10}$, the mixture function $\xi_{a,b}$ defined in Corollary 7.7.29 (recall this implicitly depends on t) is ε-strictly replica symmetric (Condition 7.7.27).*

Proof. We will first show $\xi'_{a,b}(0) \leq N^{-4/5}$. We calculate:

$$\begin{aligned}
\xi'_{a,b}(0) &= r_{a,b}^2 \xi'_t\left(\|v(a,b)\|^2\right) - \frac{r_{a,b}^2 \xi'_t\left(q_{\boldsymbol{m}}\left(1 + \frac{a}{\sqrt{N}}\right)\right)^2}{\xi'_t(q_{\boldsymbol{m}})} \\
&\leq \xi'_t\left(q_{\boldsymbol{m}} + \frac{2aq_{\boldsymbol{m}}}{\sqrt{N}} + O\left(\frac{a^2+b^2}{N}\right)\right) - \frac{\xi'_t\left(q_{\boldsymbol{m}}\left(1 + \frac{a}{\sqrt{N}}\right)\right)^2}{\xi'_t(q_{\boldsymbol{m}})}.
\end{aligned}$$

Here, the inequality follows from the fact that $\|v(a,b)\|^2 = q_{\boldsymbol{m}}\left(1 + \frac{a}{\sqrt{N}}\right)^2 + O\left(\frac{a^2+b^2}{N}\right)$, ξ'_t is non-decreasing, and $r_{a,b}^2 \leq 1$. Now, we have, using the $O(1)$-Lipschitzness of ξ'_t,

$$\xi'_t\left(\|v(a,b)\|^2\right) = \xi'_t(q_{\boldsymbol{m}}) + \xi''_t(q_{\boldsymbol{m}}) \cdot \frac{2aq_{\boldsymbol{m}}}{\sqrt{N}} + O\left(\frac{a^2+b^2}{N}\right)$$

and

$$\xi'_t\left(q_{\boldsymbol{m}}\left(1 + \frac{a}{\sqrt{N}}\right)\right)^2 = \xi'_t(q_{\boldsymbol{m}})^2 + \frac{2aq_{\boldsymbol{m}}}{\sqrt{N}} \cdot \xi''_t(q_{\boldsymbol{m}}) \cdot \xi'_t(q_{\boldsymbol{m}}) + O\left(\xi'_t(q_{\boldsymbol{m}}) \cdot \frac{a^2+b^2}{N}\right).$$

Thus $\xi'_{a,b}(0) = O(\frac{a^2+b^2}{N})$. Setting ε sufficiently small ensures $\xi'_{a,b}(0) \leq N^{-4/5}$. Next, we show $(\xi_{a,b})_{\sim 1}$ satisfies Condition 7.7.16. Since ξ satisfies (SL), there exists sufficiently small $\varepsilon = \varepsilon(\xi)$ such that $\xi''(q) \leq \frac{1-\varepsilon}{(1-q)^2}$ for all $q \in [0,1)$. Then,

$$\begin{aligned}
\xi''_{a,b}(q) &= \left(1 - \|v(a,b)\|^2\right)^2 \xi''_t\Big(\|v(a,b)\|^2 + \left(1 - \|v(a,b)\|^2\right)q\Big) \\
&\leq \left(1 - \|v(a,b)\|^2\right)^2 \cdot \frac{1-\varepsilon}{\left(1 - \|v(a,b)\|^2 - \left(1 - \|v(a,b)\|^2\right)q\right)^2} \\
&= \frac{1-\varepsilon}{(1-q)^2} \leq \frac{1}{(1-q)^2} - \varepsilon.
\end{aligned}$$

Integrating twice shows

$$(\xi_{a,b})_{\sim 1}(q) + q + \log(1-q) \leq \frac{1}{2}\varepsilon q^2. \qquad \square$$

We are now ready to bound $\mathsf{Cov}(\mu_{H_{\mathsf{TAP}}})$. First, recall that for any distribution μ over $\mathbb{R}^N$ and any vector $v \in \mathbb{R}^N$, we have $\mathsf{Cov}(\mu) \preceq \mathbb{E}_{\sigma \sim \mu}(\sigma - v)(\sigma - v)^\top$. Thus it suffices to bound

$$\begin{aligned}
&\mathop{\mathbb{E}}_{\sigma \sim \mu_{H_{\mathsf{TAP}}}} (\sigma - \boldsymbol{m})(\sigma - \boldsymbol{m})^\top \\
&\quad = \mathop{\mathbb{E}}_{(a,b)\sim\nu} \mathop{\mathbb{E}}_{\sigma \sim \mu_{a,b}} (\sigma - v(a,b) + v(a,b) - \boldsymbol{m})(\sigma - v(a,b) + v(a,b) - \boldsymbol{m})^\top \\
&\quad \preceq 2 \mathop{\mathbb{E}}_{(a,b)\sim\nu} \mathop{\mathbb{E}}_{\sigma \sim \mu_{a,b}} (\sigma - v(a,b))(\sigma - v(a,b))^\top + 2 \mathop{\mathbb{E}}_{(a,b)\sim\nu} (v(a,b) - \boldsymbol{m})(v(a,b) - \boldsymbol{m})^\top. \qquad (7.10)
\end{aligned}$$

We will bound the spectral norm of each of the above terms below. We employ the following statement for proving the desired bounds, proved in Section 7.8.

Theorem 7.7.32. *Suppose $\varepsilon > 0$, and ξ is ε-strictly replica symmetric (Condition 7.7.27). Then, there exist $c = c(\varepsilon)$ and $C = C(\varepsilon)$ such that the following hold with probability $1 - e^{-cN^{1/5}}$.*

(a) We have $\nabla^2 H_N(0) \preceq (1 + \xi''(0) - \varepsilon^2/8)I_N$ and

$$\left|\log Z_N - \frac{N\xi(1)}{2} - \frac{N\xi''(0)}{4} - \frac{\log(1-\xi''(0))}{2} + \frac{1}{2}\log\det\left((1+\xi''(0))I_N - \nabla^2 H_N(0)\right)\right| \le 1.$$

(b) The Gibbs measure satisfies $\| \mathbb{E}_{\mu_{H_N}} \sigma\sigma^\top \|_{\mathsf{op}} \le C(1 + \gamma_1^2 N)$.

Bounding the first term. Observe that:

$$\mathop{\mathbb{E}}_{\sigma\sim\mu_{a,b}} (\sigma - v(a,b))(\sigma - v(a,b))^\top = r_{a,b}^2 Q_{a,b} \mathop{\mathbb{E}}_{\tau\sim\mu_{H_{a,b}}} \tau\tau^\top Q_{a,b}^\top .$$

Thus,

$$\left\|\mathop{\mathbb{E}}_{\sigma\sim\mu_{a,b}} (\sigma - v(a,b))(\sigma - v(a,b))^\top\right\| = r_{a,b}^2 \left\|\mathop{\mathbb{E}}_{\tau\sim\mu_{H_{a,b}}} \tau\tau^\top\right\| \le Cr_{a,b}^2(1 + a^2 + b^2)\,. \tag{7.11}$$

where the inequality follows from Lemma 7.7.31 and Theorem 7.7.32.

Our next goal is to control the fluctuations of (a,b). By definition, if $\widehat{Z}_{a,b}$ is the partition function of the distribution with Hamiltonian $H_{a,b}$ defined by $H_{a,b}(\tau) = H_{\mathsf{TAP}}(v(a,b) + r_{a,b}Q\tau)$, we have

$$\nu(a,b) \propto \exp\left(\log \widehat{Z}_{a,b} + (N-4)\log r_{a,b} + \mathop{\mathbb{E}}_{\mu_{\mathsf{TAP}}} H_{\mathsf{TAP}}(v(a,b)) + \sqrt{N} g_{a,b}\right). \tag{7.12}$$

We will show that $(a,b) \sim \nu$, conditioned on $|a|, |b| \le \varepsilon N^{1/10}$ for ε as in Lemma 7.7.31, is subgaussian with variance $O(1)$. We will also show in Lemma 7.7.39 that ν places very little mass outside the set $|a|, |b| \le \varepsilon N^{1/10}$.

Lemma 7.7.33. *Let ε be as in Lemma 7.7.31. On an event with probability $1 - e^{-cN^{1/5}}$, the following holds. The density of (a,b) under ν, conditioned on $|a|, |b| \le \varepsilon N^{1/10}$, is given by*

$$\nu(a,b) \propto \exp\left(N\widehat{E}_{a,b} + \sqrt{N} g_{a,b} + \mathsf{Error}_{a,b}^{(1)} + \mathsf{Error}_{a,b}^{(2)} + \Delta_{a,b}\right),$$

where $|\Delta_{a,b}| \le 1$ and

$$\begin{aligned}
\widehat{E}_{a,b} &= \frac{\xi_{a,b}(1)}{2} + \log r_{a,b} + \frac{1}{N}\mathop{\mathbb{E}}_{\mu_{\mathsf{TAP}}} H_{\mathsf{TAP}}(v(a,b)) - \frac{\xi_t(1)}{2} \\
&= \frac{1}{2}\left(\log r_{a,b}^2 - \xi_t(\|v(a,b)\|^2) - r_{a,b}^2 \cdot \frac{\xi_t'\left(q_{\boldsymbol{m}}\left(1 + \frac{a}{\sqrt{N}}\right)\right)^2}{\xi_t'(q_{\boldsymbol{m}})}\right. \\
&\quad - \gamma(q_{\boldsymbol{x}}) \cdot \frac{\xi_t'\left(q_{\boldsymbol{m}}\left(1 + \frac{a}{\sqrt{N}}\right)\right)}{\xi_t'(q_{\boldsymbol{m}})} \cdot \left(\left(1 + \frac{b}{\sqrt{N}}\right) - q_{\boldsymbol{m}} \cdot \frac{\xi_t''(q_{\boldsymbol{m}})}{\gamma'(q_{\boldsymbol{m}})} \cdot \left(1 + \frac{a}{\sqrt{N}}\right)\right) \\
&\quad \left. + \xi_t\left(q_{\boldsymbol{x}}\left(1 + \frac{b}{\sqrt{N}}\right)\right) + \frac{\gamma\left(q_{\boldsymbol{m}}\left(1 + \frac{a}{\sqrt{N}}\right)\right)}{\gamma'(q_{\boldsymbol{m}})} \cdot \left((1 - q_{\boldsymbol{m}})\xi_t''(q_{\boldsymbol{m}}) + \frac{1}{1 - q_{\boldsymbol{m}}}\right)\right).
\end{aligned}$$

for the error terms

$$\mathsf{Error}_{a,b}^{(1)} = \left(\log \widehat{Z}_{a,b} - \frac{N\xi_{a,b}(1)}{2} - \frac{N\xi_{a,b}''(0)}{4} + \frac{1}{2}\log\det\left((1 + \xi_{a,b}''(0))\mathrm{Id} - \nabla^2 H_N(v(a,b))\right)\right) - 4\log r_{a,b}$$

and

$$\mathsf{Error}_{a,b}^{(2)} = \frac{N\xi_{a,b}''(0)}{4} - \frac{1}{2}\log\det\left((1 + \xi_{a,b}''(0))\mathrm{Id} - \nabla^2 H_N(v(a,b))\right)$$

The above follows by expanding out all the terms in the expression (7.12) for the density of ν and evaluating the term $\log \widehat{Z}_{a,b}$ using Theorem 7.7.32, which applies because the models $\xi_{a,b}$ for $|a|,|b| \le \varepsilon N^{1/10}$ are ε-strictly replica symmetric by Lemma 7.7.31. Because the conclusion of Theorem 7.7.32 holds with probability $1-e^{-cN^{1/5}}$, we may evaluate $\log \hat{Z}_{a,b}$ over a $1/\mathsf{poly}(N)$-net of such (a,b) via a union bound, and then infer the estimate for all such a,b by a standard continuity argument.

Lemma 7.7.34. $\nabla \widehat{E}_{a,b}\Big|_{(a,b)=(0,0)} = 0.$

Lemma 7.7.35. *There exist constants $\eta,\varepsilon > 0$ such that for all $|a|,|b| \le \varepsilon\sqrt{N}$, $N\nabla^2 \widehat{E}_{a,b} \preceq -\eta \mathrm{Id}$.*

The above follow from routine calculations, which we defer to Appendix 7.B.

Lemma 7.7.36. *For every constant $\iota > 0$, there is a constant c such that with probability $1-e^{-cN}$, for all a,b, we have $|g_{a,b} - g_{0,0}| \le \iota \frac{a^2+b^2}{\sqrt{N}}$.*

Lemma 7.7.37. *With probability $1-e^{-cN^{1/5}}$, $|\mathsf{Error}^{(1)}_{a,b}| = O(1)$ uniformly for all $|a|,|b| < \varepsilon N^{1/10}$, for ε as in Lemma 7.7.31.*

Proof. We shall show this very high probability bound for a fixed a,b. Constructing a net over the relevant a,b and performing a union bound over this net allows us to extend this to a uniform bound for all a,b; we omit the details. We may write the error term as

$$\begin{aligned}\mathsf{Error}^{(1)}_{a,b} = \log \widehat{Z}_{a,b} - \frac{N\xi_{a,b}(1)}{2} - \frac{\log\left(1-\xi''_{a,b}(0)\right)}{2} \\ - \frac{N\xi''_{a,b}(0)}{4} + \frac{1}{2}\log\det\left((1+\xi''_{a,b}(0))\mathrm{Id} - \nabla^2 H_{a,b}(0)\right) + O(1).\end{aligned}$$

Due to the bound on a and b, the above is $O(1)$ with very high probability by Theorem 7.7.32(a). The desideratum follows. $\square$

Lemma 7.7.38. *For any sufficiently small $\iota > 0$, with probability at least $1-e^{-cN}$, $\left|\mathsf{Error}^{(2)}_{a,b} - \mathsf{Error}^{(2)}_{0,0}\right| = O(1)$ for all $a,b < \iota N^{1/4}$.*

We relegate the proof of the above to the appendix Appendix 7.B. The idea of the proof is that the Hessian $\nabla^2 H_{a,b}(0)$ does not deviate too much for small variations in a,b – the first order terms in the deviation end up being cancelled by the $\xi''_{a,b}(0)/2$ term, while the second order terms are $O(1)$.

Lemma 7.7.39. *Let $\mathcal{E}_\iota$ be as in Lemma 7.7.23. With probability $1-e^{-cN^{1/5}}$, either $\mathcal{E}_\iota$ does not hold, or the following holds. For ε as in Lemma 7.7.31,*

$$\Pr_{(a,b)\sim\nu}\left[|a| \le \varepsilon N^{1/10} \text{ and } |b| \le \varepsilon N^{1/10}\right] \ge 1-e^{-cN^{1/5}}.$$

Proof. On the event $\mathcal{E}_\iota$, we have

$$\Pr_{(a,b)\sim\nu}\left[|a| > \iota N^{1/2} \text{ or } |b| > \iota N^{1/2}\right] \le e^{-cN}.$$

Thus, let

$$T = \left\{(a,b) \in \mathbb{R}^2 : |a| \in [\varepsilon N^{1/10}, \iota N^{1/2}] \text{ or } |b| \in [\varepsilon N^{1/10}, \iota N^{1/2}]\right\}.$$

It suffices to show that $\Pr_{(a,b)\sim\nu}[(a,b)\in T] \le e^{-cN^{1/5}}$ with probability $1-e^{-cN^{1/5}}$. Recall the density of $(a,b)\sim\nu$ is given by (7.12), and that $\mathbb{E}\,\widehat{Z}_{a,b} = e^{N\xi_{a,b}(1)/2}$. Thus, for $\widehat{\mathbb{E}}$ denoting expectation with respect to the $\widehat{Z}_{a,b}$ alone,

$$\begin{aligned}&\widehat{\mathbb{E}}\int_T \exp\left(\log \widehat{Z}_{a,b} + (N-4)\log r_{a,b} + \mathop{\mathbb{E}}_{\mu_{\mathsf{TAP}}} H_{\mathsf{TAP}}(v(a,b)) + \sqrt{N} g_{a,b}\right)\mathsf{d}(a,b) \\ &= \int_T \exp\left(N\widehat{E}_{a,b} + \frac{N\xi_t(1)}{2} + \sqrt{N} g_{a,b} - 4\log r_{a,b}\right)\mathsf{d}(a,b).\end{aligned}$$

On the event in Lemma 7.7.36, we have, for any constant $\iota > 0$,

$$\sqrt{N} g_{a,b} \le \sqrt{N} g_{0,0} + \iota(a^2 + b^2).$$

By Lemma 7.7.35,

$$N\widehat{E}_{a,b} \le N\widehat{E}_{0,0} - \eta(a^2 + b^2).$$

Combining shows that

$$N\widehat{E}_{a,b} + \sqrt{N} g_{a,b} - 4\log r_{a,b} \le N\widehat{E}_{0,0} + \sqrt{N} g_{0,0} - \frac{\eta}{2}(a^2 + b^2) + O(1).$$

Combining shows

$$\begin{aligned}
&\widehat{\mathbb{E}} \int_T \exp\left(\log \widehat{Z}_{a,b} + (N-4)\log r_{a,b} + \mathop{\mathbb{E}}_{\mu_{\mathsf{TAP}}} H_{\mathsf{TAP}}(v(a,b)) + \sqrt{N} g_{a,b}\right) \mathsf{d}(a,b) \\
&\le e^{-cN^{1/5}} \exp\left(N\widehat{E}_{0,0} + \frac{N\xi_t(1)}{2} + \sqrt{N} g_{0,0}\right)
\end{aligned}$$

and therefore with probability $1 - e^{-cN^{1/5}/2}$ over the $\widehat{Z}_{a,b}$,

$$\begin{aligned}
&\int_T \exp\left(\log \widehat{Z}_{a,b} + (N-4)\log r_{a,b} + \mathop{\mathbb{E}}_{\mu_{\mathsf{TAP}}} H_{\mathsf{TAP}}(v(a,b)) + \sqrt{N} g_{a,b}\right) \mathsf{d}(a,b) \\
&\le e^{-cN^{1/5}/2} \exp\left(N\widehat{E}_{0,0} + \frac{N\xi_t(1)}{2} + \sqrt{N} g_{0,0}\right)
\end{aligned} \tag{7.13}$$

On the other hand, Lemma 7.7.33 implies that with probability $1 - e^{-cN^{1/5}}$,

$$\begin{aligned}
&\log \widehat{Z}_{0,0} + (N-4)\log r_{0,0} + \mathop{\mathbb{E}}_{\mu_{\mathsf{TAP}}} H_{\mathsf{TAP}}(v(0,0)) + \sqrt{N} g_{0,0} \\
&= N\widehat{E}_{0,0} + \frac{N\xi_t(1)}{2} + \sqrt{N} g_{0,0} + \mathsf{Error}^{(1)}_{0,0} + \mathsf{Error}^{(2)}_{0,0} + O(1),
\end{aligned}$$

and Lemma 7.7.37 implies $|\mathsf{Error}^{(1)}_{0,0}| = O(1)$ with probability $1 - e^{-cN^{1/5}}$. Furthermore, Lemma 7.8.3 below implies that $|\mathsf{Error}^{(2)}_{0,0}| \le N^{1/10}$ with probability $1 - e^{-cN^{1/5}}$. Thus

$$\begin{aligned}
&\log \widehat{Z}_{0,0} + (N-4)\log r_{0,0} + \mathop{\mathbb{E}}_{\mu_{\mathsf{TAP}}} H_{\mathsf{TAP}}(v(0,0)) + \sqrt{N} g_{0,0} \\
&\ge N\widehat{E}_{0,0} + \frac{N\xi_t(1)}{2} + \sqrt{N} g_{0,0} - 2N^{1/10},
\end{aligned}$$

and standard continuity arguments imply that for $T' = \{(a,b) : |a|, |b| \le N^{-10}\}$,

$$\begin{aligned}
&\int_{T'} \exp\left(\log \widehat{Z}_{a,b} + (N-4)\log r_{a,b} + \mathop{\mathbb{E}}_{\mu_{\mathsf{TAP}}} H_{\mathsf{TAP}}(v(a,b)) + \sqrt{N} g_{a,b}\right) \mathsf{d}(a,b) \\
&\ge e^{-3N^{1/10}} \exp\left(N\widehat{E}_{0,0} + \frac{N\xi_t(1)}{2} + \sqrt{N} g_{0,0}\right).
\end{aligned}$$

Comparing with (7.13) implies the conclusion, after adjusting c. □

Lemma 7.7.40. *With probability $1 - e^{-cN^{1/5}}$, either $\mathcal{E}_\iota$ does not hold or the following holds. There exists a random variable X over $\mathbb{R}^2$ (coupled with ν) such that the following holds for $(a,b) \sim \nu$.*

(a) With probability at least $1 - e^{-cN^{1/5}}$, $X = (a,b)$.

(b) X has mean $O(1)$ and is $O(1)$-subgaussian.

Proof. This is an immediate corollary of Lemmas 7.7.34, 7.7.35, 7.7.36, 7.7.37, 7.7.38, 7.7.39, setting X to be the random variable that is equal to (a,b) if $|a|, |b| < \varepsilon N^{1/10}$, and 0 otherwise. □

We are now finally prepared to bound $\mathsf{Cov}(\mu)$.

Lemma 7.7.25. *Let $\boldsymbol{x} \in S_N$, let $\mathcal{S}_\iota$ be as in Lemma 7.7.23, and let $\boldsymbol{m} \in \mathcal{S}_\iota$. Then for an absolute constant $K > 0$,*

$$\Pr_{H_{\mathsf{TAP}} \sim \mu_{\mathsf{TAP},\boldsymbol{x},\boldsymbol{m}}} [\|\mathsf{Cov}(\mu_{H_{\mathsf{TAP}}})\| \geq K \wedge \mathcal{E}_\iota] \leq e^{-cN^{1/5}}.$$

Proof. Note that $|a|, |b| \leq 2\sqrt{N}$ almost surely. Let X be as in Lemma 7.7.40. This lemma implies that with probability at least $1 - e^{-cN^{1/5}}$ over the randomness of the Hamiltonian,

$$\begin{aligned}\mathop{\mathbb{E}}_{(a,b)\sim\nu}\left[a^2 + b^2\right] &= \mathbb{E}[\|X\|^2] + \mathop{\mathbb{E}}_{(a,b)\sim\nu}\left[\mathbf{1}[X \neq (a,b)](a^2+b^2)\right] \\ &\leq O(1) + \Pr(X \neq (a,b)) \cdot 8N = O(1).\end{aligned}$$

Thus, by plugging in (7.11) along with this observation into (7.10), we get that the following holds with probability at least $1 - e^{-cN^{1/5}}$

$$\begin{aligned}\|\mathsf{Cov}(\mu)\| &\leq 2C \mathop{\mathbb{E}}_{(a,b)\sim\nu}(1 + a^2 + b^2) + \left\|2 \mathop{\mathbb{E}}_{(a,b)\sim\nu}(v(a,b) - \boldsymbol{m})(v(a,b) - \boldsymbol{m})^\top\right\| \\ &\leq O(1) + 2 \mathop{\mathbb{E}}_{(a,b)\sim\nu} \|v(a,b) - \boldsymbol{m}\|^2 \\ &\leq O(1) + 2 \mathop{\mathbb{E}}_{(a,b)\sim\nu}\left[O(a^2+b^2)\right] \\ &\leq O(1).\end{aligned}$$

□

7.8 High-probability covariance bound of replica symmetric spherical spin glass

In this section we prove the main technical input to the proofs in Section 7.7.4. This takes the form of a high-probability bound on the partition function and covariance matrix (in fact, second moment matrix) of a spherical spin glass in the replica symmetric phase.

In this section, we let H_N be defined as in (7.8), with a linear term corresponding to an external field:

$$H_N(\sigma) := \sum_{p\geq 1} \frac{\gamma_p}{N^{(p-1)/2}} \sum_{i_1,\ldots,i_p=1}^N \boldsymbol{g}_{i_1,\ldots,i_p} \sigma_{i_1} \cdots \sigma_{i_p}.$$

We recall $\xi_{\sim 1}(q) = \sum_{p\geq 2} \gamma_p^2 q^p$ denotes the part of ξ without the linear term, and let

$$\xi_{\sim 2}(q) = \gamma_1^2 q + \sum_{p\geq 3} \gamma_p^2 q^p$$

denote the part of ξ excluding the degree 2 term.

The results in this section hold under the following condition, which we restate for reference.

Condition 7.7.27 (ε-strict replica symmetry)**.** *We say ξ is ε-strictly replica symmetric if $\gamma_1^2 \leq N^{-4/5}$ and $\xi_{\sim 1}$ satisfies Condition 7.7.16.*

Throughout this section, we treat $\varepsilon > 0$ as a constant and let $O_\varepsilon(1)$ denote a quantity bounded depending on ε.

Theorem 7.7.32. *Suppose $\varepsilon > 0$, and ξ is ε-strictly replica symmetric (Condition 7.7.27). Then, there exist $c = c(\varepsilon)$ and $C = C(\varepsilon)$ such that the following hold with probability $1 - e^{-cN^{1/5}}$.*

(a) We have $\nabla^2 H_N(0) \preceq (1 + \xi''(0) - \varepsilon^2/8) I_N$ and

$$\left|\log Z_N - \frac{N\xi(1)}{2} - \frac{N\xi''(0)}{4} - \frac{\log(1 - \xi''(0))}{2} + \frac{1}{2}\log\det\left((1+\xi''(0))I_N - \nabla^2 H_N(0)\right)\right| \leq 1.$$

(b) *The Gibbs measure satisfies* $\| \mathbb{E}_{\mu_{H_N}} \sigma\sigma^\top \|_{\mathsf{op}} \le C(1+\gamma_1^2 N)$.

In the below proofs, we allow the constants c and C to change from line to line, but they will always be uniform in ε. We always set C sufficiently large depending on ε, and then c sufficiently small depending on ε, C.

Theorem 7.7.32 will be proved through the following pair of propositions. We introduce the degree-2 Hamiltonian

$$H_{N,2}(\sigma) \coloneqq \frac{\gamma_2}{N^{1/2}} \sum_{i_1,i_2=1}^{N} \boldsymbol{g}_{i_1,i_2}\sigma_{i_1}\sigma_{i_2} = \frac{1}{2}\langle \nabla^2 H_N(0)\sigma,\sigma\rangle. \tag{7.14}$$

Similarly let $H_{N,\sim 2}(\sigma) = H_N(\sigma) - H_{N,2}(\sigma)$ be the non degree-2 part of $H_N(\sigma)$. Define the degree-2 Gibbs measure and partition function by

$$\mathsf{d}\mu_{H_{N,2}}(\sigma) = \frac{\exp(H_{N,2}(\sigma))}{Z_{N,2}}\mathsf{d}\rho(\sigma), \qquad Z_{N,2} = \int_{S_N} \exp(H_{N,2}(\sigma))\mathsf{d}\rho(\sigma).$$

Throughout this section, we will let $\mathbb{E}$ denote expectation with respect to the disorder coefficients $\boldsymbol{g}_{i_1,\dots,i_p}$, while $\langle\cdot\rangle$ denotes averaging with respect to $\sigma\sim\mu_{H_N}$ (or several i.i.d. samples $\sigma^1,\sigma^2,\dots$ from this measure). Similarly, let $\langle\cdot\rangle_2$ denote Gibbs average with respect to $\mu_{H_{N,2}}$.

Note that $\nabla^2 H_N(0)$ depends on H_N only through $H_{N,2}$.

Proposition 7.8.1 (Concentration of degree-2 partition function; proved in Subsection 7.8.1)**.** *With probability* $1-e^{-cN}$ *over* $H_{N,2}$, *we have* $\nabla^2 H_N(0) \preceq (1+\xi''(0)-\varepsilon^2/8)I_N$ *and*

$$\left|\log Z_{N,2} - \frac{N\xi''(0)}{2} - \frac{\log(1-\xi''(0))}{2} + \frac{1}{2}\log\det\big((1+\xi''(0))I - \nabla^2 H_N(0)\big)\right| \le 1/2.$$

The following is proved in Subsections 7.8.2 and 7.8.3.

Proposition 7.8.2. *There is a* $H_{N,2}$*-measurable event with probability* $1-e^{-cN^{1/5}}$ *on which the following holds with probability* $1-e^{-cN^{1/5}}$ *over* $H_{N,\sim 2}$.

1. *The partition functions* Z_N, $Z_{N,2}$ *satisfy*

$$\left|\log\frac{Z_N}{Z_{N,2}} - \frac{N\xi_{\sim 2}(1)}{2}\right| \le 1/2.$$

2. *The Gibbs measure satisfies* $\|\langle\sigma\sigma^\top\rangle\|_{\mathsf{op}} \le C(1+\gamma_1^2 N)$.

Proof of Theorem 7.7.32. Immediate from Propositions 7.8.1 and 7.8.2, since $\xi(1) = \xi_{\sim 2}(1) + \frac{1}{2}\xi''(0)$. $\square$

We also show the following concentration of the log determinant in Theorem 7.7.32.

Lemma 7.8.3 (Proved in Subsection 7.8.1)**.** *There exists a* $H_{N,2}$*-measurable random variable* X *that the following holds.*

1. *With probability* $1-e^{-cN}$, $X = \log\det((1+\xi''(0))I - \nabla^2 H_N(0)) - N\xi''(0)/2$.
2. X *has mean* $O_\varepsilon(1)$ *and is* $O_\varepsilon(1)$*-subgaussian.*

This implies the quantitative contiguity between the planted and null models, which we restate below for convenience.

Proposition 7.7.17 (Quantitative contiguity)**.** *Under Condition 7.7.16, there exists* $c=c(\varepsilon)>0$ *such that for any event* $\mathcal{E}$, *if* $\mu_{\mathsf{pl}}(\mathcal{E}) = p$, *then* $\mu_{\mathsf{null}}(\mathcal{E}) \le e^{-cN^{1/5}} + e^{\frac{1}{c}\sqrt{\log\frac{2}{p}}}p$.

Proof. Let $\mathcal{E}_{\mathsf{good}}$ be intersection of the event in Theorem 7.7.32, the event

$$X = \log\det((1+\xi''(0))I - \nabla^2 H_N(0)) - N\xi''(0)/2$$

from Lemma 7.8.3, and the event $X \le t$, for some $t > 0$ to be determined. Then, after adjusting $c = c(\varepsilon)$ as necessary,

$$\mu_{\mathsf{null}}(\overline{\mathcal{E}_{\mathsf{good}}}) \le e^{-cN^{1/5}} + \mathbb{P}(X > t) \le e^{-cN^{1/5}} + e^{-c(t-\frac{1}{c})_+^2}.$$

Note that $\log \mathbb{E}\, Z_N = N\xi(1)/2$, while on the event $\mathcal{E}_{\mathsf{good}}$,

$$\log Z_N = \frac{N\xi(1) - X + O_\varepsilon(1)}{2} \ge \frac{N\xi(1) - t - \frac{1}{c}}{2}.$$

Thus $\frac{\mathbb{E}\, Z_N}{Z_N} \le e^{\frac{1}{2}(t+\frac{1}{c})}$. So,

$$\begin{aligned}
\mu_{\mathsf{null}}(\mathcal{E}) &\le \mu_{\mathsf{null}}(\overline{\mathcal{E}_{\mathsf{good}}}) + \int \frac{\mathbb{E}\, Z_N}{Z_N} \mathbf{1}[H_N \in \mathcal{E} \cap \mathcal{E}_{\mathsf{good}}] \mathsf{d}\mu_{\mathsf{pl}}(H_N) \\
&\le e^{-cN^{1/5}} + e^{-c(t-\frac{1}{c})_+^2} + e^{\frac{1}{2}(t+\frac{1}{c})} p.
\end{aligned}$$

We then take $t = \frac{1}{c} + \sqrt{\frac{1}{c}\log\frac{1}{p}}$, so that this is bounded by

$$e^{-cN^{1/5}} + \left(1 + e^{\frac{1}{c} + \sqrt{\frac{1}{c}\log\frac{1}{p}}}\right) p\,.$$

Further adjusting c proves the desired bound. □

7.8.1 Concentration of degree-2 partition function

We write $A = \frac{1}{2}\nabla^2 H_N(0) = \frac{\sqrt{\xi''(0)}}{2} M$. It is straightforward to check that M is distributed as a sample from $\mathsf{GOE}(N)$.

Fact 7.8.4. *We have $\xi''(0) \le 1 - \varepsilon$.*

Proof. Writing (7.7) as

$$\xi_{\sim 1}(q) + q + \log(1-q) \le -\varepsilon q^2/2$$

and Taylor expanding around $q = 0$ implies the result. □

Fact 7.8.5. *With probability $1 - e^{-cN}$, $\nabla^2 H_N(0) \preceq (1 + \xi''(0) - \varepsilon^2/8)I$.*

Proof. With probability $1 - e^{-cN}$, we have $\lambda_{\max}(M) \le 2 + \varepsilon^2/8$. Then,

$$\begin{aligned}
\lambda_{\max}\left((1+\xi''(0) - \varepsilon^2/8)I - \nabla^2 H_N(0)\right) &\ge 1 + \xi''(0) - \varepsilon^2/8 - \sqrt{\xi''(0)}(2 + \varepsilon^2/8) \\
&= (1 - \sqrt{\xi''(0)})^2 - \varepsilon^2(1 + \sqrt{\xi''(0)})/8 \\
&> \varepsilon^2/4 - \varepsilon^2/4 = 0
\end{aligned}$$

by Fact 7.8.4. □

For $\gamma \in (\lambda_{\max}(A), +\infty)$, define

$$G(\gamma) = \gamma - \frac{1}{2N}\log\det(\gamma I - A). \tag{7.15}$$

Note that

$$G'(\gamma) = 1 - \frac{1}{2N}\mathsf{Tr}(\gamma I - A)^{-1}$$

is continuous and increasing, with $\lim_{\gamma\downarrow\lambda_{\max}(A)} G'(\gamma) = -\infty$ and $\lim_{\gamma\uparrow+\infty} G'(\gamma) = 1$. Thus G' has a unique root γ_* in $(\lambda_{\max}(A), +\infty)$. The following lemma is a consequence of [HMP24, Lemma 7.3], which is proved by an analysis of a Laplace transform of the free energy also used in [BL16].

Lemma 7.8.6. *With probability* $1-e^{-cN}$ *over* $H_{N,2}$,

$$Z_{N,2} = (1+O(N^{-c}))\sqrt{\frac{2}{G''(\gamma_*)}}(2e)^{-N/2}\exp(NG(\gamma_*)). \tag{7.16}$$

Proof. Recalling (7.14), we have

$$H_{N,2}(\sigma) = \frac{\sqrt{\xi''(0)}}{2}\langle M\sigma,\sigma\rangle,$$

and Fact 7.8.4 implies the factor $\sqrt{\xi''(0)}/2$ is bounded away from $1/2$. Then [HMP24, Lemma 7.3] (with $u=0$) implies the result. $\square$

Define $\gamma_0 = (1+\xi''(0))/2$. The next lemma shows that, although the variable γ_* in (7.16) is random, we may approximate it deterministically by γ_0.

Lemma 7.8.7. *For sufficiently large* C *depending on* ε*, and sufficiently small* c *depending on* ε, C*, with probability* $1-e^{-cN}$ *the following holds for all* $\gamma \in [\gamma_0 - N^{-1/2}, \gamma_0 + N^{-1/2}]$.

1. $|G'(\gamma_0)| \le 1/(C\sqrt{N})$.
2. $|\frac{G''(\gamma)}{2/(1-\xi''(0))} - 1| \le N^{-1/3}$.

Proof. Let $\mathsf{d}\rho_{\mathsf{sc}}(x) = \frac{1}{2\pi}\mathbf{1}[|x|\le 2]\sqrt{4-x^2}\ \mathsf{d}x$ denote Wigner's semicircle law, and

$$f_1(x) = 1 - \frac{1}{1+\xi''(0)-2\sqrt{\xi''(0)}x}, \qquad f_2(x) = \frac{2}{(1+\xi''(0)-2\sqrt{\xi''(0)}x)^2}.$$

For $k\in[2]$, let

$$L_k = \int f_k(x)\ \mathsf{d}\rho_{\mathsf{sc}}(x).$$

We will show that with probability $1-e^{-cN}$, for each $k\in[2]$,

$$\left|G^{(k)}(\gamma_0) - L_k\right| \le \frac{1}{C\sqrt{N}}. \tag{7.17}$$

Recall that $M\sim \mathsf{GOE}(N)$. For $f:\mathbb{R}\to\mathbb{R}$, define the spectral trace

$$\mathsf{Tr} f(M) = \sum_{i=1}^N f(\lambda_i(M)).$$

Note that $G^{(k)}(\gamma_0) = N^{-1}\cdot \mathsf{Tr} f_k(M)$. Define

$$\widetilde{f}_k(x) = f_k(\min(x, 2+\varepsilon^2/8)).$$

By the proof of Fact 7.8.5, $1+\xi''(0)-2\sqrt{\xi''(0)}x \ge \varepsilon^2/8$ for $x\le 2+\varepsilon^2/8$, so $\widetilde{f}_k$ is $O_\varepsilon(1)$-Lipschitz. Moreover, $\lambda_{\max}(M)\le 2+\varepsilon^2/8$ with probability $1-e^{-cN}$, and on this event $\mathsf{Tr} f_k(M) = \mathsf{Tr}\widetilde{f}_k(M)$.

By [GZ00, Lemma 1.2(b)], if we write $M_{i,i} = \sqrt{2/N}Z_{i,i}$, $M_{i,j} = \sqrt{1/N}Z_{i,j}$, then $\mathsf{Tr}\widetilde{f}_k(M)$ is a $O_\varepsilon(1)$-Lipschitz function of the standard gaussians $(Z_{i,j})_{1\le i\le j\le N}$. Thus $\mathsf{Tr}\widetilde{f}_k(M)$ is $O_\varepsilon(1)$-subgaussian, i.e.

$$\mathbb{P}(|\mathsf{Tr}\widetilde{f}_k(M) - \mathbb{E}\,\mathsf{Tr}\widetilde{f}_k(M)| \ge t) \le 2e^{-t^2/C}$$

for some $C = O_\varepsilon(1)$. By [BY05, Theorem 1.1],

$$\mathsf{Tr}\widetilde{f}_k(M) - NL_k$$

converges in distribution to a gaussian with mean and variance $O_\varepsilon(1)$. Combined with subgaussianity of $\mathsf{Tr}\widetilde{f}_k(M)$, this implies

$$|\mathbb{E}\,\mathsf{Tr}\widetilde{f}_k(M) - NL_k| = O_\varepsilon(1).$$

It follows that (after possibly increasing $C = O_\varepsilon(1)$),

$$\mathbb{P}(|\mathsf{Tr}\widetilde{f}_k(M) - NL_k| \geq t) \leq 2e^{-(t-C)_+^2/C}.$$

Thus

$$\begin{aligned}\mathbb{P}(|G^{(k)} - L_k| \geq t) &\leq \mathbb{P}(\mathsf{Tr}f_k(M) \neq \mathsf{Tr}\widetilde{f}_k(M)) + \mathbb{P}(|\mathsf{Tr}\widetilde{f}_k(M) - NL_k| \geq Nt)\\ &\leq e^{-cN} + 2e^{-(Nt-C)_+^2/C}.\end{aligned}$$

Plugging in $t = 1/(C\sqrt{N})$ proves (7.17). Next, direct calculations show $L_1 = 0$, $L_2 = \frac{2}{1-\xi''(0)}$. The former directly implies conclusion (1), and the latter implies

$$\left|G''(\gamma_0) - \frac{2}{1-\xi''(0)}\right| \leq \frac{1}{C\sqrt{N}}.$$

Moreover, on the probability $1 - e^{-cN}$ event that $\lambda_{\max}(M) \leq 2 + \varepsilon^2/8$, $G^{(3)}(\gamma) = O_\varepsilon(1)$ for all $\gamma \in [\gamma_0 - N^{-1/2}, \gamma_0 + N^{-1/2}]$. This implies the conclusion (2). □

Lemma 7.8.3 is proved by the same method, and we present the proof here.

Proof of Lemma 7.8.3. Let

$$f_0(x) = \log(1 + \xi''(0) - \sqrt{\xi''(0)}x).$$

An elementary calculation shows that

$$L_0 := \int f_0(x)\,\mathsf{d}\rho_{\mathsf{sc}}(X) = \xi''(0)/2.$$

Proceeding as in the above proof, we have

$$\log\det\Big((1+\xi''(0))I - \sqrt{\xi''(0)}\nabla^2 H_N(0)\Big) = \mathsf{Tr}f_0(M).$$

If we take $\widetilde{f}_0(x) = f_0(\min(x, 2+\varepsilon^2/8))$, then $\mathsf{Tr}f_0(M) = \mathsf{Tr}\widetilde{f}_0(M)$ with probability $1 - e^{-cN}$. The same proof shows $\mathsf{Tr}\widetilde{f}_0(M)$ is $O_\varepsilon(1)$-subgaussian, and

$$|\mathbb{E}\,\mathsf{Tr}\widetilde{f}_0(M) - NL_0| = O_\varepsilon(1).$$

Thus we may take $X = \mathsf{Tr}\widetilde{f}_0(M) - NL_0 = \mathsf{Tr}\widetilde{f}_0(M) - N\xi''(0)/2$. □

Proof of Proposition 7.8.1. The assertion $\nabla^2 H_N(0) \preceq (1 + \xi''(0) - \varepsilon^2/8)I_N$ is proved in Fact 7.8.5. Suppose the events in Lemma 7.8.6 and 7.8.7 occur. Since γ_* is the solution to $G'(\gamma_*) = 0$, we have

$$|\gamma_0 - \gamma_*| \leq |G'(\gamma_0)| \cdot \frac{1-\xi''(0)}{2(1-N^{-1/3})} \leq \frac{1}{C\sqrt{N}}.$$

So,

$$N|G(\gamma_0) - G(\gamma_*)| \leq \frac{N}{2}|\gamma_0 - \gamma_*|^2 \sup_{\gamma\in[\gamma_0 - N^{-1/2}, \gamma_0 + N^{-1/2}]} G''(\gamma) \leq \frac{1}{C^2}\cdot\frac{2(1+N^{-1/3})}{1-\xi''(0)} \leq \frac{3}{C^2\varepsilon}.$$

Moreover, $\xi''(\gamma_0)/\xi''(\gamma_*) = 1 + O(N^{-1/3})$. Combining with Lemma 7.8.6 shows that, for some Δ satisfying $|\Delta| \leq \frac{3}{C^2\varepsilon}$,

$$\begin{aligned}Z_{N,2} &= (1+O(N^{-c}))e^{\Delta}\sqrt{\frac{2}{G''(\gamma_0)}}(2e)^{-N/2}\exp(NG(\gamma_0))\\ &= (1+O(N^{-c}))e^{\Delta}\sqrt{1-\xi''(0)}(2e)^{-N/2}\exp(N\gamma_0)\det\left(\gamma_0 I - \frac{1}{2}\nabla^2 H_N(0)\right)^{-1/2}\\ &= (1+O(N^{-c}))e^{\Delta}\sqrt{1-\xi''(0)}\exp(N\xi''(0)/2)\det\big((1+\xi''(0))I - \nabla^2 H_N(0)\big)^{-1/2}.\end{aligned}$$

Taking a logarithm and setting C sufficiently large concludes the proof. □

Finally, the following concentration estimates for samples from $\mu_{H_{N,2}}$ will be useful in the sequel. This is proved similarly to [HMP24, Lemma 7.5], and we defer the proof to Subsection 7.B.3.

Let $v_1, \dots, v_N$ denote the (unit) eigenvectors of $\nabla^2 H_N(0)$. These are well defined on the almost sure event that all eigenvalues of $\nabla^2 H_N(0)$ have multiplicity 1.

Proposition 7.8.8. *With probability $1 - e^{-cN}$ over $H_{N,2}$, the following holds. Let $\sigma, \sigma^1, \sigma^2 \sim \mu_{H_{N,2}}$, and let $W = \langle \sigma, v_i \rangle$ for any $i \in [N]$, or $W = \langle \sigma^1, \sigma^2 \rangle / \sqrt{N}$. Then:*

1. *For any $0 \le t \le N^{1/5}$, $\mathbb{P}(|W| \ge t) \le 3e^{-ct^2}$.*
2. *For any $k \in 2\mathbb{N}$, there exists $C_k > 0$ independent of N such that $\langle W^k \rangle_2 \le C_k$.*

In particular, part (2) implies $\|\langle \sigma\sigma^\top \rangle_2\|_{\mathsf{op}} \le C$.

7.8.2 Conditional positive probability bounds for non degree-2 part

In this subsection, we prove the following propositions, which establish a weaker version of Proposition 7.8.2 with positive instead of high probability.

Proposition 7.8.9. *There is a $H_{N,2}$-measurable event with probability $1 - e^{-cN^{1/5}}$ on which, with probability $1 - N^{-1/15}$ over $H_{N,\sim 2}$,*

$$\left| \log \frac{Z_N}{Z_{N,2}} - \frac{N\xi_{\sim 2}(1)}{2} \right| = O(N^{-1/15}).$$

Proposition 7.8.10. *There is a $H_{N,2}$-measurable event with probability $1 - e^{-cN^{1/5}}$ on which, with probability $1/2$ over $H_{N,\sim 2}$,*

$$\left\| \frac{Z_N}{Z_{N,2} e^{N\xi_{\sim 2}(1)/2}} \langle \sigma\sigma^\top \rangle - \langle \sigma\sigma^\top \rangle_2 \right\|_F^2 \le C(1 + \gamma_1^4 N^2).$$

In conjunction with Propositions 7.8.8 and 7.8.9, the above immediately implies a positive probability bound on the second moment matrix $\langle \sigma\sigma^\top \rangle$.

Both propositions rely on the following truncation to Z_N developed in [HS23b], which allows one to estimate Z_N via the second moment method throughout the strictly RS regime. As shown in the following lemma, this truncation does not significantly affect the first moment; at the same time, it will force the second moment to be dominated by pairs of nearly-orthogonal points.

Lemma 7.8.11. *The following holds for sufficiently small $c > 0$ depending on ε. Let*

$$T = T(H_N) := \left\{ \sigma \in S_N : \int_{S_N} \mathbf{1}[|R(\sigma, \tau)| \ge N^{-2/5}] e^{H_N(\tau)} \mathsf{d}\rho(\tau) \le e^{N\xi(1)/2 - cN^{1/5}} \right\}.$$

Then, we have:

$$\mathbb{E} \int_{S_N} \mathbf{1}[\sigma \notin T] e^{H_N(\sigma)} \mathsf{d}\rho(\sigma) \le e^{N\xi(1)/2 - cN^{1/5}}, \tag{7.18}$$

$$\mathbb{E} \int_{S_N} \mathbf{1}[\sigma \notin T] e^{H_{N,2}(\sigma)} \mathsf{d}\rho(\sigma) \le e^{N\xi''(0)/4 - cN^{1/5}}, \tag{7.19}$$

$$\mathbb{E} \int_{S_N} \mathbf{1}[\sigma^1 \notin T, |R(\sigma^1, \sigma^2)| \le 3N^{-2/5}] e^{H_N(\sigma^1) + H_N(\sigma^2)} \mathsf{d}\rho^{\otimes 2}(\sigma^1, \sigma^2) \le e^{N\xi(1) - cN^{1/5}}, \tag{7.20}$$

$$\mathbb{E} \int_{S_N} \mathbf{1}[\sigma^1 \notin T, |R(\sigma^1, \sigma^2)| \le 3N^{-2/5}] e^{H_{N,2}(\sigma^1) + H_N(\sigma^2)} \mathsf{d}\rho^{\otimes 2}(\sigma^1, \sigma^2) \le e^{N\xi(1)/2 + N\xi''(0)/4 - cN^{1/5}}. \tag{7.21}$$

The proof of the above lemma is very similar to [HS23b, Proposition 3.1] and [HMP24, Lemma 7.9], and we defer it to Subsection 7.B.3. As a corollary, we can get control on the first two moments of Z_N with respect to the randomness in $H_{N,\sim 2}$. To this end, let $\mathbb{E}_{\sim 2}$ denote expectation with respect to $H_{N,\sim 2}$.

Corollary 7.8.12. *There is a $H_{N,2}$-measurable event with probability $1-e^{-cN^{1/5}}$ on which the following holds. For*

$$\widetilde{T}=\widetilde{T}(H_{N,\sim 2}):=\left\{\sigma\in S_N:\langle \mathbf{1}[|R(\sigma,\tau)|\geq N^{-2/5}]e^{H_{N,\sim 2}(\tau)}\rangle_2\leq e^{N\xi_{\sim 2}(1)/2-cN^{1/5}}\right\}, \tag{7.22}$$

where the Gibbs average is with respect to $\tau\sim\langle\cdot\rangle_2$, we have

$$\begin{aligned}
\mathop{\mathbb{E}}_{\sim 2}\left\langle \mathbf{1}[\sigma\notin\widetilde{T}]e^{H_{N,\sim 2}(\sigma)}\right\rangle_2 &\leq e^{N\xi_{\sim 2}(1)/2-cN^{1/5}},\\
\mathop{\mathbb{E}}_{\sim 2}\left\langle \mathbf{1}[\sigma\notin\widetilde{T}]\right\rangle_2 &\leq e^{-cN^{1/5}},\\
\mathop{\mathbb{E}}_{\sim 2}\left\langle \mathbf{1}[\sigma^1\notin\widetilde{T},|R(\sigma^1,\sigma^2)|\leq 3N^{-2/5}]e^{H_{N,\sim 2}(\sigma^1)+H_{N,\sim 2}(\sigma^2)}\right\rangle_2 &\leq e^{N\xi_{\sim 2}(1)-cN^{1/5}},\\
\mathop{\mathbb{E}}_{\sim 2}\left\langle \mathbf{1}[\sigma^1\notin\widetilde{T},|R(\sigma^1,\sigma^2)|\leq 3N^{-2/5}]e^{H_{N,\sim 2}(\sigma^2)}\right\rangle_2 &\leq e^{N\xi_{\sim 2}(1)/2-cN^{1/5}}.
\end{aligned}$$

Proof. By Proposition 7.8.1 and Lemma 7.8.3, with probability $1-e^{-cN^{2/5}}$ over $H_{N,2}$,

$$Z_{N,2}\geq e^{N\xi''(0)/4-cN^{1/5}/2}.$$

On this event, for $\sigma\in T$ where T is as in Lemma 7.8.11,

$$\begin{aligned}
\left\langle \mathbf{1}[|R(\sigma,\tau)|\geq N^{-2/5}]e^{H_{N,\sim 2}(\tau)}\right\rangle_2 &= \frac{1}{Z_{N,2}}\int_{S_N}\mathbf{1}[|R(\sigma,\tau)|\geq N^{-2/5}]e^{H_N(\tau)}\mathsf{d}\rho(\tau)\\
&\leq e^{N\xi_{\sim 2}(1)/2-cN^{1/5}/2}.
\end{aligned}$$

Here we recall $\xi(1)/2-\xi''(0)/4=\xi_{\sim 2}(1)/2$. So, $\sigma\in\widetilde{T}(H_{N,\sim 2},c/2)$, where this denotes $\widetilde{T}$ defined with $c/2$ in place of c. Therefore $T\subseteq\widetilde{T}(H_{N,\sim 2},c/2)$.

By Markov's inequality and Lemma 7.8.11, with probability $1-e^{-cN^{1/5}/4}$ over $H_{N,2}$,

$$\mathop{\mathbb{E}}_{\sim 2}\int_{S_N}\mathbf{1}[\sigma\notin T]e^{H_N(\sigma)}\mathsf{d}\rho(\sigma)\leq e^{N\xi(1)/2-3cN^{1/5}/4}$$

On the intersection of these events,

$$\begin{aligned}
\mathop{\mathbb{E}}_{\sim 2}\left\langle \mathbf{1}[\sigma\notin\widetilde{T}(H_{N,\sim 2},c/2)]e^{H_{N,\sim 2}(\sigma)}\right\rangle_2 &\leq \mathop{\mathbb{E}}_{\sim 2}\left\langle \mathbf{1}[\sigma\notin T]e^{H_{N,\sim 2}(\sigma)}\right\rangle_2\\
&=\frac{1}{Z_{N,2}}\mathop{\mathbb{E}}_{\sim 2}\int_{S_N}\mathbf{1}[\sigma\notin T]e^{H_N(\sigma)}\mathsf{d}\rho(\sigma)\\
&\leq e^{N\xi_{\sim 2}(1)/2-cN^{1/5}/4}.
\end{aligned}$$

The first conclusion follows by adjusting c, and the other two conclusions follow similarly. □

For the rest of this subsection, we condition on a realization of $H_{N,2}$ satisfying the following good event.

Definition 7.8.13. Let E_2 denote the $H_{N,2}$-measurable event that the events in Proposition 7.8.8 and Corollary 7.8.12 hold. This occurs with probability $1-e^{-cN^{1/5}}$.

We can now prove Proposition 7.8.9.

Proof of Proposition 7.8.9. Let $\widetilde{T}$ be as in Corollary 7.8.12. We can write

$$\frac{Z_N}{Z_{N,2}}=\langle e^{H_{N,\sim 2}(\sigma)}\rangle_2=X_1+X_2, \tag{7.23}$$

where

$$X_1=\left\langle \mathbf{1}[\sigma\in\widetilde{T}]e^{H_{N,\sim 2}(\sigma)}\right\rangle_2, \qquad X_2=\left\langle \mathbf{1}[\sigma\notin\widetilde{T}]e^{H_{N,\sim 2}(\sigma)}\right\rangle_2.$$

We will show that that X_2 is much smaller than $\mathbb{E}_{\sim 2} X_1$ with high probability, and then control the fluctuations of X_1. For all $\sigma \in S_N$, $\mathbb{E}_{\sim 2}[e^{H_{N,\sim 2}(\sigma)}] = e^{N\xi_{\sim 2}(1)/2}$, so Corollary 7.8.12 implies

$$(1 - e^{-cN^{1/5}})e^{N\xi_{\sim 2}(1)/2} \le \mathop{\mathbb{E}}_{\sim 2}[X_1] \le e^{N\xi_{\sim 2}(1)/2}. \tag{7.24}$$

On the other hand, by Corollary 7.8.12 and Markov's inequality, with probability $1 - e^{-cN^{1/5}/2}$ over $H_{N,\sim 2}$,

$$X_2 \le e^{N\xi_{\sim 2}(1)/2 - cN^{1/5}/2}, \tag{7.25}$$

so $X_2 \le e^{-cN^{1/5}/2}\,\mathbb{E}_{\sim 2} X_1$, as desired.

We now control the fluctuations of X_1 by estimating

$$\mathsf{Var}_{\sim 2}[X_1] \coloneqq \mathop{\mathbb{E}}_{\sim 2}[X_1^2] - \mathop{\mathbb{E}}_{\sim 2}[X_1]^2.$$

Then, for $\sigma^1, \sigma^2 \sim \mu_{H_{N,2}}$,

$$\mathop{\mathbb{E}}_{\sim 2}[X_1^2] = \mathop{\mathbb{E}}_{\sim 2}\left\langle \mathbf{1}[\sigma^1, \sigma^2 \in \widetilde{T}]e^{H_{N,\sim 2}(\sigma^1)+H_{N,\sim 2}(\sigma^2)}\right\rangle_2 \le \mathop{\mathbb{E}}_{\sim 2}[Y_1] + \mathop{\mathbb{E}}_{\sim 2}[Y_2],$$

where

$$\begin{aligned} Y_1 &= \left\langle \mathbf{1}[|R(\sigma^1,\sigma^2)| \le N^{-2/5}]e^{H_{N,\sim 2}(\sigma^1)+H_{N,\sim 2}(\sigma^2)}\right\rangle_2, \\ Y_2 &= \left\langle \mathbf{1}[\sigma^1 \in \widetilde{T}, |R(\sigma^1,\sigma^2)| \ge N^{-2/5}]e^{H_{N,\sim 2}(\sigma^1)+H_{N,\sim 2}(\sigma^2)}\right\rangle_2. \end{aligned} \tag{7.26}$$

By the definition of $\widetilde{T}$ and (7.24),

$$\mathop{\mathbb{E}}_{\sim 2}[Y_2] \le \mathop{\mathbb{E}}_{\sim 2}\left\langle \mathbf{1}[\sigma^1 \in \widetilde{T}]e^{H_{N,\sim 2}(\sigma^1)}\right\rangle_2 e^{N\xi_{\sim 2}(1)/2 - cN^{1/5}} \le e^{N\xi_{\sim 2}(1) - cN^{1/5}}. \tag{7.27}$$

We further calculate

$$\mathop{\mathbb{E}}_{\sim 2}[Y_1] = e^{N\xi_{\sim 2}(1)}\left\langle \mathbf{1}[|R(\sigma^1,\sigma^2)| \le N^{-2/5}]e^{N\xi_{\sim 2}(R(\sigma^1,\sigma^2))}\right\rangle_2.$$

Recall that in Theorem 7.7.32, we assumed $\gamma_1^2 \le N^{-4/5}$. Thus, for $|R| \le N^{-2/5}$,

$$\xi_{\sim 2}(R) = \gamma_1^2 R + O(R^3) = O(N^{-6/5}).$$

It follows that $\mathbb{E}_{\sim 2}[Y_1] \le (1 + O(N^{-1/5}))e^{N\xi_{\sim 2}(1)}$.

Combining the above estimates shows $\mathbb{E}_{\sim 2}[X_1^2] \le (1 + O(N^{-1/5}))e^{N\xi_{\sim 2}(1)}$. Further combining with the lower bound in (7.24) shows

$$\mathsf{Var}_{\sim 2}[X_1] = O(N^{-1/5})e^{N\xi_{\sim 2}(1)}.$$

By Chebyshev's inequality, with probability $1 - N^{-1/15}/2$,

$$|X_1 - \mathop{\mathbb{E}}_{\sim 2}[X_1]| = O(N^{-1/15})e^{N\xi_{\sim 2}(1)/2}$$

Union bounding with the event in (7.25), and recalling (7.24), we conclude that with probability $1 - N^{-1/15}$,

$$\frac{Z_N}{Z_{N,2}} = (1 + O(N^{-1/15}))e^{N\xi_{\sim 2}(1)/2}. \qquad \square$$

The above proof also implies the following estimate, which will be useful in the sequel.

Corollary 7.8.14. *On the event E_2 (Definition 7.8.13), with probability $1 - e^{-cN^{1/5}}$ over $H_{N,\sim 2}$,*

$$\int \mathbf{1}[|R(\sigma^1,\sigma^2)| \ge N^{-2/5}\}e^{H_N(\sigma^1)+H_N(\sigma^2)}\mathsf{d}\rho^{\otimes 2}(\sigma^1,\sigma^2) \le Z_{N,2}^2 e^{N\xi_{\sim 2}(1) - cN^{1/5}}.$$

Proof. Dividing through by $Z_{N,2}^2$, it suffices to show, for $\sigma^1, \sigma^2 \sim \mu_{H_{N,2}}$,

$$\left\langle \mathbf{1}[|R(\sigma^1,\sigma^2)| \geq N^{-2/5}] e^{H_{N,\sim 2}(\sigma^1)+H_{N,\sim 2}(\sigma^2)} \right\rangle_2 \leq e^{N\xi_{\sim 2}(1)-cN^{1/5}}.$$

The left-hand side is bounded by $Y_2 + Y_3$, where Y_2 is as in (7.26) and

$$Y_3 := \left\langle \mathbf{1}[\sigma^1 \notin \widetilde{T}] e^{H_{N,\sim 2}(\sigma^1)+H_{N,\sim 2}(\sigma^2)} \right\rangle_2 = \left\langle \mathbf{1}[\sigma \notin \widetilde{T}] e^{H_{N,\sim 2}(\sigma)} \right\rangle_2 \left\langle e^{H_{N,\sim 2}(\sigma)} \right\rangle_2.$$

By (7.27), $\mathbb{E}_{\sim 2}[Y_2] \leq e^{N\xi_{\sim 2}(1)-cN^{1/5}}$. By Corollary 7.8.12,

$$\mathop{\mathbb{E}}_{\sim 2} \left\langle \mathbf{1}[\sigma \notin \widetilde{T}] e^{H_{N,\sim 2}(\sigma)} \right\rangle_2 \leq e^{N\xi_{\sim 2}(1)/2-cN^{1/5}}, \qquad \mathop{\mathbb{E}}_{\sim 2} \langle e^{H_{N,\sim 2}(\sigma)} \rangle_2 \leq e^{N\xi_{\sim 2}(1)/2}.$$

So, the following estimates each hold with probability $1 - e^{-cN^{1/5}/4}$ over $H_{N,\sim 2}$:

$$Y_2 \leq e^{N\xi_{\sim 2}(1)-cN^{1/5}/2}, \qquad \left\langle \mathbf{1}[\sigma \notin \widetilde{T}] e^{H_{N,\sim 2}(\sigma)} \right\rangle_2 \leq e^{N\xi_{\sim 2}(1)/2-3cN^{1/5}/4},$$
$$\langle e^{H_{N,\sim 2}(\sigma)} \rangle_2 \leq e^{N\xi_{\sim 2}(1)/2+cN^{1/5}/4}.$$

The conclusion follows on the intersection of these events, after adjusting c. ☐

We now turn to the proof of Proposition 7.8.10. By rotational invariance of gaussians, we may assume $\nabla^2 H_N(0)$ is diagonal while keeping the law of $H_{N,\sim 2}$ unchanged. For $i,j \in [N]$, define

$$X_{i,j} = \left\langle \sigma_i \sigma_j \left(e^{H_{N,\sim 2}(\sigma) - N\xi_{\sim 2}(1)/2} - \mathbf{1}[i=j] \right) \right\rangle_2,$$

and note that this equals the (i,j) entry of the matrix appearing in Proposition 7.8.10. For $\sigma \in \mathbb{R}^N$ and $i \in [N]$, let $\sigma_{\sim i} \in \mathbb{R}^{N-1}$ denote σ with coordinate i omitted. Similarly, for $i \neq j$, let $\sigma_{\sim i,j} \in \mathbb{R}^{N-2}$ denote σ with coordinates i and j omitted, and by slight abuse of notation let $\sigma_{\sim i,i} = \sigma_{\sim i}$. For $i,j \in [N]$ (possibly with $i = j$) define analogously to $\widetilde{T}$

$$\widetilde{T}_{i,j} := \left\{ \sigma \in S_N : \left\langle \mathbf{1}[|R(\sigma_{\sim i,j}, \tau_{\sim i,j})| \geq 2N^{-2/5}] e^{H_{N,\sim 2}(\tau)} \right\rangle_2 \leq e^{N\xi_{\sim 2}(1)/2-cN^{1/5}} \right\}, \tag{7.28}$$

where we recall the Gibbs average is with respect to $\tau \sim \langle \cdot \rangle_2$. Then define

$$\begin{aligned}
\widetilde{X}_{i,j} &= \left\langle \mathbf{1}[|\sigma_i|, |\sigma_j| \leq \log N, \sigma \in \widetilde{T}_{i,j}] \sigma_i \sigma_j \left(e^{H_{N,\sim 2}(\sigma) - N\xi_{\sim 2}(1)/2} - \mathbf{1}[i=j] \right) \right\rangle_2 \\
\widehat{X}_{i,j} &= \Big\langle \mathbf{1}[|\sigma_i^1|, |\sigma_j^1|, |\sigma_i^2|, |\sigma_j^2| \leq \log N, |R(\sigma^1_{\sim i,j}, \sigma^2_{\sim i,j})| \leq 2N^{-2/5}] \\
&\qquad \sigma_i^1 \sigma_j^1 \sigma_i^2 \sigma_j^2 \left(e^{H_{N,\sim 2}(\sigma^1) - N\xi_{\sim 2}(1)/2} - \mathbf{1}[i=j] \right) \left(e^{H_{N,\sim 2}(\sigma^2) - N\xi_{\sim 2}(1)/2} - \mathbf{1}[i=j] \right) \Big\rangle_2.
\end{aligned}$$

Note that $\widehat{X}_{i,j}$ is the contribution to $X_{i,j}^2$ coming from σ^1, σ^2 that are both not localized to coordinate i or j and have small overlap. The following two lemmas reduce the task of controlling $X_{i,j}^2$ to bounding $\mathbb{E}_{\sim 2} \widehat{X}_{i,j}$. They are proved by manipulating the typicality truncations $\widetilde{T}$ and $\widetilde{T}_{i,j}$ similarly to the proofs above; we defer these proofs to Subsection 7.B.3.

Lemma 7.8.15. *For each $i,j \in [N]$, with probability $1 - e^{-c\log^2 N}$ over $H_{N,\sim 2}$,*

$$X_{i,j}^2 \leq 2\widetilde{X}_{i,j}^2 + e^{-c\log^2 N}.$$

Lemma 7.8.16. *For each $i,j \in [N]$,*

$$\mathop{\mathbb{E}}_{\sim 2} \widetilde{X}_{i,j}^2 \leq \mathop{\mathbb{E}}_{\sim 2} \widehat{X}_{i,j} + e^{-cN^{1/5}}.$$

We now turn to bounding the $\mathbb{E}_{\sim 2} \widehat{X}_{i,j}$. This is achieved by the following pair of propositions.

Proposition 7.8.17. *For any $i \in [N]$, we have $\mathbb{E}_{\sim 2} \widehat{X}_{i,i} \le C(N\gamma_1^4 + N^{-1})$.*

Proposition 7.8.18. *For any distinct $i, j \in [N]$, we have $\mathbb{E}_{\sim 2} \widehat{X}_{i,j} \le C(\gamma_1^4 + N^{-2})$.*

Throughout the next two proofs, $\langle \cdot \rangle_2$ denotes expectation w.r.t. $\sigma^1, \sigma^2 \sim \langle \cdot \rangle_2$, and we write $R = R(\sigma^1, \sigma^2)$, $R_{\sim i} = R(\sigma^1_{\sim i}, \sigma^2_{\sim i})$, and $R_{\sim i,j} = R(\sigma^1_{\sim i,j}, \sigma^2_{\sim i,j})$.

Proof of Proposition 7.8.17. By direct calculation,

$$
\begin{aligned}
\mathop{\mathbb{E}}_{\sim 2} \widehat{X}_{i,i} = \mathop{\mathbb{E}}_{\sim 2} \Big\langle & \mathbf{1}[|\sigma_i^1|, |\sigma_i^2| \le \log N, |R_{\sim i}| \le 2N^{-2/5}] \\
& (\sigma_i^1)^2 (\sigma_i^2)^2 \left(e^{H_{N,\sim 2}(\sigma^1) - N\xi_{\sim 2}(1)/2} - 1\right)\left(e^{H_{N,\sim 2}(\sigma^2) - N\xi_{\sim 2}(1)/2} - 1\right) \Big\rangle_2 \\
= \Big\langle & \mathbf{1}[|\sigma_i^1|, |\sigma_i^2| \le \log N, |R_{\sim i}| \le 2N^{-2/5}] (\sigma_i^1)^2 (\sigma_i^2)^2 \left(e^{N\xi_{\sim 2}(R)} - 1\right) \Big\rangle_2 .
\end{aligned}
$$

In view of Proposition 7.8.8, σ_i^1 and σ_i^2 are subgaussian of scale $O(1)$ and R is subgaussian of scale $O(N^{-1/2})$. We will see that the above integral is dominated by $|\sigma_i^1| \asymp |\sigma_i^2| \asymp 1$ and $|R| \asymp N^{-1/2}$, in which case Taylor expanding $e^{N\xi_{\sim 2}(R)}$ shows this integral has the desired scale. Formally, we can write the above integral as $Y_{i,i}^{(1)} + Y_{i,i}^{(2)}$, where

$$
\begin{aligned}
Y_{i,i}^{(1)} &= \Big\langle \mathbf{1}[|\sigma_i^1|, |\sigma_i^2| \le \log N, |R_{\sim i}| \le 2N^{-1/2}\log N] (\sigma_i^1)^2 (\sigma_i^2)^2 \left(e^{N\xi_{\sim 2}(R)} - 1\right) \Big\rangle_2 \\
Y_{i,i}^{(2)} &= \Big\langle \mathbf{1}[|\sigma_i^1|, |\sigma_i^2| \le \log N, 2N^{-1/2}\log N \le |R_{\sim i}| \le 2N^{-2/5}] (\sigma_i^1)^2 (\sigma_i^2)^2 \left(e^{N\xi_{\sim 2}(R)} - 1\right) \Big\rangle_2
\end{aligned}
$$

are the contributions from $|R_{\sim i}|$ smaller and larger than $2N^{-1/2}\log N$.

We first address $Y_{i,i}^{(2)}$. Note that on the event in the indicator in $Y_{i,i}^{(2)}$, $|R| \le 3N^{-2/5}$. Thus, as $\gamma_1^2 \le N^{-4/5}$,

$$
N|\xi_{\sim 2}(R)| \le N\gamma_1^2 R + O(NR^3) = O(N^{-1/5}).
$$

It follows that $|e^{N\xi_{\sim 2}(R)} - 1| \le 1$. Then, by Cauchy–Schwarz,

$$
\begin{aligned}
Y_{i,i}^{(2)} &\le \Big\langle \mathbf{1}[|R| \ge N^{-1/2}\log N] (\sigma_i^1)^2 (\sigma_i^2)^2 \Big\rangle_2 \\
&\le \Big\langle \mathbf{1}[|R| \ge N^{-1/2}\log N] \Big\rangle_2^{1/2} \big\langle (\sigma_i^1)^2 (\sigma_i^2)^2 \big\rangle_2^{1/2} \\
&\le e^{-c\log^2 N} \cdot O(1) = e^{-c\log^2 N}, \qquad (7.29)
\end{aligned}
$$

where we have used Proposition 7.8.8(1) for the tail probability and Proposition 7.8.8(2) for the coordinate moments.

Next we turn to $Y_{i,i}^{(1)}$. On the event in the indicator in $Y_{i,i}^{(1)}$, $|R| \le 3N^{-1/2}\log N$, so

$$
N|\xi_{\sim 2}(R)| \le N\gamma_1^2 R + O(NR^3) = O(N^{-3/10}\log N),
$$

where we recall $\gamma_1^2 \le N^{-4/5}$. Thus, Taylor expanding the exponential and $\xi_{\sim 2}$,

$$
\begin{aligned}
Y_{i,i}^{(1)} = \Big\langle & \mathbf{1}[|\sigma_i^1|, |\sigma_i^2| \le \log N, |R_{\sim i}| \le 2N^{-1/2}\log N] (\sigma_i^1)^2 (\sigma_i^2)^2 \\
& \left(N\xi_{\sim 2}(R) + \tfrac{1}{2}N^2\xi_{\sim 2}(R)^2 + \tfrac{1}{6}N^3\xi_{\sim 2}(R)^3\right) \Big\rangle_2 + O(N^{-6/5}\log^8 N) \\
= \Big\langle & \mathbf{1}[|\sigma_i^1|, |\sigma_i^2| \le \log N, |R_{\sim i}| \le 2N^{-1/2}\log N] (\sigma_i^1)^2 (\sigma_i^2)^2 \\
& \left(N(\gamma_1^2 R + \gamma_3^2 R^3 + \gamma_4^2 R^4) + \tfrac{1}{2}N^2(\gamma_1^2 R + \gamma_3^2 R^3)^2 + \tfrac{1}{6}N^3\gamma_1^6 R^3\right) \Big\rangle_2 + O(N^{-11/10}\log^9 N).
\end{aligned}
$$

By exchangeability of $(\sigma_i^1, -\sigma_i^1)$, $(\sigma_i^2, -\sigma_i^2)$, and $(R_{\sim i}, -R_{\sim i})$, all the odd degree in R terms vanish, leaving

$$Y_{i,i}^{(1)} = \tfrac{1}{2}N^2\gamma_1^4 Q_2 + (N\gamma_4^2 + N^2\gamma_1^2\gamma_3^2)Q_4 + \tfrac{1}{2}N^2\gamma_3^4 Q_6 + o(N^{-1}),$$

and where we have introduced the notation

$$Q_k := \left\langle \mathbf{1}[|\sigma_i^1|, |\sigma_i^2| \le \log N, |R_{\sim i}| \le 2N^{-1/2}\log N](\sigma_i^1)^2(\sigma_i^2)^2 R^k \right\rangle_2.$$

By Cauchy–Schwarz and Proposition 7.8.8, for each $k \in \{2, 4, 6\}$,

$$Q_k \le \left\langle (\sigma_i^1)^4(\sigma_i^2)^4 \right\rangle_2^{1/2} \left\langle R^{2k} \right\rangle_2^{1/2} = O(N^{-k/2}).$$

This implies $Y_{i,i}^{(1)} \le C(N\gamma_1^4 + N^{-1})$. Combining with the bound (7.29) on $Y_{i,i}^{(2)}$ implies the result. □

Proof of Proposition 7.8.18. We calculate as above

$$\begin{aligned}
\mathop{\mathbb{E}}_{\sim 2} \widehat{X}_{i,j} &\le \mathop{\mathbb{E}}_{\sim 2} \Big\langle \mathbf{1}[|\sigma_i^1|, |\sigma_j^1|, |\sigma_i^2|, |\sigma_j^2| \le \log N, |R_{\sim i,j}| \le 2N^{-2/5}] \\
&\qquad \sigma_i^1\sigma_j^1\sigma_i^2\sigma_j^2 e^{H_{N,\sim 2}(\sigma^1) + H_{N,\sim 2}(\sigma^2) - N\xi_{\sim 2}(1)} \Big\rangle_2 \\
&= \left\langle \mathbf{1}[|\sigma_i^1|, |\sigma_j^1|, |\sigma_i^2|, |\sigma_j^2| \le \log N, |R_{\sim i,j}| \le 2N^{-2/5}]\sigma_i^1\sigma_j^1\sigma_i^2\sigma_j^2 e^{N\xi_{\sim 2}(R)} \right\rangle_2.
\end{aligned}$$

Our strategy for evaluating this will be similar as above, except that because this integral contains $\sigma_i^1\sigma_j^1\sigma_i^2\sigma_j^2$ instead of $(\sigma_i^1)^2(\sigma_i^2)^2$, we will need to expand the exponential more carefully to obtain cancellations in these terms. Formally, we write the above integral as $Y_{i,j}^{(1)} + Y_{i,j}^{(2)}$ for

$$\begin{aligned}
Y_{i,j}^{(1)} &= \left\langle \mathbf{1}[|\sigma_i^1|, |\sigma_j^1|, |\sigma_i^2|, |\sigma_j^2| \le \log N, |R_{\sim i,j}| \le 2N^{-1/2}\log N]\sigma_i^1\sigma_j^1\sigma_i^2\sigma_j^2 e^{N\xi_{\sim 2}(R)} \right\rangle_2, \\
Y_{i,j}^{(2)} &= \left\langle \mathbf{1}[|\sigma_i^1|, |\sigma_j^1|, |\sigma_i^2|, |\sigma_j^2| \le \log N, 2N^{-1/2}\log N \le |R_{\sim i,j}| \le 2N^{-2/5}]\sigma_i^1\sigma_j^1\sigma_i^2\sigma_j^2 e^{N\xi_{\sim 2}(R)} \right\rangle_2.
\end{aligned}$$

Identically to the previous proof, on the event in the indicator in $Y_{i,j}^{(2)}$ we have $N|\xi_{\sim 2}(R)| = O(N^{-1/5})$, so $e^{N\xi_{\sim 2}(R)} \le 2$. Then, by Cauchy–Schwarz and Proposition 7.8.8,

$$\begin{aligned}
|Y_{i,j}^{(2)}| &\le 2\left\langle \mathbf{1}[|R| \ge N^{-1/2}\log N]|\sigma_i^1\sigma_j^1\sigma_i^2\sigma_j^2| \right\rangle_2 \\
&\le 2\langle \mathbf{1}[|R| \ge N^{-1/2}\log N]\rangle_2^{1/2} \left\langle (\sigma_i^1)^4(\sigma_i^2)^4 \right\rangle_2^{1/4} \left\langle (\sigma_j^1)^4(\sigma_j^2)^4 \right\rangle_2^{1/4} \\
&\le 2e^{-c\log^2 N} \cdot O(1) \cdot O(1) = e^{-c\log^2 N}.
\end{aligned}$$

To address $Y_{i,j}^{(1)}$, define $\Delta_i = \sigma_i^1\sigma_i^2/N$ and $\Delta_j = \sigma_j^1\sigma_j^2/N$, the contributions to R coming from the ith and jth coordinate, respectively. Then, by exchangeability of $(\sigma_i^1, -\sigma_i^1)$ and $(\sigma_j^1, -\sigma_j^1)$,

$$\begin{aligned}
4Y_{i,j}^{(1)} = \Big\langle & \mathbf{1}[|\sigma_i^1|, |\sigma_j^1|, |\sigma_i^2|, |\sigma_j^2| \le \log N, |R_{\sim i,j}| \le 2N^{-1/2}\log N] \\
& \sigma_i^1\sigma_j^1\sigma_i^2\sigma_j^2 (e^{N\xi_{\sim 2}(R_{\sim i,j} + \Delta_i + \Delta_j)} - e^{N\xi_{\sim 2}(R_{\sim i,j} + \Delta_i - \Delta_j)} - e^{N\xi_{\sim 2}(R_{\sim i,j} - \Delta_i + \Delta_j)} + e^{N\xi_{\sim 2}(R_{\sim i,j} - \Delta_i - \Delta_j)}) \Big\rangle_2.
\end{aligned}$$

Note that on the event in this indicator, $|R_{\sim i,j} \pm \Delta_i \pm \Delta_j| \le 3N^{-1/2}\log N$. Define

$$\kappa(x) = e^{N\xi_{\sim 2}(x)},$$

and note that

$$\begin{aligned}
\sup_{|x| \le 3N^{-1/2}\log N} \kappa^{(4)}(x) = \sup_{|x| \le 3N^{-1/2}\log N} \Big(& N\xi_{\sim 2}^{(4)}(x) + 4N^2\xi_{\sim 2}'(x)\xi_{\sim 2}^{(3)}(x) + 3N^2\xi_{\sim 2}''(x)^2 + \\
& + 6N^3\xi_{\sim 2}'(x)^2\xi_{\sim 2}''(x) + N^4\xi_{\sim 2}'(x)^4\Big)\kappa(x) = O(N^{6/5}),
\end{aligned}$$

where we have used that $\sup_{|x|\le 3N^{-1/2}\log N}\kappa(x)\le 2$ and $\gamma_1^2\le N^{-4/5}$. Since $|\Delta_i|,|\Delta_j|\le N^{-1}\log^2 N$ on the event in the indicator, for $s_i,s_j\in\{\pm 1\}$,

$$\begin{aligned}e^{N\xi_{\sim 2}(R_{\sim i,j}+s_i\Delta_i+s_j\Delta_j)} &= \kappa(R_{\sim i,j})+\kappa'(R_{\sim i,j})(s_i\Delta_i+s_j\Delta_j)+\tfrac{1}{2}\kappa''(R_{\sim i,j})(s_i\Delta_i+s_j\Delta_j)^2\\ &\quad+\tfrac{1}{6}\kappa^{(3)}(R_{\sim i,j})(s_i\Delta_i+s_j\Delta_j)^3+O(N^{6/5})\cdot(2N^{-1}\log^2 N)^4.\end{aligned}$$

It follows that

$$\begin{aligned}&e^{N\xi_{\sim 2}(R_{\sim i,j}+\Delta_i+\Delta_j)}-e^{N\xi_{\sim 2}(R_{\sim i,j}+\Delta_i-\Delta_j)}-e^{N\xi_{\sim 2}(R_{\sim i,j}-\Delta_i+\Delta_j)}+e^{N\xi_{\sim 2}(R_{\sim i,j}-\Delta_i-\Delta_j)}\\ &=4\kappa''(R_{\sim i,j})\Delta_i\Delta_j+O(N^{-14/5}\log^8 N),\end{aligned}$$

and thus

$$\begin{aligned}Y^{(1)}_{i,j} &= \Big\langle \mathbf{1}[|\sigma_i^1|,|\sigma_j^1|,|\sigma_i^2|,|\sigma_j^2|\le\log N, |R_{\sim i,j}|\le 2N^{-1/2}\log N]\sigma_i^1\sigma_j^1\sigma_i^2\sigma_j^2\Delta_i\Delta_j\kappa''(R_{\sim i,j})\Big\rangle_2\\ &\quad+O(N^{-14/5}\log^{12} N)\\ &=\Big\langle \mathbf{1}[|\sigma_i^1|,|\sigma_j^1|,|\sigma_i^2|,|\sigma_j^2|\le\log N, |R_{\sim i,j}|\le 2N^{-1/2}\log N]\\ &\quad(\sigma_i^1\sigma_j^1\sigma_i^2\sigma_j^2)^2\left(N^{-1}\xi''_{\sim 2}(R_{\sim i,j})+\xi'_{\sim 2}(R_{\sim i,j})^2\right)e^{N\xi_{\sim 2}(R_{\sim i,j})}\Big\rangle_2+o(N^{-2}).\end{aligned}$$

On the event in this indicator, $e^{N\xi_{\sim 2}(R_{\sim i,j})}\le 2$, and therefore $\xi'_{\sim 2}$ and $\xi''_{\sim 2}$ can be Taylor expanded to obtain

$$Y^{(1)}_{i,j}\le Y^{(3)}_{i,j}+Y^{(4)}_{i,j}+o(N^{-2}),$$

where

$$\begin{aligned}Y^{(3)}_{i,j} &= 6\gamma_3^2N^{-1}\Big\langle \mathbf{1}[|\sigma_i^1|,|\sigma_j^1|,|\sigma_i^2|,|\sigma_j^2|\le\log N, |R_{\sim i,j}|\le 2N^{-1/2}\log N]\\ &\quad(\sigma_i^1\sigma_j^1\sigma_i^2\sigma_j^2)^2R_{\sim i,j}e^{N\xi_{\sim 2}(R_{\sim i,j})}\Big\rangle_2,\\ Y^{(4)}_{i,j} &= \mathop{\mathbb{E}}_{\mu_{H_{N,2}}}\Big\langle \mathbf{1}[|\sigma_i^1|,|\sigma_j^1|,|\sigma_i^2|,|\sigma_j^2|\le\log N, |R_{\sim i,j}|\le 2N^{-1/2}\log N]\\ &\quad(\sigma_i^1\sigma_j^1\sigma_i^2\sigma_j^2)^2\left(12\gamma_4^2N^{-1}R_{\sim i,j}^2+(\gamma_1^2+3\gamma_3^2R_{\sim i,j}^2)^2\right)e^{N\xi_{\sim 2}(R_{\sim i,j})}\Big\rangle_2\end{aligned}$$

On the event in these indicators, we further have

$$|R^2-R_{\sim i,j}^2|=|R-R_{\sim i,j}||R+R_{\sim i,j}|\le(|\Delta_i|+|\Delta_j|)\cdot 5N^{-1/2}\log N=O(N^{-3/2}\log^3 N). \tag{7.30}$$

From this it readily follows that

$$\begin{aligned}Y^{(4)}_{i,j} &\le 2\Big\langle \mathbf{1}[|\sigma_i^1|,|\sigma_j^1|,|\sigma_i^2|,|\sigma_j^2|\le\log N, |R_{\sim i,j}|\le 2N^{-1/2}\log N]\\ &\quad(\sigma_i^1\sigma_j^1\sigma_i^2\sigma_j^2)^2\left(12\gamma_4^2N^{-1}R^2+(\gamma_1^2+3\gamma_3^2R^2)^2\right)\Big\rangle_2+o(N^{-2})\\ &=2\gamma_1^4\widetilde{Q}_0+(12\gamma_1^2\gamma_3^2+24\gamma_4^2N^{-1})\widetilde{Q}_2+18\gamma_3^2\widetilde{Q}_4+o(N^{-2}),\end{aligned}$$

where

$$\widetilde{Q}_k=\Big\langle \mathbf{1}[|\sigma_i^1|,|\sigma_j^1|,|\sigma_i^2|,|\sigma_j^2|\le\log N, |R_{\sim i,j}|\le 2N^{-1/2}\log N](\sigma_i^1\sigma_j^1\sigma_i^2\sigma_j^2)^2R^k\Big\rangle_2.$$

By Cauchy–Schwarz and Proposition 7.8.8, for each $k\in\{0,2,4\}$,

$$\widetilde{Q}_k\le\left\langle(\sigma_i^1)^8(\sigma_i^2)^8\right\rangle_2^{1/4}\left\langle(\sigma_j^1)^8(\sigma_j^2)^8\right\rangle_2^{1/4}\left\langle R^{2k}\right\rangle_2^{1/2}=O(N^{-k/2}).$$

This implies $Y_{i,j}^{(4)} \le C(\gamma_1^4 + N^{-2})$. To control $Y_{i,j}^{(3)}$, we recall that $|N\xi_{\sim 2}(R_{\sim i,j})| = O(N^{-3/10}\log N)$ and Taylor expand the exponential:

$$\begin{aligned}
Y_{i,j}^{(3)} &= 6\gamma_3^2 N^{-1}\Big\langle \mathbf{1}[|\sigma_i^1|,|\sigma_j^1|,|\sigma_i^2|,|\sigma_j^2| \le \log N, |R_{\sim i,j}| \le 2N^{-1/2}\log N] \\
&\qquad (\sigma_i^1\sigma_j^1\sigma_i^2\sigma_j^2)^2 R_{\sim i,j}(1 + N\xi_{\sim 2}(R_{\sim i,j}))\Big\rangle_2 + o(N^{-2}) \\
&= 6\gamma_3^2 N^{-1}\Big\langle \mathbf{1}[|\sigma_i^1|,|\sigma_j^1|,|\sigma_i^2|,|\sigma_j^2| \le \log N, |R_{\sim i,j}| \le 2N^{-1/2}\log N] \\
&\qquad (\sigma_i^1\sigma_j^1\sigma_i^2\sigma_j^2)^2 (R_{\sim i,j} + N\gamma_1^2 R_{\sim i,j}^2 + N\gamma_3^2 R_{\sim i,j}^4)\Big\rangle_2 + o(N^{-2}).
\end{aligned}$$

By exchangeability of $(R_{\sim i,j}, -R_{\sim i,j})$, the contribution of the term $R_{\sim i,j}$ vanishes. By (7.30), we can further estimate $R_{\sim i,j}^2$ with R^2, obtaining

$$\begin{aligned}
Y_{i,j}^{(3)} &= 6\gamma_3^2\Big\langle \mathbf{1}[|\sigma_i^1|,|\sigma_j^1|,|\sigma_i^2|,|\sigma_j^2| \le \log N, |R_{\sim i,j}| \le 2N^{-1/2}\log N] \\
&\qquad (\sigma_i^1\sigma_j^1\sigma_i^2\sigma_j^2)^2 (\gamma_1^2 R^2 + \gamma_3^2 R^4)\Big\rangle_2 + o(N^{-2}) \\
&= 6\gamma_1^2\gamma_3^2 \widetilde{Q}_2 + 6\gamma_3^4 \widetilde{Q}_4 + o(N^{-2}) \le C(\gamma_1^2 N^{-1} + N^{-2}).
\end{aligned}$$

Combining all of the above estimates concludes the proof. □

Proof of Proposition 7.8.10. By a union bound, the event in Lemma 7.8.15 holds for all $i,j \in [N]$ with probability $1 - e^{-c\log^2 N}$ (over $H_{N,\sim 2}$). On this event,

$$\left\| \frac{Z_N}{Z_{N,2}e^{N\xi_{\sim 2}(1)/2}}\langle \sigma\sigma^\top\rangle - \langle \sigma\sigma^\top\rangle_2 \right\|_F^2 = \sum_{i,j=1}^N X_{i,j}^2 \le 2\sum_{i,j=1}^N \widetilde{X}_{i,j}^2 + e^{-c\log^2 N}. \tag{7.31}$$

Combining Lemma 7.8.16 and Propositions 7.8.17 and 7.8.18 shows that

$$\mathop{\mathbb{E}}_{\sim 2} \sum_{i,j=1}^N \widetilde{X}_{i,j}^2 \le 2C(N^2\gamma_1^4 + 1) + e^{-cN^{1/5}} \le 3C(N^2\gamma_1^4 + 1).$$

Thus, with probability $2/3$ over $H_{N,\sim 2}$, $\sum_{i,j=1}^N \widetilde{X}_{i,j}^2 \le 9C(N^2\gamma_1^4 + 1)$. Combining with (7.31) and taking a final union bound shows that with probability $1/2$ over $H_{N,\sim 2}$,

$$\left\| \frac{Z_N}{Z_{N,2}e^{N\xi_{\sim 2}(1)/2}}\langle \sigma\sigma^\top\rangle - \langle \sigma\sigma^\top\rangle_2 \right\|_F^2 \le 18C(N^2\gamma_1^4 + 1) + e^{-c\log^2 N} \le 20C(N^2\gamma_1^4 + 1).$$

The result follows after adjusting C. □

7.8.3 From positive to very high probability

In this section, we boost the positive probability bound on the second moment matrix to a very high probability bound. To this end, we will show that an appropriate proxy function for the second moment matrix is very Lipschitz. This will imply the desired concentration by standard gaussian concentration.

Let $\boldsymbol{g} \in \mathbb{R}^{N+N^3+N^4+\cdots+N^{p_*}}$ be the vectorized collection of all gaussian interactions corresponding to $H_{N,\sim 2}$. Throughout, as in the previous subsection, we will condition on the event E_2 from Definition 7.8.13 over $H_{N,2}$, which holds with probability $1 - e^{-cN^{1/5}}$. We define the following functions of $\boldsymbol{g}$:

$$\begin{aligned}
F_1(\boldsymbol{g}) &= B(1 + \gamma_1^2 N)\cdot\left(1 - |\log\frac{Z_N}{Z_{N,2}} - \frac{N\xi_{\sim 2}(1)}{2}|\right) \\
F_2(\boldsymbol{g}) &= B(1 + \gamma_1^2 N) - \left\|\langle \sigma\sigma^\top\rangle\right\|_{\mathsf{op}} \\
F(\boldsymbol{g}) &= \max(\min(F_1(\boldsymbol{g}), F_2(\boldsymbol{g})), 0)\,,
\end{aligned}$$

where B is a sufficiently large constant (specified in the proof of Lemma 7.8.19). If we can show that with high probability over $\boldsymbol{g}$, $\min(F_1(\boldsymbol{g}), F_2(\boldsymbol{g})) \geq 0$, then the conclusion follows. Indeed, from $F_2(\boldsymbol{g}) \geq 0$, we immediately obtain $\left\|\langle\sigma\sigma^\top\rangle\right\|_{\mathsf{op}} \leq B(1+\gamma_1^2 N)$. F_1 allows control over the free energy of the p-spin model in terms of that of the corresponding 2-spin model. This gives good control over the overlaps (in a manner to be made precise shortly), which is crucial for establishing the high probability statement. It is also important earlier in this section, in showing that the free energy of the p-spin model concentrates well.

Towards this, we first start with the positive probability statement, which was essentially established in the previous subsections.

Lemma 7.8.19. *There exists a constant $B > 0$ such that with probability at least $\frac{1}{3}$, we have $F(\boldsymbol{g}) \geq \frac{B}{2}(1+\gamma_1^2 N)$.*

Proof. By Proposition 7.8.9, with probability $1 - O(N^{-1/15})$ over $\boldsymbol{g}$, we have $F_1(\boldsymbol{g}) \geq \frac{B}{2}(1+\gamma_1^2 N)$, so $\frac{Z_N}{Z_{N,2}e^{N\xi_{\sim 2}(1)/2}} \geq e^{-1/2}$. Intersecting this with the event from Proposition 7.8.10 implies that with probability at least $\frac{1}{3}$,

$$
\begin{aligned}
\left\|\langle\sigma\sigma^\top\rangle\right\|_{\mathsf{op}} &\leq e^{1/2}\left\|\langle\sigma\sigma^\top\rangle_2\right\|_{\mathsf{op}} + e^{1/2}\sqrt{C(1+\gamma_1^4 N^2)} \\
&\leq \frac{B}{2}(1+\gamma_1^2 N)\,,
\end{aligned}
$$

where we have used E_2 to apply Proposition 7.8.8 and after appropriately picking B. ☐

Let $\mathcal{E}$ denote the $H_{N,2}$-measurable event from Corollary 7.8.14:

$$
\{\boldsymbol{g} : \int \mathbf{1}[|R(\sigma^1,\sigma^2)| \geq N^{-2/5}\} e^{H_N(\sigma^1)+H_N(\sigma^2)} \mathsf{d}\rho^{\otimes 2}(\sigma^1,\sigma^2) \leq Z_{N,2}^2 e^{N\xi_{\sim 2}(1) - cN^{1/5}}\}\,,
$$

which holds with probability $1 - e^{-cN^{1/5}}$ over $\boldsymbol{g}$. The key observation is that this gives us good control on the overlaps.

Lemma 7.8.20. *On $\mathcal{E}$, if $F_1(\boldsymbol{g}) \geq 0$, then for any $0 \leq p \leq \log^2 N$, we have*

$$
\left\|\langle\sigma^{\otimes p}\rangle\right\|_F^2 \leq O(N^{3p/5}).
$$

Proof. By splitting up the expectation based on whether $|R(\sigma^1,\sigma^2)| \geq N^{-2/5}$, on $\mathcal{E}$ we have

$$
\begin{aligned}
\left\|\langle\sigma^{\otimes p}\rangle\right\|_F^2 &= \langle N^p R(\sigma^1,\sigma^2)^p\rangle \\
&\leq N^{3p/5} + N^{2p}\frac{1}{Z_N^2}\int \mathbf{1}[|R(\sigma^1,\sigma^2)| \geq N^{-2/5}\} e^{H_N(\sigma^1)+H_N(\sigma^2)} \mathsf{d}\rho^{\otimes 2}(\sigma^1,\sigma^2) \\
&\leq N^{3p/5} + N^{2p}\frac{Z_{N,2}^2}{Z_N^2} e^{N\xi_{\sim 2}(1) - cN^{1/5}} && \text{(Definition of } \mathcal{E}) \\
&\leq N^{3p/5} + N^{2p} e^{1-cN^{1/5}}\,,
\end{aligned}
$$

where the last line used $F_1(\boldsymbol{g}) \geq 0$. Since $p \leq \log^2 N$, the above quantity is $O(N^{3p/5})$, as desired. ☐

The above is a crucial input to prove Lipschitzness of F on $\mathcal{E}$.

Lemma 7.8.21. *The function F is $O((1+\gamma_1^2 N)N^{-1/10})$-Lipschitz restricted to $\mathcal{E}$.*

Before we prove this, let us see how it implies Proposition 7.8.2, restated for convenience.

Proposition 7.8.2. *There is a $H_{N,2}$-measurable event with probability $1 - e^{-cN^{1/5}}$ on which the following holds with probability $1 - e^{-cN^{1/5}}$ over $H_{N,\sim 2}$.*

1. *The partition functions Z_N, $Z_{N,2}$ satisfy*

$$\left|\log \frac{Z_N}{Z_{N,2}} - \frac{N\xi_{\sim 2}(1)}{2}\right| \le 1/2.$$

2. *The Gibbs measure satisfies $\|\langle \sigma\sigma^\top \rangle\|_{\mathsf{op}} \le C(1+\gamma_1^2 N)$.*

Proof of Proposition 7.8.2. By Kirszbraun's extension theorem, we can extend F to $\widetilde{F}$ such that each $\widetilde{F}$ has the same Lipschitz constant as F and agrees with F on $\mathcal{E}$. We can now apply gaussian concentration to $\widetilde{F}$ to conclude that

$$\Pr\left[|\widetilde{F}(\boldsymbol{g}) - \mathbb{E}\,\widetilde{F}(\boldsymbol{g})| \ge \tfrac{B}{4}(1+\gamma_1^2 N)\right] \ge 1 - e^{-cN^{1/5}}\,. \tag{7.32}$$

By Lemma 7.8.19, with probability at least $\frac{1}{3}$, we have $F(\boldsymbol{g}) \ge \frac{B}{2}(1+\gamma_1^2 N)$. Upon further intersection with $\mathcal{E}$ (where $\widetilde{F}(\boldsymbol{g}) = F(\boldsymbol{g})$) and the event from (7.32), we conclude $\mathbb{E}\,\widetilde{F}(\boldsymbol{g}) \ge \frac{B}{4}(1+\gamma_1^2 N)$. Thus,

$$\begin{aligned}
\Pr[F(\boldsymbol{g}) = 0] &\le \Pr[\mathcal{E}^c] + \Pr[\widetilde{F}(\boldsymbol{g}) = 0] \\
&\le e^{-cN^{1/5}} + \Pr\left[|\widetilde{F}(\boldsymbol{g}) - \mathbb{E}\,\widetilde{F}(\boldsymbol{g})| \ge \frac{B(1+\gamma_1^2 N)}{4}\right] \\
&\le e^{-cN^{1/5}},
\end{aligned}$$

after adjusting c.

□

Finally, let us prove the Lipschitz bound.

Proof of Lemma 7.8.21. The set $\mathcal{E}$ is a convex set in $\boldsymbol{g}$. Indeed, $e^{H_N(\sigma)}$ is a convex function of $\boldsymbol{g}$, so the LHS of the inequality $\mathcal{E}$ is convex in $\boldsymbol{g}$, whereas the RHS does not depend on $\boldsymbol{g}$, and sublevel sets of convex functions are convex. Furthermore, F is absolutely continuous (hence differentiable almost everywhere), so to prove F is Lipschitz on $\mathcal{E}$ it suffices to bound $\|\nabla F\|$ on $\mathcal{E}$, wherever it is defined.

The easier case is if $\min(F_1(\boldsymbol{g}), F_2(\boldsymbol{g})) < 0$. In this case, $F(\boldsymbol{g}) = 0$ in an open neighborhood of $\boldsymbol{g}$, so $\nabla F(\boldsymbol{g}) = 0$ identically. Therefore, for the rest of the proof, assume $\min(F_1(\boldsymbol{g}), F_2(\boldsymbol{g})) \ge 0$. We will compute the gradient of the F_i's, and to simplify the calculation, we will take the gradient with respect to $\boldsymbol{g}_p \in \mathbb{R}^{N^p}$ corresponding to the degree-p disorder in $H_{N,\sim 2}$.

For $F_1(\boldsymbol{g})$, note that its only dependence on $\boldsymbol{g}$ is via $\log Z_N$, we have

$$\left\|\nabla_{\boldsymbol{g}_p} F_1(\boldsymbol{g})\right\| = B(1+\gamma_1^2 N)\left\|\nabla_{\boldsymbol{g}_p} \log Z_N\right\| = B(1+\gamma_1^2 N)\cdot \frac{\gamma_p}{N^{(p-1)/2}}\left\|\langle \sigma^{\otimes p}\rangle\right\|_F,$$

and since $F_1(\boldsymbol{g}) \ge 0$, we can apply Lemma 7.8.20 to conclude that

$$\begin{aligned}
\|\nabla_{\boldsymbol{g}} \log Z_N\|^2 &\lesssim \sum_{p \in [p_*]\setminus\{2\}} \gamma_p^2 N^{-(p-1)} \cdot N^{3p/5} \\
&\le \gamma_1^2 \cdot N^{3/5} + \sum_{p\ge 3} \gamma_p^2 N^{1-2p/5} \\
&\lesssim \gamma_1^2 \cdot N^{3/5} + N^{-1/5}
\end{aligned} \tag{7.33}$$

Since $\gamma_1^2 \le N^{-4/5}$, we conclude that $\|\nabla_{\boldsymbol{g}} F_1(\boldsymbol{g})\| \lesssim (1+\gamma_1^2 N)N^{-1/10}$, as desired.

Turning now to $F_2(\boldsymbol{g})$, we observe that $\left\|\langle \sigma\sigma^\top\rangle\right\|_{\mathsf{op}} = \langle \langle u,\sigma\rangle^2\rangle$, where u is the top eigenvector of $\langle \sigma\sigma^\top\rangle$ with $\|u\|_2 = 1$. By the envelope theorem, we can evaluate the gradient with u fixed. For any $v \in \mathbb{R}^{N^p}$ with $\|v\|_2 = 1$, we will upper bound $\langle v, \nabla_{\boldsymbol{g}_p} \langle \langle u,\sigma\rangle^2\rangle\rangle$. Applying the quotient rule yields

$$\begin{aligned}
\langle v, \nabla_{\boldsymbol{g}_p} \langle \langle u,\sigma\rangle^2\rangle\rangle &= \left\langle \langle u,\sigma\rangle^2 \langle v, \nabla_{\boldsymbol{g}_p} H(\sigma)\rangle\right\rangle - \left\langle \langle u,\sigma\rangle^2\right\rangle\left\langle \langle v, \nabla_{\boldsymbol{g}_p} H(\sigma)\rangle\right\rangle \\
&= \frac{\gamma_p}{N^{(p-1)/2}}\left(\left\langle \langle u,\sigma\rangle^2 \langle v, \sigma^{\otimes p}\rangle\right\rangle - \left\langle \langle u,\sigma\rangle^2\right\rangle\left\langle \langle v, \sigma^{\otimes p}\rangle\right\rangle\right)\,.
\end{aligned}$$

Consider the first term $\left\langle \langle u,\sigma\rangle^2 \langle v,\sigma^{\otimes p}\rangle \right\rangle$. Using Hölder's inequality with $q = 1+\log N$ and $q' = 1+\frac{1}{\log N}$, we see

$$\begin{aligned}
\left\langle \langle u,\sigma\rangle^2 \langle v,\sigma^{\otimes p}\rangle \right\rangle &\le \left\langle \langle u,\sigma\rangle^{2q'} \right\rangle^{1/q'} \left\langle \langle v,\sigma^{\otimes p}\rangle^q \right\rangle^{1/q} \\
&\le N^{1/\log N} \left\langle \langle u,\sigma\rangle^2 \right\rangle \langle v^{\otimes q}, \left\langle \sigma^{\otimes pq} \right\rangle \rangle^{1/q} \\
&\lesssim (1+\gamma_1^2 N) \langle v^{\otimes q}, \left\langle \sigma^{\otimes pq} \right\rangle \rangle^{1/q} \\
&\le (1+\gamma_1^2 N)\cdot N^{3p/10}\,,
\end{aligned}$$

where in the second to last line we have used $F_2(\boldsymbol{g}) \ge 0$ to apply the bound $\left\langle \langle u,\sigma\rangle^2 \right\rangle \le O(1+\gamma_1^2 N)$, and in the last line we have used $F_1(\boldsymbol{g}) \ge 0$, along with $pq \le O(\log N)$, to apply Lemma 7.8.20. The same argument upper bounds the contribution of the second term as $O\left((1+\gamma_1^2 N)N^{3p/10}\right)$. These bounds (aside from the common factor of $O(1+\gamma_1^2 N)$, which we can pull out), exactly match the ones used in the calculation as carried out for $F_1(\boldsymbol{g})$ in (7.33). Hence, the same argument ultimately yields $\|\nabla_{\boldsymbol{g}} F_2(\boldsymbol{g})\| \lesssim (1+\gamma_1^2 N)N^{-1/10}$, completing the proof. $\square$

Appendix

7.A Annealed Glauber dynamics on discrete domains

In this section, we collect the analogous results for weak functional inequalities for Glauber dynamics.

Definition 7.A.1 (Weak Poincaré for the hypercube)**.** We say π on $\{\pm 1\}^n$ satisfies a $(\rho_{\mathrm{PI}}, \varepsilon)$-weak Poincaré inequality for Glauber dynamics if for all functions f,

$$\mathsf{Var}_\pi[f] \le \frac{1}{\rho_{\mathrm{PI}}} \cdot \mathcal{E}(f,f) + \varepsilon \cdot \operatorname{osc}(f)^2.$$

Similarly, we say π satisfies a $(\rho_{\mathrm{LS}}, \varepsilon)$-weak modified log-Sobolev inequality if for all functions f,

$$\mathsf{Ent}_\pi[f] \le \frac{1}{\rho_{\mathrm{LS}}} \cdot \mathcal{E}(f, \log f) + \varepsilon \cdot \operatorname{osc}(\sqrt{f})^2.$$

Remark 7.A.2. The above definition is related to the continuous setting by using the discrete gradient, which can be bounded by $\operatorname{osc}(f)^2$.

Remark 7.A.3. A weak Poincaré inequality with sufficiently good parameters implies a true Poincaré inequality for Glauber dynamics. Indeed, any low conductance cut limits on the region of valid $(\rho_{\mathrm{PI}}, \varepsilon)$. Hence, by Cheeger, one can conclude that Glauber satisfies a true Poincaré inequality, with some loss in parameters.

First, we will need a concavity property for the Dirichlet form for Glauber dynamics, which is well-known. We provide a proof for the sake of self-containedness.

Fact 7.A.4. *Let π be a distribution on $\{\pm 1\}^n$, and $\pi = \mathbb{E}_{\boldsymbol{z}\sim\rho}\, \pi_{\boldsymbol{z}}$ a measure decomposition of π. Then $\mathcal{E}_\pi(f,f) \ge \mathbb{E}_{\boldsymbol{z}\sim\rho}\, \mathbb{E}_{\pi_{\boldsymbol{z}}}(f,f)$.*

Proof. For Glauber dynamics on the hypercube, we have

$$\begin{aligned}
\mathcal{E}_\pi(f,f) &= \frac{1}{n} \sum_{\|x-y\|_1 = 2} \frac{\pi(x)\pi(y)}{\pi(x)+\pi(y)} (f(x) - f(y))^2 \\
&= \frac{1}{n} \sum_{\|x-y\|_1 = 2} \frac{\mathbb{E}_{\boldsymbol{z}\sim\rho}\, \pi_{\boldsymbol{z}}(x)\, \mathbb{E}_{\boldsymbol{z}\sim\rho}\, \pi_{\boldsymbol{z}}(y)}{\mathbb{E}_{\boldsymbol{z}\sim\rho}\, \pi_{\boldsymbol{z}}(x) + \mathbb{E}_{\boldsymbol{z}\sim\rho}\, \pi_{\boldsymbol{z}}(y)} (f(x) - f(y))^2 \\
&\ge \mathop{\mathbb{E}}_{\boldsymbol{z}\sim\rho} \mathcal{E}_{\pi_{\boldsymbol{z}}}(f,f),
\end{aligned}$$

where the last line follows from concavity of the map $(a,b) \mapsto \frac{ab}{a+b}$ for $a, b > 0$. ◻

The following lemma transfers a true Poincaré inequality on π to a weak Poincaré inequality on π' for Glauber dynamics on the hypercube.

Lemma 7.A.5. *Let π, π' be distributions on $\{\pm 1\}^n$ such that π satisfies a ρ_{PI}-Poincaré inequality for Glauber dynamics and $\mathrm{TV}(\pi, \pi') \le \delta$. Then, π' satisfies a weak $(\rho_{\mathrm{PI}}, 2\delta)$-Poincaré inequality for Glauber dynamics.*

Proof. By definition, there exists a coupling $\mathcal{C}$ of (π, π') such that for $(x, x') \sim \mathcal{C}$, $\Pr[x \neq x'] \leq \delta$. The main difference in the proof compared to the Langevin case Lemma 7.4.9 is that the Dirichlet form comparison is now bounded in terms of $\operatorname{osc}(f)^2$ rather than $\sup \|\nabla f\|^2$. Indeed, by Fact 7.4.2 we have $\mathcal{E}_\pi(f, f) = \mathbb{E}_{\boldsymbol{x} \sim \pi} \mathbb{E}_{\boldsymbol{y} \sim P\boldsymbol{x}}(f(\boldsymbol{x}) - f(\boldsymbol{y}))^2$, and for any fixed x the function $\mathbb{E}_{\boldsymbol{y} \sim P x}(f(x) - f(\boldsymbol{y}))^2 \leq \operatorname{osc}(f)^2$, so the coupling allows to conclude that $\mathcal{E}_{\pi'}(f, f) \geq \mathcal{E}_\pi(f, f) - \delta \cdot \operatorname{osc}(f)^2$. Similarly, we can deduce that $\mathsf{Var}_\pi[f] \geq \mathsf{Var}_{\pi'} - \delta \cdot \operatorname{osc}(f)^2$. Hence, we have

$$\begin{aligned}
\mathcal{E}_{\pi'}(f, f) &\geq \mathcal{E}_\pi(f, f) - \delta \cdot \operatorname{osc}(f)^2 \\
&\geq \rho_{\mathrm{PI}} \cdot \mathsf{Var}_\pi[f] - \delta \cdot \operatorname{osc}(f)^2 && (\pi \text{ satisfies PI}) \\
&\geq \rho_{\mathrm{PI}} \cdot \mathsf{Var}_{\pi'}[f] - \delta \cdot \operatorname{osc}(f)^2 \left(1 + \rho_{\mathrm{PI}}\right) .
\end{aligned}$$

□

Remark 7.A.6. The above two results also hold more generally if P is the Markov chain associated to a Doob localization scheme (cf. [CE22, Section 2.3]), such as when P is Glauber dynamics for a general product domain.

Lemma 7.A.7. *Let π be a distribution over $\{\pm 1\}^n$, and $\pi = \mathbb{E}_{\boldsymbol{z} \sim \rho} \pi_{\boldsymbol{z}}$ a measure decomposition of π such that*

- *for all functions f, $\mathbb{E}_{\boldsymbol{z} \sim \rho} \mathsf{Var}_{\pi_{\boldsymbol{z}}}[f] \geq C_{\mathrm{Var}} \mathsf{Var}_\pi[f]$, and*
- *with probability $1 - \eta$ over $\boldsymbol{z} \sim \rho$, $\pi_{\boldsymbol{z}}$ satisfies a $(\rho_{\mathrm{PI}}, \delta)$-weak Poincaré inequality with respect to Glauber.*

Then, π satisfies a $\left(\rho_{\mathrm{PI}} C_{\mathrm{Var}}, \frac{\delta + \eta}{C_{\mathrm{Var}}}\right)$-weak Poincaré inequality.

Proof. The proof is the same as that of Lemma 7.4.10, except in the Langevin case we have $\mathcal{E}_\pi(f, f) = \mathbb{E}_{\boldsymbol{z} \sim \rho} \mathcal{E}_{\pi_{\boldsymbol{z}}}(f, f)$, whereas here we apply Fact 7.A.4 to get the desired inequality. □

Finally, we record the following simple observation connecting weak functional inequalities in discrete domains.

Fact 7.A.8. *Let π be a distribution on a finite state space Ω, and set $C_\pi = \frac{1 - 2\pi_{\min}}{\log(1/\pi_{\min} - 1)}$. If π satisfies a $(\rho_{\mathrm{PI}}, \varepsilon)$-weak Poincaré inequality, then π also satisfies a $\left(4\rho_{\mathrm{PI}} C_\pi, \frac{\varepsilon}{C_\pi}\right)$-weak MLSI and a $(\rho_{\mathrm{PI}} C_\pi, \frac{\varepsilon}{C_\pi})$-weak LSI.*

Proof. For finite state spaces, it is well-known that the LSI of the complete graph Markov chain $P_{K(\pi)}$ has $\rho_{\mathrm{LS}} = \frac{1 - 2\pi_{\min}}{\log(1/\pi_{\min} - 1)}$ (see e.g., [DSC96]). Furthermore, observe that $\mathcal{E}_{P_{K(\pi)}}(f, f) = \mathsf{Var}_\pi[f]$. Hence,

$$\begin{aligned}
\frac{1}{\rho_{\mathrm{PI}}} \mathcal{E}(f, f) &\geq \mathsf{Var}_\pi[f] - \varepsilon \cdot \operatorname{osc}(f)^2 && (\text{Weak PI}) \\
&\geq C_\pi \mathsf{Ent}_\pi[f^2] - \varepsilon \cdot \operatorname{osc}(f)^2, && (\text{LSI of } P_{K(\pi)})
\end{aligned}$$

which establishes the weak LSI. For the weak MLSI, one applies the inequality $4\mathcal{E}(f, f) \leq \mathcal{E}(f^2, \log f^2)$, whose proof reduces to checking the two-variable inequality $4(\sqrt{u} - \sqrt{v})^2 \leq (u - v) \log \frac{u}{v}$ for positive u, v. □

7.B Deferred calculations for spherical spin glasses

7.B.1 The TAP Hamiltonian

In this subsection, we will prove Lemma 7.7.26, which we restate for convenience.

Lemma 7.7.26. *The law of Hamiltonian $H_{\mathsf{TAP}} \sim \mu_{\mathsf{TAP},\boldsymbol{x},\boldsymbol{m}}$ is described by a Gaussian process $(H_{\mathsf{TAP}}(\sigma))_{\sigma\in S_N}$ defined by*

$$\mathbb{E}\ H_{\mathsf{TAP}}(\sigma) = N\xi_t(R(x,\sigma)) - \langle \boldsymbol{x}, v(\sigma)\rangle \cdot \xi_t'(q_{\boldsymbol{x}}) - \frac{\xi_t'(R(\boldsymbol{m},\sigma))}{\gamma'(q_{\boldsymbol{m}})} \cdot \langle \boldsymbol{m},\sigma\rangle \cdot \left(\theta'(q_{\boldsymbol{m}}) - \frac{1}{1-q_{\boldsymbol{m}}}\right)$$
$$\frac{1}{N}\mathsf{Cov}(H_{\mathsf{TAP}}(\sigma), H_{\mathsf{TAP}}(\sigma')) = \xi_t(R(\sigma,\sigma')) - R(\sigma,\sigma')\frac{\xi_t'(R(\boldsymbol{m},\sigma))\xi_t'(R(\boldsymbol{m},\sigma'))}{\xi_t'(q_{\boldsymbol{m}})} + \frac{\xi_t''(q_{\boldsymbol{m}})}{\gamma'(q_{\boldsymbol{m}})\xi_t'(q_{\boldsymbol{m}})}\gamma(R(\boldsymbol{m},\sigma))\gamma(R(\boldsymbol{m},\sigma')),$$

where

$$v(\sigma) \coloneqq \frac{\xi_t'(R(\boldsymbol{m},\sigma))}{\xi_t'(q_{\boldsymbol{m}})}\left[I - \frac{\xi_t''(q_{\boldsymbol{m}})}{\gamma'(q_{\boldsymbol{m}})} \cdot \frac{\boldsymbol{m}\boldsymbol{m}^\top}{N}\right]\sigma$$
$$\gamma(q) \coloneqq q \cdot \xi_t'(q)\,.$$

To prove the above, we will need the following formulas for any p-spin Hamiltonian H_N with mixture function ξ.

Fact 7.B.1. *For any $u, v, m \in \mathbb{R}^N$, we have:*

$$\frac{1}{N}\,\mathbb{E}\langle u, \nabla H_N(m)\rangle\langle v, \nabla H_N(m)\rangle = R(u,v)\xi'(R(m,m)) + R(m,u)R(m,v)\xi''(R(m,m))\,.$$

Proof. Once we write the derivative as its definition as a limit, the order of the limit and the expectation operator can be swapped by the dominated convergence theorem.

$$\begin{aligned}\frac{1}{N}\,\mathbb{E}\langle u, \nabla H_N(m)\rangle\langle v, \nabla H_N(m)\rangle &= \frac{1}{N}\,\mathbb{E}\lim_{\delta,\varepsilon\to 0}\frac{H_N(m+\delta u) - H_N(m)}{\delta} \cdot \frac{H_N(m+\varepsilon v) - H_N(m)}{\varepsilon}\\ &= \frac{1}{N}\lim_{\delta,\varepsilon\to 0}\frac{1}{\delta\varepsilon}\,\mathbb{E}(H_N(m+\delta u) - H_N(m)) \cdot (H_N(m+\varepsilon v) - H_N(m))\\ &= \lim_{\delta,\varepsilon\to 0}\frac{1}{\delta\varepsilon}[\xi(R(m+\delta u, m+\varepsilon v)) - \xi(R(m+\delta u, m)) - \xi(R(m, m+\varepsilon v)) + \xi(R(m,m))]\\ &= R(u,v)\xi'(R(m,m)) + R(m,u)R(m,v)\xi''(R(m,m))\,.\end{aligned}$$

□

Fact 7.B.2. *For any $u, v, m \in \mathbb{R}^N$, we have:*

$$\frac{1}{N}\,\mathbb{E}\langle u, \nabla H_N(m)\rangle H_N(v) = R(u,v)\xi'(R(m,v))\,.$$

The proof of the above is analogous to the proof of Fact 7.B.1, and hence omitted.We now prove Lemma 7.7.26.

Proof of Lemma 7.7.26. The distribution of $H_{\mathsf{TAP}}(\sigma)$ is the same as that of $H_{N,t}(\sigma)|\boldsymbol{x}, \nabla\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m}) = 0$. Recall that $H_{N,t}(\sigma) = N\xi_t(R(\boldsymbol{x},\sigma)) + \widetilde{H}(\sigma)$ where $\widetilde{H}(\sigma)$ is a centered Gaussian process. Next, observe that conditioning on $\nabla\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m}) = 0$ is the same as conditioning on

$$\nabla\widetilde{H}(\boldsymbol{m}) = -\boldsymbol{x} \cdot \xi_t'(q_{\boldsymbol{x}}) - \boldsymbol{m} \cdot \left(\theta'(q_{\boldsymbol{m}}) - \frac{1}{1-q_{\boldsymbol{m}}}\right). \tag{7.34}$$

Observe that $(H_{\mathsf{TAP}}(\sigma))_{\sigma\in S_N}$ is a Gaussian process, since it is obtained by conditioning on another Gaussian process satisfying affine constraints. First, observe that we can write

$$H_{\mathsf{TAP}}(\sigma) = N\xi_t(R(\boldsymbol{x},\sigma)) + \widetilde{H}_{\mathsf{TAP}}(\sigma) \tag{7.35}$$

where $\widetilde{H}_{\mathsf{TAP}}(\sigma) = \widetilde{H}(\sigma)|\boldsymbol{x}, \nabla\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m}) = 0$. To understand the behavior of $\widetilde{H}_{\mathsf{TAP}}(\sigma)$, we break $\widetilde{H}(\sigma)$ into a sum of two terms: one term for its projection onto the space $U := \left\{\left\langle \nabla\widetilde{H}(\boldsymbol{m}), u\right\rangle : u \in \mathbb{R}^N\right\}$, and the part that is orthogonal to U, and thus independent of $\nabla\widetilde{H}(\boldsymbol{m})$. Concretely, let us write

$$\widetilde{H}(\sigma) = \left\langle \nabla\widetilde{H}(\boldsymbol{m}), v(\sigma)\right\rangle + \left(\widetilde{H}(\sigma) - \left\langle \nabla\widetilde{H}(\boldsymbol{m}), v(\sigma)\right\rangle\right). \tag{7.36}$$

This is true for any $v(\sigma)$, but we have set up the definition such that the second summand is independent of $\nabla\widetilde{H}(\boldsymbol{m})$. To verify this, since these two random variables are each mean 0, it suffices to check that for any $u \in S_N$,

$$\mathbb{E}\left\langle \nabla\widetilde{H}(\boldsymbol{m}), u\right\rangle \left(\widetilde{H}(\sigma) - \left\langle \nabla\widetilde{H}(\boldsymbol{m}), v(\sigma)\right\rangle\right) = 0\,.$$

By Facts 7.B.1 and 7.B.2, the left-hand-side of the above is:

$$R(u,\sigma)\xi_t'(R(\boldsymbol{m},\sigma)) - R(u, v(\sigma))\xi_t'(q_{\boldsymbol{m}}) - R(\boldsymbol{m}, u)R(\boldsymbol{m}, v(\sigma))\xi''(q_{\boldsymbol{m}})\,.$$

We would like $v(\sigma)$ to be such that this is 0 for *all* u. Setting u orthogonal to $\boldsymbol{m}$ and σ shows that we must have $v(\sigma)$ in the subspace spanned by $\boldsymbol{m}$ and σ.

Suppose that $v(\sigma) = \alpha\sigma + \beta\boldsymbol{m}$. Then, plugging this into the above requires that

$$\begin{aligned}
0 &= R(\sigma,u)\xi_t'(R(\boldsymbol{m},\sigma)) - (\alpha R(\sigma,u) + \beta R(\boldsymbol{m},u))\,\xi_t'(q_{\boldsymbol{m}}) - R(\boldsymbol{m},u)\,(\alpha R(\boldsymbol{m},\sigma) + \beta q_{\boldsymbol{m}})\,\xi_t''(q_{\boldsymbol{m}}) \\
&= R(\sigma,u)\,(\xi_t'(R(\boldsymbol{m},\sigma)) - \alpha\xi_t'(q_{\boldsymbol{m}})) \\
&\quad - R(\boldsymbol{m},u)\,(\beta\xi_t'(q_{\boldsymbol{m}}) - \alpha R(\boldsymbol{m},\sigma)\xi_t''(q_{\boldsymbol{m}}) - \beta q_{\boldsymbol{m}}\xi_t''(q_{\boldsymbol{m}}))\,.
\end{aligned}$$

Since this is true for all u, each of these two terms must be 0. That is,

$$\alpha = \frac{\xi_t'(R(\boldsymbol{m},\sigma))}{\xi_t'(q_{\boldsymbol{m}})}$$

and

$$\beta = -\alpha\cdot\frac{R(\boldsymbol{m},\sigma)\xi_t''(q_{\boldsymbol{m}})}{\xi_t'(q_{\boldsymbol{m}}) + q_{\boldsymbol{m}}\xi_t''(q_{\boldsymbol{m}})},$$

so

$$\begin{aligned}
v(\sigma) &= \frac{\xi_t'(R(\boldsymbol{m},\sigma))}{\xi_t'(q_{\boldsymbol{m}})}\left(\sigma - \boldsymbol{m}\cdot\frac{R(\boldsymbol{m},\sigma)\xi_t''(q_{\boldsymbol{m}})}{\xi_t'(q_{\boldsymbol{m}}) + q_{\boldsymbol{m}}\xi_t''(q_{\boldsymbol{m}})}\right) \\
&= \frac{\xi_t'(R(\boldsymbol{m},\sigma))}{\xi_t'(q_{\boldsymbol{m}})}\left(\mathrm{Id} - \frac{R(\boldsymbol{m},\sigma)\xi_t''(q_{\boldsymbol{m}})}{\xi_t'(q_{\boldsymbol{m}}) + q_{\boldsymbol{m}}\xi_t''(q_{\boldsymbol{m}})}\cdot\frac{\boldsymbol{m}\boldsymbol{m}^\top}{N}\right)\sigma
\end{aligned}$$

as defined.

Now returning to (7.36), when we condition on $\boldsymbol{x}$ and $\nabla\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m}) = 0$, by plugging in (7.34), we get

$$\begin{aligned}
\widetilde{H}_{\mathsf{TAP}}(\sigma) = -\langle \boldsymbol{x}, v(\sigma)\rangle\cdot\xi_t'(q_{\boldsymbol{x}}) - \langle \boldsymbol{m}, v(\sigma)\rangle\cdot\left(\theta'(q_{\boldsymbol{m}}) - \frac{1}{1-q_{\boldsymbol{m}}}\right) \\
+ \left(\widetilde{H}(\sigma) - \left\langle \nabla\widetilde{H}(\boldsymbol{m}), v(\sigma)\right\rangle\right)
\end{aligned} \tag{7.37}$$

We use $\widehat{H}(\sigma)$ to denote the random variable $\widetilde{H}(\sigma) - \left\langle \nabla\widetilde{H}(\boldsymbol{m}), v(\sigma)\right\rangle$, whose distribution remains unaffected by the conditioning, as this random variable is independent of $\boldsymbol{x}$ and the event $\nabla\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m}) = 0$. Since $\widehat{H}(\sigma)$ is centered, our expression for $\mathbb{E}\,H_{\mathsf{TAP}}(\sigma)$ follows from (7.35) and (7.37), and the observation that

$$R(\boldsymbol{m}, v(\sigma)) = \frac{\xi_t'(R(\boldsymbol{m},\sigma))}{\gamma'(q_{\boldsymbol{m}})}\cdot R(\boldsymbol{m},\sigma).$$

It remains to compute $N^{-1}\mathsf{Cov}(H_{\mathsf{TAP}}(\sigma), H_{\mathsf{TAP}}(\sigma'))$ for any $\sigma, \sigma' \in S_N$. Observe that this is equal to $N^{-1}\,\mathbb{E}\,\widehat{H}(\sigma)\widehat{H}(\sigma')$. By Facts 7.B.1 and 7.B.2, we have that this is equal to:

$$\begin{aligned}
\xi_t(R(\sigma,\sigma')) - R(v(\sigma),\sigma')\xi_t'(R(\boldsymbol{m},\sigma')) - R(v(\sigma'),\sigma)\xi_t'(R(\boldsymbol{m},\sigma)) \\
+ R(v(\sigma), v(\sigma'))\xi_t'(q_{\boldsymbol{m}}) + R(\boldsymbol{m}, v(\sigma))R(\boldsymbol{m}, v(\sigma'))\xi''(q_{\boldsymbol{m}})\,.
\end{aligned}$$

The formula for the covariance can be obtained from the above by expanding $v(\sigma)$. $\square$

Next, we look at the mixture function of these "TAP planted distributions on slices".

Corollary 7.7.29. *For a fixed choice of a and b, the Gaussian process $(H_{\mathsf{TAP}}(v(a,b)+r_{a,b}Q\tau))_{\tau\in S_{N-2}}$ is described by the following law.*

- *Let $H_{a,b}$ be a spherical p-spin Hamiltonian with mixture function $\xi_{a,b}$ given by:*
$$\xi_{a,b}(s) \coloneqq \xi_t\Big(\|v(a,b)\|^2 + r_{a,b}^2 s\Big) - \xi_t\Big(\|v(a,b)\|^2\Big) - s\cdot\frac{r_{a,b}^2\xi_t'\Big(q_{\boldsymbol{m}}\cdot\Big(1+\frac{a}{\sqrt{N}}\Big)\Big)^2}{\xi_t'(q_{\boldsymbol{m}})}.$$
- *Let $V(a,b)\coloneqq \xi_t\Big(\|v(a,b)\|^2\Big) - \|v(a,b)\|^2\cdot\frac{\xi_t'\left(\left(1+\frac{a}{\sqrt{N}}\right)q_{\boldsymbol{m}}\right)^2}{\xi_t'(q_{\boldsymbol{m}})} + \frac{\xi_t''(q_{\boldsymbol{m}})}{\gamma'(q_{\boldsymbol{m}})\xi_t'(q_{\boldsymbol{m}})}\cdot\gamma\left(\left(1+\frac{a}{\sqrt{N}}\right)q_{\boldsymbol{m}}\right)^2$.*

The law of $H_{\mathsf{TAP}}(v(a,b)+r_{a,b}Q\tau)$ is the same as that of $H_{a,b}(\tau)+\sqrt{N}\cdot g_{a,b}+\mathbb{E}_{\mu_{\mathsf{TAP}}} H_{\mathsf{TAP}}(v(a,b)+r_{a,b}Q\tau)$ where $g_{a,b}$ is a centered Gaussian of variance $V(a,b)$ independent of $H_{a,b}$.

Proof. Let $\tau,\tau'\in S_{N-2}$, and
$$\sigma = v(a,b)+r_{a,b}Q\tau \text{ and } \sigma' = v(a,b)+r_{a,b}Q\tau'.$$
Recall from Lemma 7.7.26 that
$$\begin{aligned}&N^{-1}\mathsf{Cov}\left(H_{\mathsf{TAP}}(\sigma),H_{\mathsf{TAP}}(\sigma')\right)\\ &\quad= \xi_t'(R(\sigma,\sigma')) - R(\sigma,\sigma')\frac{\xi_t'(R(\boldsymbol{m},\sigma))\xi_t'(R(\boldsymbol{m},\sigma')}{\xi_t'(q_{\boldsymbol{m}})} + \frac{\xi_t''(q_{\boldsymbol{m}})}{\gamma'(q_{\boldsymbol{m}})\xi_t'(q_{\boldsymbol{m}})}\gamma(R(\boldsymbol{m},\sigma))\gamma(R(\boldsymbol{m},\sigma')).\end{aligned}$$
By the definition of σ and σ', we have $R(\boldsymbol{m},\sigma)=R(\boldsymbol{m},\sigma')=R(\boldsymbol{m},v(a,b))=\left(1+\frac{a}{\sqrt{N}}\right)q_{\boldsymbol{m}}$, and $R(\sigma,\sigma')=R\left(\|v(a,b)\|^2+r_{a,b}^2R(\tau,\tau')\right)$, since Q is an isometry. As a result,
$$\begin{aligned}&N^{-1}\mathsf{Cov}\left(H_{\mathsf{TAP}}(\sigma),H_{\mathsf{TAP}}(\sigma')\right)\\ &\quad= \xi_t'(\|v(a,b)\|^2+r_{a,b}^2R(\tau,\tau')) - \left(\|v(a,b)\|^2+r_{a,b}^2R(\tau,\tau')\right)\frac{\xi_t'\left(\left(1+\frac{a}{\sqrt{N}}\right)q_{\boldsymbol{m}}\right)^2}{\xi_t'(q_{\boldsymbol{m}})}\\ &\qquad+\frac{\xi_t''(q_{\boldsymbol{m}})}{\gamma'(q_{\boldsymbol{m}})\xi_t'(q_{\boldsymbol{m}})}\gamma\left(\left(1+\frac{a}{\sqrt{N}}\right)q_{\boldsymbol{m}}\right)^2.\end{aligned}$$
This may be written as
$$N^{-1}\mathsf{Cov}\left(H_{\mathsf{TAP}}(\sigma),H_{\mathsf{TAP}}(\sigma')\right) = \xi_{a,b}(R(\tau,\tau'))+V(a,b).$$
This implies that $H_{\mathsf{TAP}}(\sigma)$ is equal to $H_{a,b}(\tau)+g_{a,b}$ for some Gaussian process $(H_{a,b}(\tau))_{\tau\in S_{N-2}}$, where $g_{a,b}$ is a centered Gaussian of variance $V(a,b)$. To complete the proof, we must show that the correlation structure of $H_{a,b}$ can be achieved by a p-spin model with mixture function $\xi_{a,b}$. To do this, it suffices to show that $\xi_{a,b}$ is indeed a valid mixture function, in that $\xi_{a,b}^{(p)}(0)\ge 0$ for all $p\ge 1$, and $\xi_{a,b}(0)=0$. The latter of these is clearly true by construction. The former is easily seen to be true for $p\ge 2$, since for such p,
$$\xi_{a,b}^{(p)}(0) = r_{a,b}^{2p}\xi_t^{(p)}(\|v(a,b)\|^2)\ge 0$$
since ξ_t is a valid mixture function. For $p=1$,
$$\begin{aligned}\xi_{a,b}'(0) &= r_{a,b}^2\cdot\left(\xi_t'\left(\|v(a,b)\|^2\right) - \frac{\xi_t'\left(q_{\boldsymbol{m}}\left(1+\frac{a}{\sqrt{N}}\right)\right)^2}{\xi_t'(q_{\boldsymbol{m}})}\right)\\ &\overset{(7.9)}{\ge} r_{a,b}^2\cdot\left(\xi_t'\left(q_{\boldsymbol{m}}\left(1+\frac{a}{\sqrt{N}}\right)^2\right) - \frac{\xi_t'\left(q_{\boldsymbol{m}}\left(1+\frac{a}{\sqrt{N}}\right)\right)^2}{\xi_t'(q_{\boldsymbol{m}})}\right)\\ &= \frac{r_{a,b}^2}{\xi_t'(q_{\boldsymbol{m}})}\cdot\left(\xi_t'\left(q_{\boldsymbol{m}}\left(1+\frac{a}{\sqrt{N}}\right)^2\right)\xi_t'(q_{\boldsymbol{m}}) - \xi_t'\left(q_{\boldsymbol{m}}\left(1+\frac{a}{\sqrt{N}}\right)\right)^2\right)\ge 0.\end{aligned}$$

In the first inequality above, we use the fact that ξ_t' is non-decreasing. The final inequality is an application of Cauchy-Schwarz. $\square$

7.B.2 Understanding concentration around the codimension-2 slice

Next, we bound the variance of $g_{a,b} - g_{0,0}$.

Lemma 7.7.36. *For every constant $\iota > 0$, there is a constant c such that with probability $1 - e^{-cN}$, for all a, b, we have $|g_{a,b} - g_{0,0}| \le \iota \frac{a^2+b^2}{\sqrt{N}}$.*

Proof of Lemma 7.7.36. The strategy is to prove that for any $a, b, a', b' \in \mathbb{R}$ of magnitude $\ll N^{1/4}$, $g_{a,b} - g_{a',b'}$ is a Gaussian of variance $O\left(\cdot \|v(a,b) - v(a',b')\|^4\right) = O\left(\frac{(a-a')^4 + (b-b')^4}{N^2}\right)$. The desideratum then immediately follows by applying Slepian's lemma on $(|g_{a,b} - g_{0,0}|)_{a,b}$ comparing it to the Gaussian process $\langle G, (v(a,b) - v(0,0))(v(a,b) - v(0,0))^\top \rangle$ for a standard Gaussian matrix G.

We carry out the calculation for $a', b' = 0$; the general case follows similarly. We have $g_{a,b} = N^{-1/2}\Big(H_{\mathsf{TAP}}(\sqrt{N} \cdot v(a,b)) - \mathbb{E}\, H_{\mathsf{TAP}}(\sqrt{N} \cdot v(a,b))\Big)$, and $g_{0,0} = N^{-1/2}(H_{\mathsf{TAP}}(\boldsymbol{m}) - \mathbb{E}\, H_{\mathsf{TAP}}(\boldsymbol{m}))$. Clearly, $g_{a,b} - g_{0,0}$ is a centered Gaussian process. As in the proof of Lemma 7.7.26, we have that the distribution of $H_{\mathsf{TAP}}(\sigma) - \mathbb{E}\, H_{\mathsf{TAP}}(\sigma)$ is the same as that of

$$\widetilde{H}(\sigma) - \left\langle \nabla \widetilde{H}(\boldsymbol{m}), v(\sigma) \right\rangle,$$

where $\widetilde{H}$ is a Hamiltonian distributed according to the mixture function ξ_t. Thus, the distribution of $g_{a,b} - g_{0,0}$ is:

$$\widetilde{H}(u(a,b)) - \left\langle \nabla \widetilde{H}(\boldsymbol{m}), v(u(a,b)) \right\rangle - \widetilde{H}(\boldsymbol{m}) + \left\langle \nabla \widetilde{H}(\boldsymbol{m}), v(\boldsymbol{m}) \right\rangle.$$

For brevity, we denote $v(a,b)$ as $\boldsymbol{m} + \varepsilon$. We express $\widetilde{H}(\boldsymbol{m} + \varepsilon)$ in its Taylor expansion, and we get:

$$\sqrt{N}(g_{a,b} - g_{0,0}) = \sum_{i \ge 1} \frac{1}{i!} \left\langle \mathrm{D}_i \widetilde{H}(\boldsymbol{m}), \varepsilon^{\otimes i} \right\rangle - \left\langle \nabla \widetilde{H}(\boldsymbol{m}), v(\boldsymbol{m} + \varepsilon) - v(\boldsymbol{m}) \right\rangle. \tag{7.38}$$

Expanding out $v(\boldsymbol{m} + \varepsilon) - v(\boldsymbol{m})$ ultimately yields:

$$v(\boldsymbol{m} + \varepsilon) - v(\boldsymbol{m}) = \varepsilon + R(\boldsymbol{m}, \varepsilon)\varepsilon \frac{\xi_t''(q_{\boldsymbol{m}})}{\gamma'(q_{\boldsymbol{m}})} - \frac{R(\boldsymbol{m}, \varepsilon)^2 \xi_t''(q_{\boldsymbol{m}})^2}{\xi_t'(q_{\boldsymbol{m}})\gamma'(q_{\boldsymbol{m}})} \boldsymbol{m}.$$

Plugging in the above into (7.38) gives:

$$\begin{aligned} \sqrt{N}(g_{a,b} - g_{0,0}) = \sum_{i \ge 2} \frac{1}{i!} \left\langle \mathrm{D}_i \widetilde{H}(\boldsymbol{m}), \varepsilon^{\otimes i} \right\rangle \\ - \left\langle \nabla \widetilde{H}(\boldsymbol{m}), \varepsilon \right\rangle R(\boldsymbol{m}, \varepsilon) \frac{\xi_t''(q_{\boldsymbol{m}})}{\gamma'(q_{\boldsymbol{m}})} - \left\langle \nabla \widetilde{H}(\boldsymbol{m}), \boldsymbol{m} \right\rangle \frac{R(\boldsymbol{m}, \varepsilon)^2 \xi_t''(q_{\boldsymbol{m}})^2}{\xi_t'(q_{\boldsymbol{m}})\gamma'(q_{\boldsymbol{m}})} \end{aligned}$$

We have an explicit expression for ε:

$$\varepsilon = \frac{a q_{\boldsymbol{m}} - b q_{\boldsymbol{x}}^2}{\sqrt{N}(q_{\boldsymbol{m}} - q_{\boldsymbol{x}}^2)} \boldsymbol{m} + \frac{q_{\boldsymbol{m}} q_{\boldsymbol{x}}}{q_{\boldsymbol{m}} - q_{\boldsymbol{x}}^2} \left(\frac{b-a}{\sqrt{N}} \right) \boldsymbol{x}.$$

This explicit expression can be used to obtain the following bounds on the variances of the above terms:

$$\begin{aligned} \mathsf{Var}\left[\frac{1}{i!} \left\langle \mathrm{D}_i \widetilde{H}(\boldsymbol{m}), \varepsilon^{\otimes i} \right\rangle \right] &\le \frac{O\left(a^{2i} + b^{2i}\right)}{N^{i-1}} \\ \mathsf{Var}\left[\left\langle \nabla \widetilde{H}(\boldsymbol{m}), \varepsilon \right\rangle R(\boldsymbol{m}, \varepsilon) \frac{\xi_t''(q_{\boldsymbol{m}})}{\gamma'(q_{\boldsymbol{m}})} \right] &= \frac{O\left(a^4 + b^4\right)}{N} \\ \mathsf{Var}\left[\left\langle \nabla \widetilde{H}(\boldsymbol{m}), \boldsymbol{m} \right\rangle \frac{R(\boldsymbol{m}, \varepsilon)^2 \xi_t''(q_{\boldsymbol{m}})}{\xi_t'(q_{\boldsymbol{m}})\gamma'(q_{\boldsymbol{m}})} \right] &\le \frac{O\left(a^4 + b^4\right)}{N} \end{aligned}$$

The expression for $\sqrt{N}(g_{a,b} - g_{0,0})$ only involves a constant number of terms, and since the first term enumerates over $i \geq 2$, and since $|a|, |b| \leq \sqrt{N}$, we have an overall bound of $\frac{O(a^4+b^4)}{N}$. Dividing by $\sqrt{N}$ gives the desired variance bound. $\square$

Lemma 7.7.34. $\nabla \widehat{E}_{a,b}\Big|_{(a,b)=(0,0)} = 0.$

Proof. Recall

$$\widehat{E}_{a,b} = \frac{1}{2}\underbrace{\left(\log r_{a,b}^2 - \xi_t(\|v(a,b)\|^2) - r_{a,b}^2 \cdot \frac{\xi_t'\left(q_{\boldsymbol{m}}\left(1+\frac{a}{\sqrt{N}}\right)\right)^2}{\xi_t'(q_{\boldsymbol{m}})}\right)}_{\text{(I)}}$$

$$+ \underbrace{\xi_t\left(q_{\boldsymbol{x}}\left(1+\frac{b}{\sqrt{N}}\right)\right) + \frac{\gamma\left(q_{\boldsymbol{m}}\left(1+\frac{a}{\sqrt{N}}\right)\right)}{\gamma'(q_{\boldsymbol{m}})} \cdot \left((1-q_{\boldsymbol{m}})\xi_t''(q_{\boldsymbol{m}}) + \frac{1}{1-q_{\boldsymbol{m}}}\right)}_{\text{(II)}}$$

$$- \frac{\gamma(q_{\boldsymbol{x}})}{\xi_t'(q_{\boldsymbol{m}})} \cdot \underbrace{\xi_t'\left(q_{\boldsymbol{m}}\left(1+\frac{a}{\sqrt{N}}\right)\right) \cdot \left(\left(1+\frac{b}{\sqrt{N}}\right) - q_{\boldsymbol{m}} \cdot \frac{\xi_t''(q_{\boldsymbol{m}})}{\gamma'(q_{\boldsymbol{m}})} \cdot \left(1+\frac{a}{\sqrt{N}}\right)\right)}_{\text{(III)}}.$$

Because

$$\|v(a,b)\|^2 = q_{\boldsymbol{m}}\left(1+\frac{a}{\sqrt{N}}\right)^2 + \frac{q_{\boldsymbol{m}}q_{\boldsymbol{x}}}{q_{\boldsymbol{m}} - q_{\boldsymbol{x}}^2} \cdot \left(\frac{a-b}{\sqrt{N}}\right)^2,$$

we have

$$\nabla\|v(a,b)\|^2\big|_{(a,b)=(0,0)} = -\nabla r_{a,b}^2\big|_{(a,b)=(0,0)} = \left(\frac{2q_{\boldsymbol{m}}}{\sqrt{N}}, 0\right).$$

We also have $r_{0,0}^2 = 1 - q_{\boldsymbol{m}}$ and $\|v(0,0)\|^2 = q_{\boldsymbol{m}}$. Let us start by computing the derivative with respect to a. We have

$$\begin{aligned}\sqrt{N} \cdot \partial_a(\text{III})|_{(a,b)=(0,0)} &= \xi_t''(q_{\boldsymbol{m}}) \cdot q_{\boldsymbol{m}} \cdot \left(1 - \frac{q_{\boldsymbol{m}}\xi_t''(q_{\boldsymbol{m}})}{\gamma'(q_{\boldsymbol{m}})}\right) + \xi_t'(q_{\boldsymbol{m}}) \cdot \left(-q_{\boldsymbol{m}} \cdot \frac{\xi_t''(q_{\boldsymbol{m}})}{\gamma'(q_{\boldsymbol{m}})}\right) \\ &= \frac{q_{\boldsymbol{m}}\xi_t''(q_{\boldsymbol{m}})}{\gamma'(q_{\boldsymbol{m}})}\left(\gamma'(q_{\boldsymbol{m}}) - q_{\boldsymbol{m}}\xi_t''(q_{\boldsymbol{m}}) - \xi_t'(q_{\boldsymbol{m}})\right) = 0.\end{aligned}$$

Next,

$$\begin{aligned}&\sqrt{N} \cdot \partial_a(\text{I})|_{(a,b)=(0,0)} \\ &\quad= \frac{1}{r_{0,0}^2} \cdot (-2q_{\boldsymbol{m}}) - \xi_t'(\|v(0,0)\|^2) \cdot (2q_{\boldsymbol{m}}) - (-2q_{\boldsymbol{m}}) \cdot \left(\frac{\xi_t'(q_{\boldsymbol{m}})^2}{\xi_t'(q_{\boldsymbol{m}})}\right) - r_{0,0}^2 \cdot \frac{2\xi_t'(q_{\boldsymbol{m}}) \cdot \xi_t''(q_{\boldsymbol{m}}) \cdot q_{\boldsymbol{m}}}{\xi_t'(q_{\boldsymbol{m}})} \\ &\quad= \frac{-2q_{\boldsymbol{m}}}{1-q_{\boldsymbol{m}}} - 2q_{\boldsymbol{m}}\xi_t'(q_{\boldsymbol{m}}) + 2q_{\boldsymbol{m}}\xi_t'(q_{\boldsymbol{m}}) - 2q_{\boldsymbol{m}}(1-q_{\boldsymbol{m}})\xi_t''(q_{\boldsymbol{m}}) \\ &\quad= \frac{-2q_{\boldsymbol{m}}}{1-q_{\boldsymbol{m}}} - 2q_{\boldsymbol{m}}(1-q_{\boldsymbol{m}})\xi_t''(q_{\boldsymbol{m}}).\end{aligned}$$

Finally,

$$\begin{aligned}&\sqrt{N} \cdot \partial_a(\text{II})|_{(a,b)=(0,0)} \\ &\quad= \frac{\gamma'(q_{\boldsymbol{m}}) \cdot q_{\boldsymbol{m}}}{\gamma'(q_{\boldsymbol{m}})} \cdot \left((1-q_{\boldsymbol{m}})\xi_t''(q_{\boldsymbol{m}}) + \frac{1}{1-q_{\boldsymbol{m}}}\right) \\ &= -\frac{1}{2} \cdot \sqrt{N} \cdot \partial_a(\text{I})|_{(a,b)=(0,0)}\end{aligned}$$

as desired. The derivative with respect to b is much simpler, since the derivative of $r_{a,b}^2$ with respect to b is 0 at $(0,0)$. Consequently, $\partial_b(\mathrm{I})|_{(a,b)=(0,0)} = 0$, $\partial_b(\mathrm{II})|_{(a,b)=(0,0)} = q_{\boldsymbol{x}}\xi_t'(q_{\boldsymbol{x}}) = \gamma(q_{\boldsymbol{x}})$, and $\partial_b(\mathrm{III})|_{(a,b)=(0,0)} = \xi_t'(q_{\boldsymbol{m}})$, completing the proof. □

Lemma 7.7.35. *There exist constants $\eta, \varepsilon > 0$ such that for all $|a|, |b| \le \varepsilon\sqrt{N}$, $N\nabla^2 \widehat{E}_{a,b} \preceq -\eta \mathrm{Id}$.*

Proof. For ease of notation, define $\widetilde{E}_{a,b} = \widehat{E}_{\sqrt{N}a,\sqrt{N}b}$, $\widetilde{r}_{a,b} = r_{\sqrt{N}a,\sqrt{N}b}$, and $\widetilde{v}(a,b) = v(\sqrt{N}a, \sqrt{N}b)$. As in the previous lemma, recall

$$\begin{aligned}\widetilde{E}_{a,b} = \frac{1}{2}&\left(\log \widetilde{r}_{a,b}^2 - \xi_t(\|\widetilde{v}(a,b)\|^2) - \widetilde{r}_{a,b}^2 \cdot \frac{\xi_t'\left(q_{\boldsymbol{m}}\left(1+a\right)\right)^2}{\xi_t'(q_{\boldsymbol{m}})}\right) \\ &+ \xi_t\left(q_{\boldsymbol{x}}\left(1+b\right)\right) + \frac{\gamma\left(q_{\boldsymbol{m}}\left(1+a\right)\right)}{\gamma'(q_{\boldsymbol{m}})} \cdot \left((1-q_{\boldsymbol{m}})\xi_t''(q_{\boldsymbol{m}}) + \frac{1}{1-q_{\boldsymbol{m}}}\right) \\ &- \frac{\gamma(q_{\boldsymbol{x}})}{\xi_t'(q_{\boldsymbol{m}})} \cdot \xi_t'\left(q_{\boldsymbol{m}}\left(1+a\right)\right) \cdot \left((1+b) - q_{\boldsymbol{m}} \cdot \frac{\xi_t''(q_{\boldsymbol{m}})}{\gamma'(q_{\boldsymbol{m}})} \cdot (1+a)\right).\end{aligned}$$

Because the Hessian is Lipschitz in all the parameters involved, it suffices to prove the negative definiteness of the Hessian at $(0,0)$, under the assumption that $q_{\boldsymbol{m}} = q_{\boldsymbol{x}} = q$, where q (formerly denoted $q_*(t)$) satisfies $\xi_t'(q) = \frac{q}{1-q}$. Under these constraints, we have $\gamma(q) = q\xi_t'(q) = \frac{q^2}{1-q}$ and $\gamma'(q) = \xi_t'(q) + q\xi_t''(q) = \frac{q}{1-q}\left(1 + (1-q)\xi_t''(q)\right)$. $\widetilde{E}_{a,b}$ simplifies as

$$\begin{aligned}\widetilde{E}_{a,b} = &\underbrace{\frac{1}{2}\left(\log \widetilde{r}_{a,b}^2 - \xi_t(\|\widetilde{v}(a,b)\|^2) - \widetilde{r}_{a,b}^2 \cdot \frac{\xi_t'\left(q\left(1+a\right)\right)^2}{\xi_t'(q)}\right)}_{(\mathrm{I})} \\ &\underbrace{+ \xi_t\left(q\left(1+b\right)\right) + \frac{\gamma\left(q\left(1+a\right)\right)}{q} \cdot \frac{1 + (1-q)^2\xi_t''(q)}{1 + (1-q)\xi_t''(q)}}_{(\mathrm{II})} \\ &\underbrace{- q\left(1+b\right) \cdot \xi_t'\left(q\left(1+a\right)\right) + q^2\left(1+a\right) \cdot \frac{\xi_t''(q)}{\gamma'(q)}}_{(\mathrm{III})}.\end{aligned}$$

We have $\partial_b^2(\mathrm{III})\big|_{(0,0)} = 0$, and

$$\partial_b^2(\mathrm{II})\big|_{(0,0)} = \xi_t''(q) \cdot q^2.$$

We have that $\partial_b \widetilde{r}_{a,b}^2\Big|_{(0,0)} = 0$, and $\partial_b^2 \widetilde{r}_{a,b}^2\Big|_{(0,0)} = \frac{-2q^2}{1-q}$. Consequently,

$$\begin{aligned}\partial_b^2(\mathrm{I})\big|_{(0,0)} &= \frac{1}{2}\left(\frac{1}{r_{0,0}^2} \cdot \partial_b^2 \widetilde{r}_{a,b}^2\big|_{(0,0)} - \xi_t'(\|v(0,0)\|^2) \cdot \partial_b^2 \|\widetilde{v}(a,b)\|^2\big|_{(0,0)} - \partial_b^2 \widetilde{r}_{a,b}^2\big|_{(0,0)} \cdot \xi_t'(q)\right) \\ &= \frac{-q^2}{(1-q)^2}.\end{aligned}$$

It follows that

$$\partial_b^2 \widehat{E}_{a,b}\Big|_{(0,0)} = \xi_t''(q) \cdot q^2 - \frac{q^2}{(1-q)^2}.$$

Similarly, we have $\partial_a\partial_b(\mathrm{II})|_{(0,0)} = 0$, and

$$\partial_a\partial_b(\mathrm{III})|_{(0,0)} = \xi_t''(q) \cdot q^2.$$

We have that $\partial_a \partial_b \widetilde{r}^2_{a,b}\Big|_{(0,0)} = \frac{-2q^2}{1-q}$. Consequently,

$$\begin{aligned}
\partial_a \partial_b (\mathrm{I})\big|_{(0,0)} &= \frac{1}{2}\left(\frac{1}{r^2_{0,0}} \cdot \partial_a \partial_b \widetilde{r}^2_{a,b}\big|_{(0,0)} - \xi'_t(\|v(0,0)\|^2) \cdot \partial_a \partial_b \|\widetilde{v}(a,b)\|^2\big|_{(0,0)} - \partial_a \partial_b \widetilde{r}^2_{a,b}\big|_{(0,0)} \cdot \xi'_t(q)\right) \\
&= \frac{-q^2}{(1-q)^2}.
\end{aligned}$$

It follows that

$$\partial_a \partial_b \widehat{E}_{a,b}\Big|_{(0,0)} = \xi''_t(q) \cdot q^2 - \frac{q^2}{(1-q)^2}.$$

It remains to compute the second derivative with respect to a. We have

$$\partial_a^2 (\mathrm{III})\big|_{(0,0)} = q^3 \xi'''_t(q).$$

We also have

$$\begin{aligned}
\partial_a^2 (\mathrm{II})\big|_{(0,0)} &= q\gamma''(q) \cdot \frac{1+(1-q)^2 \xi''_t(q)}{1+(1-q)\xi''_t(q)} \\
&= q \cdot \frac{1+(1-q)^2 \xi''_t(q)}{1+(1-q)\xi''_t(q)} \cdot (2\xi''_t(q) + q\xi'''_t(q)) \\
&\le q^2 \xi'''_t(q) + 2q\xi''_t(q).
\end{aligned}$$

We have $\partial_a^2 \widetilde{r}^2_{a,b}\Big|_{(0,0)} = -\left(2q + \frac{2q^2}{1-q}\right) = -\frac{2q}{1-q}$ and $\partial_a \widetilde{r}^2_{a,b}\Big|_{(0,0)} = -2q$. Finally,

$$\begin{aligned}
\partial_a^2 (\mathrm{I})\big|_{(0,0)} &= \frac{1}{r^2_{0,0}} \cdot \frac{-2q}{1-q} - \frac{1}{r^4_{0,0}} \cdot (-2q)^2 - \xi'_t(q) \cdot \frac{-2q}{1-q} - \xi''_t(q) \cdot (-2q)^2 - \frac{-2q}{1-q} \cdot \xi'_t(q) \\
&\quad - 2 \cdot (-2q) \cdot \frac{2\xi'_t(q)\xi''_t(q) q}{\xi'_t(q)} - (1-q) \cdot \frac{2q^2(\xi'_t(q)\xi'''_t(q) + \xi''_t(q)^2)}{\xi'_t(q)} \\
&= \frac{-2q}{(1-q)^2} - \frac{4q^2}{(1-q)^2} - 4q^2\xi''_t(q) + 8q^2 \xi''_t(q) - 2q^2(1-q)\xi'''_t(q) - 2q^2(1-q) \cdot \frac{\xi''_t(q)^2}{\xi'_t(q)} \\
&= \frac{-2q(2q+1)}{(1-q)^2} + 4q^2 \xi''_t(q) - 2q^2(1-q)\xi'''_t(q) - 2q(1-q)^2 \xi''_t(q)^2.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\partial_a^2 \widehat{E}_{a,b}\Big|_{(0,0)} &\le -q^3 \xi'''_t(q) + q^2 \xi'''_t(q) + 2q\xi''_t(q) - \frac{q(2q+1)}{(1-q)^2} \\
&\quad + 2q^2 \xi''_t(q) - q^2(1-q)\xi'''_t(q) - q(1-q)^2 \xi''_t(q)^2 \\
&= \frac{-q(2q+1)}{(1-q)^2} + 2q\xi''_t(q) + 2q^2 \xi''_t(q) - q(1-q)^2 \xi''_t(q)^2.
\end{aligned}$$

To conclude, let us check that the Hessian is negative definite. Because $\xi''_t(q) \cdot q^2 - \frac{q^2}{(1-q)^2} < 0$ by the SL condition (SL), it suffices to check that

$$\partial_a^2 \widehat{E}_{a,b}\Big|_{(0,0)} < q^2 \xi''_t(q) - \frac{q^2}{(1-q)^2}.$$

This is true if and only if

$$q(1-q)^2 \xi''_t(q)^2 - q(q+2)\xi''_t(q) + \frac{q(q+1)}{(1-q)^2} > 0.$$

It is not difficult to see that this is true if $\xi''_t(q)$ is less than the smaller root of the above quadratic, which is equal to

$$\frac{(q+2) - \sqrt{(q+2)^2 - 4(q+1)}}{2(1-q)^2} = \frac{1}{(1-q)^2}.$$

This is true by the SL condition (SL), concluding the proof. $\square$

Next, we shall prove Lemma 7.7.38. Recall the definition

$$\mathsf{Error}^{(2)}_{a,b} = \frac{N\xi''_{a,b}(0)}{4} - \frac{1}{2}\log\det\left((1+\xi''_{a,b}(0))\mathrm{Id} - \nabla^2 H_{a,b}(0)\right),$$

where $\nabla^2 H_{a,b}(0)$ is equal to the restriction of $r_{a,b}^2 \cdot \nabla^2 H_{\mathsf{TAP}}(v(a,b))$ restricted to the codimension-2 subspace orthogonal to $\boldsymbol{m}$ and $\boldsymbol{x}$.

Lemma 7.7.38. *For any sufficiently small $\iota > 0$, with probability at least $1 - e^{-cN}$, $\left|\mathsf{Error}^{(2)}_{a,b} - \mathsf{Error}^{(2)}_{0,0}\right| = O(1)$ for all $a, b < \iota N^{1/4}$.*

Let us start by computing the correlation structure of the random matrices $\nabla^2 H_{a,b}(0)$. Note that $\nabla^2 H_{a,b}(0)$ is an $(N-2)$-dimensional GOE matrix scaled by $\sqrt{\xi''_{a,b}(0)}$.

Fact 7.B.3. *For σ^1, σ^2,*

$$\begin{aligned}\frac{1}{N}\,\mathbb{E}\langle \nabla^2 H_{\mathsf{TAP}}(\sigma^1), u^1 \otimes u^2\rangle\langle \nabla^2 H_{\mathsf{TAP}}(\sigma^2), v^1 \otimes v^2\rangle& \\ = \xi''_t\left(R(\sigma^1,\sigma^2)\right)\cdot\left(R(u^1,v^2)\cdot R(u^2,v^1) + R(u^1,v^1)\cdot R(u^2,v^2)\right)&.\end{aligned}$$

In particular,

$$\frac{1}{N}\,\mathbb{E}\langle \nabla^2 H_{\mathsf{TAP}}(\sigma^1), \nabla^2 H_{\mathsf{TAP}}(\sigma^2)\rangle = \xi''_t\left(R(\sigma^1,\sigma^2)\right).$$

The above follows from calculations similar to those involved in the proofs of Facts 7.B.1 and 7.B.2; we omit the details.

Proof of Lemma 7.7.38. We shall prove the statement for a fixed a, b; a union bound over a, b implies the boundedness for all $a, b \le \iota N^{1/4}$.

Recalling that $\nabla^2 H_{a,b}(0) = r_{a,b}^2 \nabla^2 H_{\mathsf{TAP}}(\sqrt{N}\cdot v(a,b))$ is a GOE matrix scaled by $\sqrt{\xi''_{a,b}(0)}$. By Fact 7.B.3,

$$\begin{aligned}\frac{1}{N}\,\mathbb{E}\langle \nabla^2 H_{a,b}(0), \nabla^2 H_{0,0}(0)\rangle &= r_{a,b}^2\cdot r_{0,0}^2\cdot\frac{1}{N}\,\mathbb{E}\langle \nabla^2 H_{\mathsf{TAP}}(\sqrt{N}\cdot v(a,b)), \nabla^2 H_{\mathsf{TAP}}(\sqrt{N}\cdot v(0,0))\rangle \\ &= r_{a,b}^2\cdot r_{0,0}^2\cdot\xi''_t\left(\langle v(a,b), v(0,0)\rangle\right).\end{aligned}$$

For comparison, we have

$$\frac{1}{N}\,\mathbb{E}\left\|\nabla^2 H_{a,b}(0)\right\|_F^2 = \xi''_{a,b}(0) = r_{a,b}^4\cdot\xi''_t\left(\|v(a,b)\|^2\right).$$

For succinctness of notation, let

$$\begin{pmatrix}\alpha_1 & \rho \\ \rho & \alpha_2\end{pmatrix} = \begin{pmatrix} r_{0,0}^4\xi''_t(\|v(0,0)\|^2) & r_{a,b}^2 r_{0,0}^2\xi''_t\left(\langle v(a,b), v(0,0)\rangle\right) \\ r_{a,b}^2 r_{0,0}^2\xi''_t\left(\langle v(a,b), v(0,0)\rangle\right) & r_{a,b}^4\xi''_t(\|v(a,b)\|^2)\end{pmatrix}$$

be the covariance structure of the scaled GOE matrices $\nabla^2 H_{0,0}(0)$ and $\nabla^2 H_{a,b}(0)$. It is not difficult to see that $\alpha_2 = \alpha_1 + O\left(\frac{a^2+b^2}{N}\right)$, and ρ is between α_1 and α_2. Also note that $\alpha_1 = \xi''_{0,0}(0)$ and $\alpha_2 = \xi''_{a,b}(0)$. Then, for some choice of GOE matrices G and $\widetilde{G}$, we may write

$$\begin{aligned}\nabla^2 H_{0,0}(0) &= \sqrt{\alpha_1}G \\ \nabla^2 H_{a,b}(0) &= \frac{\rho}{\sqrt{\alpha_1}}G + \sqrt{\alpha_2 - \frac{\rho^2}{\alpha_1}}\widetilde{G}.\end{aligned}$$

We thus have

$$2\left(\mathsf{Error}_{0,0}^{(2)} - \mathsf{Error}_{a,b}^{(2)}\right) = \frac{N\alpha_1}{2} - \frac{N\alpha_2}{2} - \log\det\left(\underbrace{(1+\alpha_1)\,\mathrm{Id} - \sqrt{\alpha_1}G}_{M_1}\right)$$
$$+ \log\det\left((1+\alpha_2)\mathrm{Id} - \frac{\rho}{\sqrt{\alpha_1}}G - \sqrt{\alpha_2 - \frac{\rho^2}{\alpha_1}}\widetilde{G}\right).$$

We may write the matrix inside the final log det as

$$\begin{aligned}(1+\alpha_2)\mathrm{Id} &- \frac{\rho}{\sqrt{\alpha_1}}G - \sqrt{\alpha_2 - \frac{\rho^2}{\alpha_1}}\widetilde{G}\\ &= ((1+\alpha_1)\mathrm{Id} - \sqrt{\alpha_1}G) + \underbrace{(\alpha_2-\alpha_1)\mathrm{Id} - \left(\frac{\rho}{\sqrt{\alpha_1}} - \sqrt{\alpha_1}\right)G}_{M_2} - \underbrace{\sqrt{\alpha_2 - \frac{\rho^2}{\alpha_1}}\widetilde{G}}_{M_3}.\end{aligned}$$

The difference of the two log det terms is thus equal to

$$\log\det\left(\mathrm{Id} + \underbrace{M_1^{-1/2}M_2M_1^{-1/2} + M_1^{-1/2}M_3M_1^{-1/2}}_{M}\right).$$

Observe that M_3 is a scaled GOE matrix independent of M_1 (and M_2). Taylor expanding the above, we shall control the trace and Frobenius norm of M. It may be verified that the higher-order terms, corresponding to higher Schatten norms, are $O(1)$. To control the trace and Frobenius norm, we shall essentially control their values in expectation. Standard concentration arguments for GOE matrices, along the lines of Lemma 7.8.6 using [GZ00, Lemma 1.2(b) and Corollary 1.6(b)], allow us to assume (with probability $1-e^{-cN}$) that the eigenvalues of G are distributed according to the semicircular distribution up to some small Wasserstein perturbation. That is, with probability $1-e^{-cN}$, denoting by $\lambda_i(G)$ the eigenvalues of G,

$$\begin{aligned}\mathsf{Tr}M_1^{-1/2}M_2M_1^{-1/2} &= \sum_{1\le i\le N} \frac{(\alpha_2-\alpha_1) - \left(\frac{\rho}{\sqrt{\alpha_1}} - \sqrt{\alpha_1}\right)\lambda_i(G)}{(1+\alpha_1) - \sqrt{\alpha_1}\lambda_i(G)}\\ &= N\cdot\int \frac{(\alpha_2-\alpha_1) - \left(\frac{\rho}{\sqrt{\alpha_1}} - \sqrt{\alpha_1}\right)u}{(1+\alpha_1) - \sqrt{\alpha_1}u}\mathsf{d}\mu_{\mathsf{sc}}(u) + O(1)\\ &= N(\alpha_2-\rho) + O(1),\end{aligned}$$

where the final equality follows from the standard semicircle integral $\int \frac{1}{x-u}\mathsf{d}\mu_{\mathsf{sc}}(u) = \frac{1}{2}\left(x - \sqrt{x^2-4}\right)$. On the other hand, because $\widetilde{G}$ is independent of G,

$$\mathsf{Tr}M_1^{-1/2}M_3M_1^{-1/2} = O(1)$$

with probability $1-e^{-cN}$. Let us next control the Frobenius norms of these matrices. Again, because $\widetilde{G}$ is independent of G, with very high probability,

$$\|M\|_F^2 = O(1) + \left\|M_1^{-1/2}M_2M_1^{-1/2}\right\|_F^2 + \left\|M_1^{-1/2}M_3M_1^{-1/2}\right\|_F^2.$$

The first squared Frobenius norm is equal to

$$\sum\left(\frac{(\alpha_2-\alpha_1) - \left(\frac{\rho}{\sqrt{\alpha_1}} - \sqrt{\alpha_1}\right)\lambda_i(G)}{(1+\alpha_1) - \sqrt{\alpha_1}\lambda_i(G)}\right)^2.$$

Let ι such that $(1+\alpha_1)-(2+\iota)\sqrt{\alpha_1} > \iota$ (this uses strict replica symmetry). Then, with probability $1-e^{-cN}$, $|\lambda_i(G)| \le 2+\iota$ for all i. Conditioned on this event happening, and recalling that $\alpha_2 - \alpha_1 = O\left(\frac{a^2+b^2}{N}\right)$, the above is $\frac{O(a^2+b^2)^2}{N}$. This is $O(1)$ for all choices of $a, b \le \iota N^{1/4}$.

We must next control the squared Frobenius norm of $M_1^{-1/2} M_3 M_1^{-1/2}$. Let us condition on a typical realization of M_1: all its eigenvalues are smaller than $2+\iota$ in magnitude, and the empirical spectral distribution is Wasserstein-close to the semicircle law in the same sense as the previous section (where we controlled the trace), in that

$$\sum \frac{1}{(1+\alpha_1)-\sqrt{\alpha_1}\lambda_i(G)} = N \cdot \int \frac{1}{(1+\alpha_1)-\sqrt{\alpha_1}u} \mathsf{d}\mu_{\mathsf{sc}}(u) + O(1) = N + O(1)$$

Because M_3 is independent of M_1, it suffices to control the expected Frobenius norm of the matrix – the true realization concentrates around its expectation to additive $O(1)$ factors. It is not difficult to see that this expectation is equal to

$$\begin{aligned}
&\frac{1}{N} \cdot \left(\alpha_2 - \frac{\rho^2}{\alpha_1}\right) \cdot \left(\sum \frac{1}{(1+\alpha_1)-\sqrt{\alpha_1}\lambda_i(G)}\right)^2 \\
&\qquad = N \cdot \left(\alpha_2 - \frac{\rho^2}{\alpha_1}\right) \cdot \left(\int \frac{1}{(1+\alpha_1)-\sqrt{\alpha_1}u} \mathsf{d}\mu_{\mathsf{sc}}(u)\right)^2 + O(1) \\
&\qquad = N \cdot \left(\alpha_2 - \frac{\rho^2}{\alpha_1}\right) + O(1).
\end{aligned}$$

Putting the pieces together and returning to the Taylor expansion, we get that with very high probability,

$$\begin{aligned}
&2\left(\mathsf{Error}^{(2)}_{0,0} - \mathsf{Error}^{(2)}_{a,b}\right) \\
&\quad = \frac{N\alpha_1}{2} - \frac{N\alpha_2}{2} + \log\det\left(\mathrm{Id} + M\right) \\
&\quad = O(1) + \frac{N\alpha_1}{2} - \frac{N\alpha_2}{2} + \mathsf{Tr}\left(M\right) - \frac{1}{2}\|M\|_F^2 \\
&\quad = O(1) + N \cdot \left(\frac{\alpha_1}{2} - \frac{\alpha_2}{2} + (\alpha_2 - \rho) - \frac{1}{2}\left(\alpha_2 - \frac{\rho^2}{\alpha_1}\right)\right) \\
&\quad = O(1) + N \cdot \left(\frac{\alpha_1}{2} - \rho + \frac{\rho^2}{2\alpha_1}\right) \\
&\quad = O(1) + \frac{N}{2} \cdot \frac{(\alpha_1 - \rho)^2}{\alpha_1}.
\end{aligned}$$

Because $\alpha_1 - \rho = O\left(\frac{a^2+b^2}{N}\right)$, this is $O(1)$ for $a, b \le \iota N^{1/4}$, completing the proof. □

7.B.3 Moment calculations for covariance bounds

We first prove subgaussian concentration for the covariance of the degree-2 part.

Proof of Proposition 7.8.8. Part (2) follows from part (1) by a standard tail integration argument. Indeed, the random variable W is bounded by $N^{1/2}$, so the contribution to $\mathbb{E}[W^k]$ from the event $|W| \ge N^{1/5}$ is bounded by

$$N^{k/2}\mathbb{P}(|W| \ge N^{1/5}) \le N^{k/2} e^{-cN^{2/5}},$$

which is vanishing for any constant k. So, we focus on proving part (1). In the case $W = \langle \sigma^1, \sigma^2 \rangle / \sqrt{N}$, this is a special case of [HMP24, Lemma 7.5] (where we take $u = 0$). We consider the case $W = \langle \sigma, v_i \rangle$. Recall that

$$H_{N,2}(\sigma) = \langle A\sigma, \sigma \rangle = \frac{\sqrt{\xi''(0)}}{2} \langle M\sigma, \sigma \rangle$$

where $M \sim \mathsf{GOE}(N)$. For $0 \le s \le N^{1/5}\log N$, we will evaluate

$$\int e^{H_{N,2}(\sigma)}\mathsf{d}\rho(\sigma) \qquad \text{and} \qquad \int e^{H_{N,2}(\sigma)+s\langle v_i,\sigma\rangle}\mathsf{d}\rho(\sigma)$$

using [HMP24, Lemma 7.3]. We recall the function $G : (\lambda_{\max}(A), +\infty)$ defined in (7.15), which we copy below for convenience, and define $\widetilde{G}$ by

$$G(\gamma) = \gamma - \frac{1}{2N}\log\det(\gamma I - A), \qquad \widetilde{G}(\gamma) = G(\gamma) + \frac{s^2}{4N(\gamma - \lambda_i(A))}.$$

Recall from below (7.15) that G' has a unique root γ_* on $(\lambda_{\max}(A), +\infty)$. By the same argument, $\widetilde{G}'$ has a unique root $\widetilde{\gamma}_*$ on the same interval. As argued in the proof of Lemma 7.8.6, the conditions of [HMP24, Lemma 7.3] apply. Applying this lemma with $u = 0$ and $u = sv_i$, respectively, shows that with probability $1 - e^{-cN}$,

$$\int e^{H_{N,2}(\sigma)}\mathsf{d}\rho(\sigma) = (1 + O(N^{-c}))\sqrt{\frac{2}{G''(\gamma_*)}}(2e)^{-N/2}\exp(NG(\gamma_*)),$$

$$\int e^{H_{N,2}(\sigma)+s\langle v_i,\sigma\rangle}\mathsf{d}\rho(\sigma) = (1 + O(N^{-c}))\sqrt{\frac{2}{\widetilde{G}''(\widetilde{\gamma}_*)}}(2e)^{-N/2}\exp(N\widetilde{G}(\widetilde{\gamma}_*)). \tag{7.39}$$

Suppose further the probability $1 - e^{-cN}$ event in Lemma 7.8.7 holds. As argued in the proof of Proposition 7.8.1, $|\gamma_* - \gamma_0| \le \frac{1}{C\sqrt{N}}$. Also, with probability $1 - e^{-cN}$, $\lambda_{\max}(A) \le \frac{\sqrt{\xi''(0)}}{2}(2 + \varepsilon^2/8)$.

We will show that on the intersection of these events $|\gamma_* - \widetilde{\gamma}_*| = O(N^{-3/5}\log^2 N)$. First note that

$$\gamma_* - \lambda_{\max}(A) \ge \gamma_0 - \lambda_{\max}(A) - \frac{1}{C\sqrt{N}} \ge \frac{1 + \xi''(0) - (2 + \varepsilon^2/8)\sqrt{\xi''(0)}}{2} - \frac{1}{C\sqrt{N}} \ge \varepsilon^2/32$$

is bounded below by a constant, as in the proof of Fact 7.8.5. Since $G'(\gamma_*) = 0$ and $\widetilde{G}'(\gamma) = G'(\gamma) - \frac{s^2}{4N(\gamma-\lambda_i(A))^2}$, we have $\widetilde{\gamma}_* \ge \gamma_*$ and so $\widetilde{\gamma}_* - \lambda_i(A)$ is also bounded below by a constant. Thus, as $s \le N^{1/5}\log N$, we have

$$0 \ge \widetilde{G}'(\gamma_*) = -\frac{s^2}{4N(\gamma - \lambda_i(A))^2} = O(N^{-3/5}\log^2 N).$$

By direct computation,

$$\widetilde{G}''(\gamma) = G''(\gamma) + \frac{s^2}{8N(\gamma - \lambda_i(A))^2}.$$

Since $\widetilde{\gamma}_* \ge \gamma_*$, we have $\gamma - \lambda_i(A)$ is bounded below by a constant for all $\gamma \in [\gamma_*, \widetilde{\gamma}_*]$. We claim that $\widetilde{\gamma}_* \in [\gamma_0 - N^{-1/2}, \gamma_0 + N^{-1/2}]$. Combining both conclusions of Lemma 7.8.7, we obtain that $G''(\gamma) \ge \Omega_\varepsilon(1)$ for $\gamma \in [\gamma_0 - N^{-1/2}, \gamma_0 + N^{-1/2}]$, and consequently the desired claim on $\widetilde{\gamma}_*$ holds.

We thus have $\widetilde{G}''(\gamma) = O_\varepsilon(1)$ in the same interval, and consequently we can conclude the stronger statement $|\gamma_* - \widetilde{\gamma}_*| = O(N^{-3/5}\log^2 N)$. Furthermore, since $\widetilde{G}^{(3)}(\gamma) = O_\varepsilon(1)$, the same logic allows us to also conclude

$$\widetilde{G}''(\widetilde{\gamma}_*)/\widetilde{G}''(\gamma_*) = 1 + O(N^{-3/5}\log^2 N).$$

By Taylor expanding $\widetilde{G}$ around $\widetilde{\gamma}_*$, we see

$$N|\widetilde{G}(\gamma_*) - \widetilde{G}(\widetilde{\gamma}_*)| \le \frac{N}{2}|\gamma_* - \widetilde{\gamma}_*|^2 \sup_{\gamma\in[\gamma_0 - N^{-1/2}, \gamma_0 + N^{-1/2}]} G''(\gamma) = O(N^{-1/5}\log^4 N).$$

The above two displays allow us to replace instances of $\widetilde{\gamma}_*$ with γ_* in (7.39), yielding

$$\int e^{H_{N,2}(\sigma)+s\langle v_i,\sigma\rangle}\mathsf{d}\rho(\sigma) = (1 + O(N^{-c}))\sqrt{\frac{2}{\widetilde{G}''(\gamma_*)}}(2e)^{-N/2}\exp(N\widetilde{G}(\gamma_*)),$$

and thus

$$\begin{aligned}
\langle e^{s\langle v_i,\sigma\rangle}\rangle_2 &= \frac{\int e^{H_{N,2}(\sigma)+s\langle v_i,\sigma\rangle}\mathsf{d}\rho(\sigma)}{\int e^{H_{N,2}(\sigma)}\mathsf{d}\rho(\sigma)} \\
&= (1+O(N^{-c}))\sqrt{\frac{G''(\gamma_*)}{\widetilde{G}''(\gamma_*)}}\exp(N(\widetilde{G}(\gamma_*)-G(\gamma_*))) \\
&= (1+O(N^{-c}))\exp(s^2/(4(\gamma_*-\lambda_i(A)))) = (1+O(N^{-c}))\exp(cs^2),
\end{aligned}$$

where the last two steps again use that $\gamma_* - \lambda_i(A)$ is bounded away from 0. The tail estimate on $W = \langle v_i, \sigma\rangle$ now follows from a standard Chernoff bound. □

We next turn to Lemma 7.8.11, proving each part in turn.

Proof of Lemma 7.8.11, (7.18). We reproduce (7.18) below for convenience:

$$\mathbb{E}\int_{S_N} \mathbf{1}[\sigma \notin T(H_N)]e^{H_N(\sigma)}\mathsf{d}\rho(\sigma) \le e^{N\xi(1)/2-cN^{1/5}}.$$

The proof follows [HS23b, Proposition 3.1], except with more precise control of overlaps between $N^{-2/5}$ and a small constant. By symmetry of the sphere, for any deterministic $\boldsymbol{x} \in S_N$,

$$\mathbb{E}\int_{S_N} \mathbf{1}[\sigma \notin T(H_N)]e^{H_N(\sigma)}\mathsf{d}\rho(\sigma) = \mathbb{E}\left[\mathbf{1}[\boldsymbol{x} \notin T(H_N)]e^{H_N(\boldsymbol{x})}\right]. \tag{7.40}$$

Let $\mu_{\mathsf{pl}}(\cdot|\boldsymbol{x})$ denote the planted model (Definition 7.7.9) conditional on spike $\boldsymbol{x}$. A Gaussian change of measure calculation implies that the right-hand side of (7.40) equals

$$e^{N\xi(1)/2}\mathbb{P}_{H_N^{\boldsymbol{x},\mathrm{I}}\sim\mu_{\mathsf{pl}}(\cdot|\boldsymbol{x})}[\boldsymbol{x} \notin T(H_N^{\boldsymbol{x},\mathrm{I}})].$$

Thus it suffices to show

$$\mathbb{P}_{H_N^{\boldsymbol{x},\mathrm{I}}\sim\mu_{\mathsf{pl}}(\cdot|\boldsymbol{x})}[\boldsymbol{x} \notin T(H_N^{\boldsymbol{x},\mathrm{I}})] \le e^{-cN^{1/5}}.$$

Recall (Remark 7.7.15) that a sample $H_N^{\boldsymbol{x},\mathrm{I}} \sim \mu_{\mathsf{pl}}(\cdot|\boldsymbol{x})$ can be generated by

$$H_N^{\boldsymbol{x},\mathrm{I}}(\sigma) = N\xi(R(\boldsymbol{x},\sigma)) + \widetilde{H}_N(\sigma), \tag{7.41}$$

where $\widetilde{H}_N \sim \mu_{\mathsf{null}}$. Furthermore, from the definition, $\boldsymbol{x} \in T(H_N^{\boldsymbol{x},\mathrm{I}})$ is equivalent to

$$\int_{S_N} \mathbf{1}[|R(\boldsymbol{x},\tau)| \ge N^{-2/5}]e^{H_N^{\boldsymbol{x},\mathrm{I}}(\tau)}\mathsf{d}\rho(\tau) \le e^{N\xi(1)/2-cN^{1/5}} \tag{7.42}$$

We will show this occurs with probability at least $1 - e^{-cN^{1/5}}$. Let ψ denote the probability density of $R(\boldsymbol{x},\tau) \in [-1,1]$, where τ is sampled from the Haar measure on S_N. Then it is known that

$$\psi(q) = \frac{1}{Z_\psi}(1-q^2)^{(N-3)/2}$$

where $Z_\psi = \Theta(N^{-1/2})$. Define the codimension-1 band

$$\mathsf{Band}(q) = \mathsf{Band}(q;\boldsymbol{x}) := \{\tau \in S_N : R(\boldsymbol{x},\tau) = q\}$$

and let

$$Z^{\boldsymbol{x},\mathrm{I}}(q) = \int_{\mathsf{Band}(q)} e^{H_N^{\boldsymbol{x},\mathrm{I}}(\tau)}\mathsf{d}\rho_q(\tau),$$

where ρ_q is the Haar measure on $\mathsf{Band}(q)$, normalized so that $\rho_q(\mathsf{Band}(q)) = \psi(q)$. Then the left-hand side of (7.42) is equal to

$$\int_{N^{-2/5}\le|q|\le 1} Z^{\boldsymbol{x},\mathrm{I}}(q)\mathsf{d}q.$$

An application of Guerra's interpolation as in [HS23b, Lemma 3.3] shows that for any $q \in [-1,1]$ and constant $\eta > 0$, with probability $1 - e^{-cN}$,

$$\frac{1}{N} \log Z^{\boldsymbol{x},\mathrm{I}}(q) \le \frac{1}{2}\left(\xi(1) + \xi(|q|) + |q| + \log(1-|q|)\right) + \eta.$$

Since $\xi_{\sim 1}$ is ε-strictly replica symmetric and $\gamma_1^2 \le N^{-4/5}$, this implies

$$\frac{1}{N} \log Z^{\boldsymbol{x},\mathrm{I}}(q) \le \frac{\xi(1)}{2} - \frac{\varepsilon q^2}{4} + 2\eta.$$

Let $\delta > 0$ be small depending on ε, and η small depending on δ. This implies that for any $|q| \ge \delta$, with probability $1 - e^{-cN}$,

$$\frac{1}{N} \log Z^{\boldsymbol{x},\mathrm{I}}(q) \le \frac{\xi(1)}{2} - \frac{\varepsilon \delta^2}{8}.$$

Taking a union bound over a N^{-1}-net of $|q| \ge \delta$ as in [HS23b, Lemma 3.4] implies that with probability $1 - e^{-cN}$,

$$\int_{\delta \le |q| \le 1} Z^{\boldsymbol{x},\mathrm{I}}(q) \mathsf{d}q \le e^{N\xi(1)/2 - cN}. \tag{7.43}$$

We address the remaining range of q by a first moment bound. Note that

$$\mathbb{E} \int_{N^{-2/5} \le |q| \le \delta} Z^{\boldsymbol{x},\mathrm{I}}(q) \mathsf{d}q = e^{N\xi(1)/2} \int_{N^{-2/5} \le |q| \le \delta} e^{N\xi(q)} \psi(q) \mathsf{d}q. \tag{7.44}$$

Recall from Fact 7.8.4 that $\xi''(0) \le 1 - \varepsilon$. Thus, for sufficiently small δ, for all $|q| \le \delta$,

$$\xi_{\sim 1}(q) + \frac{1}{2}\log(1-q^2) \le -\varepsilon q^2/4.$$

Thus, for all $N^{-2/5} \le |q| \le \delta$,

$$\begin{aligned}\frac{1}{N}\log\left(e^{N\xi(q)}\psi(q)\right) = \xi(q) + \frac{1}{N}\log\psi(q) &= \gamma_1^2 q + \xi_{\sim 1}(q) + \frac{1}{2}\log(1-q^2) + O(N^{-1}\log N) \\ &\le \gamma_1^2 q - \varepsilon q^2/4 + O(N^{-1}\log N) \le -\varepsilon N^{-4/5}/8.\end{aligned}$$

Combining with (7.44) shows

$$\mathbb{E} \int_{N^{-2/5} \le |q| \le \delta} Z^{\boldsymbol{x},\mathrm{I}}(q) \mathsf{d}q \le e^{N\xi(1)/2 - cN^{1/5}},$$

so by Markov's inequality, with probability $1 - e^{-cN^{1/5}/2}$,

$$\int_{N^{-2/5} \le |q| \le \delta} Z^{\boldsymbol{x},\mathrm{I}}(q) \mathsf{d}q \le e^{N\xi(1)/2 - cN^{1/5}/2}.$$

Combining with (7.43) proves (7.42) after adjusting c. $\square$

Proof of Lemma 7.8.11, (7.19). By the same argument leading to (7.42), it suffices to prove

$$\int_{S_N} \mathbf{1}[|R(\boldsymbol{x},\tau)| \ge N^{-2/5}] e^{H_N^{\boldsymbol{x},II}(\tau)} \mathsf{d}\rho(\tau) \le e^{N\xi(1)/2 - cN^{1/5}} \tag{7.45}$$

holds with probability at least $1 - e^{-cN^{1/5}}$, where now

$$H_N^{\boldsymbol{x},II}(\sigma) = N\gamma_2^2 R(\boldsymbol{x},\sigma)^2 + \widetilde{H}_N(\sigma),$$

i.e. we have replaced the spike in $H_N^{\boldsymbol{x},\mathrm{I}}$ with only its degree-2 part. Then $H_N^{\boldsymbol{x},\mathrm{II}}(\sigma) \le H_N^{\boldsymbol{x},\mathrm{I}}(\sigma)$ almost surely for all σ such that $R(\boldsymbol{x},\sigma) \ge 0$, so (7.42) implies

$$\int_{S_N} \mathbf{1}[R(\boldsymbol{x},\tau) \ge N^{-2/5}] e^{H_N^{\boldsymbol{x},\mathrm{II}}(\tau)} \mathsf{d}\rho(\tau) \le e^{N\xi(1)/2 - cN^{1/5}}$$

with probability $1 - e^{-cN}$. Moreover, by symmetry of the degree-2 spike,

$$\int_{S_N} \mathbf{1}[R(\boldsymbol{x},\tau) \ge N^{-2/5}] e^{H_N^{\boldsymbol{x},\mathrm{II}}(\tau)} \mathsf{d}\rho(\tau) \overset{d}{=} \int_{S_N} \mathbf{1}[R(\boldsymbol{x},\tau) \le -N^{-2/5}] e^{H_N^{\boldsymbol{x},\mathrm{II}}(\tau)} \mathsf{d}\rho(\tau).$$

This implies (7.45) and thus (7.19). □

Proof of Lemma 7.8.11, (7.20). This follows from a slightly more complex form of the same strategy, where the planted Hamiltonian now has two spikes. For $\boldsymbol{x}^1, \boldsymbol{x}^2 \in S_N$, let

$$H_N^{\boldsymbol{x}^1,\boldsymbol{x}^2,III}(\sigma) = N\xi(R(\boldsymbol{x}^1,\sigma)) + N\xi(R(\boldsymbol{x}^2,\sigma)) + \widetilde{H}_N(\sigma).$$

By the same gaussian change of measure argument as above, the expectation in the left-hand side of (7.20) equals

$$e^{N\xi(1)} \int \mathbf{1}[|R(\sigma^1,\sigma^2)| \le 3N^{-2/5}] \mathbb{P}(\sigma^1 \notin T(H_N^{\sigma^1,\sigma^2,III})) \mathsf{d}\rho^{\otimes 2}(\sigma^1,\sigma^2).$$

Thus it suffices to show that for all $\boldsymbol{x}^1, \boldsymbol{x}^2 \in S_N$ with $|R(\boldsymbol{x}^1,\boldsymbol{x}^2)| \le 3N^{-2/5}$,

$$\int_{S_N} \mathbf{1}[|R(\boldsymbol{x}^1,\tau)| \ge N^{-2/5}] e^{H_N^{\boldsymbol{x}^1,\boldsymbol{x}^2,III}(\tau)} \mathsf{d}\rho(\tau) \le e^{N\xi(1)/2 - cN^{1/5}} \tag{7.46}$$

with probability at least $1 - e^{-cN^{1/5}}$. Let $\lambda = R(\boldsymbol{x}^1,\boldsymbol{x}^2) \in [-3N^{-2/5}, 3N^{2/5}]$ and

$$\boldsymbol{x}^2 = \lambda \boldsymbol{x}^1 + \sqrt{1-\lambda^2}\boldsymbol{x}^2_\perp,$$

where $\boldsymbol{x}^2_\perp \in S_N$ and $R(\boldsymbol{x}^1,\boldsymbol{x}^2_\perp) = 0$. Let ψ_2 denote the probability density of $(R(\boldsymbol{x}^1,\tau), R(\boldsymbol{x}^2_\perp,\tau)) \in [1,1]^2$, where τ is sampled from the Haar measure on S_N. It is known that

$$\psi_2(q) = \frac{\mathbf{1}[q_1^2 + q_2^2 \le 1]}{Z_{\psi_2}} (1 - q_1^2 - q_2^2)^{(N-4)/2}$$

where $Z_\psi = \Theta(N^{-1})$. Define the codimension-2 band

$$\mathsf{Band}(q_1,q_2) = \mathsf{Band}(q_1,q_2;\boldsymbol{x}^1,\boldsymbol{x}^2) \coloneqq \{\tau \in S_N : R(\boldsymbol{x}^1,\tau) = q_1, R(\boldsymbol{x}^2_\perp,\tau) = q_2\}.$$

and let

$$Z^{\boldsymbol{x}^1,\boldsymbol{x}^2,III}(q_1,q_2) = \int_{\mathsf{Band}(q_1,q_2)} e^{H_N^{\boldsymbol{x}^1,\boldsymbol{x}^2,III}(\tau)} \mathsf{d}\rho_{q_1,q_2}(\tau),$$

where ρ_{q_1,q_2} is the Haar measure on $\mathsf{Band}(q_1,q_2)$, normalized so that $\rho_{q_1,q_2}(\mathsf{Band}(q_1,q_2)) = \psi_2(q_1,q_2)$. Then the left-hand side of (7.46) is equal to

$$\int \mathbf{1}[|q_1| \ge N^{-2/5}] Z^{\boldsymbol{x}^1,\boldsymbol{x}^2,III}(q_1,q_2) \mathsf{d}(q_1,q_2). \tag{7.47}$$

Note that

$$\frac{1}{N} \log Z^{\boldsymbol{x}^1,\boldsymbol{x}^2,III}(q_1,q_2) = \xi(q_1) + \xi\left(\lambda q_1 + \sqrt{1-\lambda^2} q_2\right) + \frac{1}{N} \log \int_{\mathsf{Band}(q_1,q_2)} e^{\widetilde{H}_N(\tau)} \mathsf{d}\rho_{q_1,q_2}(\tau).$$

Let $\widetilde{q} = \widetilde{q}(q_1,q_2) \coloneqq \sqrt{q_1^2 + q_2^2}$. Applying Guerra's interpolation as in [HS23b, Lemma 3.3] shows that for any $\eta > 0$, with probability $1 - e^{-cN}$

$$\frac{1}{N} \log \int_{\mathsf{Band}(q_1,q_2)} e^{\widetilde{H}_N(\tau)} \mathsf{d}\rho_{q_1,q_2}(\tau) \le \frac{1}{2} \left(\xi(1) - \xi(\widetilde{q}) + \widetilde{q} + \log(1 - \widetilde{q})\right) + \eta.$$

and thus

$$\frac{1}{N}\log Z^{\boldsymbol{x}^1,\boldsymbol{x}^2,III}(q_1,q_2) \le \xi(q_1) + \xi\left(\lambda q_1 + \sqrt{1-\lambda^2}q_2\right) + \frac{1}{2}\left(\xi(1) - \xi(\widetilde{q}) + \widetilde{q} + \log(1-\widetilde{q})\right) + \eta$$
$$\le \xi_{\sim 1}(|q_1|) + \xi_{\sim 1}(|q_2|) + \frac{1}{2}\left(\xi(1) - \xi_{\sim 1}(\widetilde{q}) + \widetilde{q} + \log(1-\widetilde{q})\right) + 2\eta.$$

Since $\xi_{\sim 1}$ only includes terms that are degree 2 or larger, $\xi_{\sim 1}(|q_1|)+\xi_{\sim 1}(|q_2|) \le \xi_{\sim 1}(\widetilde{q})$. Thus the last display is bounded by

$$\frac{1}{2}\left(\xi(1) + \xi_{\sim 1}(\widetilde{q}) + \widetilde{q} + \log(1-\widetilde{q})\right) + 2\eta \le \frac{\xi(1)}{2} - \frac{\varepsilon\widetilde{q}^2}{4} + 2\eta.$$

Arguing as in the proof of equation (7.18) then shows that for any $\delta > 0$ depending only on ε,

$$\int \mathbf{1}[\widetilde{q}(q_1,q_2) > \delta] Z^{\boldsymbol{x}^1,\boldsymbol{x}^2,III}(q_1,q_2)\mathsf{d}(q_1,q_2) \le e^{N\xi(1)/2 - cN} \tag{7.48}$$

with probability $1-e^{-cN}$. The remaining part of the integral (7.47) has expectation

$$\mathbb{E}\int \mathbf{1}[|q_1| \ge N^{-2/5}, \widetilde{q}(q_1,q_2) \le \delta] Z^{\boldsymbol{x}^1,\boldsymbol{x}^2,III}(q_1,q_2)\mathsf{d}(q_1,q_2)$$
$$= e^{N\xi(1)/2}\int \mathbf{1}[|q_1| \ge N^{-2/5}, \widetilde{q}(q_1,q_2) \le \delta] e^{N\xi(q_1) + N\xi(\lambda q_1 + \sqrt{1-\lambda^2}q_2)}\psi_2(q_1,q_2)\mathsf{d}(q_1,q_2). \tag{7.49}$$

Recall $\gamma_1^2 \le N^{-4/5}$. For all (q_1,q_2) in this indicator,

$$\begin{aligned}
&\frac{1}{N}\log\left(e^{N\xi(q_1) + N\xi(\lambda q_1 + \sqrt{1-\lambda^2}q_2)}\psi_2(q_1,q_2)\right) \\
&= \xi(q_1) + \xi\left(\lambda q_1 + \sqrt{1-\lambda^2}q_2\right) + \frac{1}{2}\log(1-\widetilde{q}^2) + O(N^{-1}\log N) \\
&\le 2N^{-4/5}\widetilde{q} - \frac{1}{2}(1-\xi''(0))\widetilde{q}^2 + O(\widetilde{q}^3 + N^{-1}\log N) \\
&\le -\frac{\varepsilon\widetilde{q}^2}{2} + 2N^{-4/5}\widetilde{q} + O(\widetilde{q}^3 + N^{-1}\log N).
\end{aligned}$$

Since $N^{-2/5} \le \widetilde{q} \le \delta$, for δ sufficiently small depending on ε this is bounded by $-cN^{-4/5}$. Combining with (7.49) shows that with probability $1-e^{-cN^{1/5}}$,

$$\int \mathbf{1}[|q_1| \ge N^{-2/5}, \widetilde{q}(q_1,q_2) \le \delta] Z^{\boldsymbol{x}^1,\boldsymbol{x}^2,III}(q_1,q_2)\mathsf{d}(q_1,q_2) \le e^{N\xi(1)/2 - cN^{1/5}}.$$

Further combining with (7.48) completes the proof. □

Proof of Lemma 7.8.11, (7.21). This is proved identically to equation (7.20), except with spiked Hamiltonian

$$H_N^{\boldsymbol{x}^1,\boldsymbol{x}^2,\mathrm{IV}}(\sigma) = N\gamma_2^2 R(\boldsymbol{x}^1,\sigma)^2 + N\xi(R(\boldsymbol{x}^2,\sigma)) + \widetilde{H}_N(\sigma),$$

i.e. the spike involving $\boldsymbol{x}^1$ is replaced with just its degree-2 component. The same argument applies and we omit details. □

We turn to the proofs of Lemmas 7.8.15 and 7.8.16. The following fact will be useful in the proofs of both lemmas.

Fact 7.B.4. *If $\sigma \in S_N$ satisfies $|\sigma_i|, |\sigma_j| \le \log N$ and $\sigma \notin \widetilde{T}_{i,j}$, then $\sigma \notin \widetilde{T}$.*

Proof. Consider any $\tau \in S_N$ satisfying $|R(\sigma_{\sim i,j}, \tau_{\sim i,j})| \ge 2N^{-2/5}$. Since $|\tau_i|, |\tau_j| \le N^{1/2}$,

$$|R(\sigma,\tau)| \ge |R(\sigma_{\sim i,j}, \tau_{\sim i,j})| - \frac{|\sigma_i||\tau_i| + |\sigma_j||\tau_j|}{N} \ge N^{-2/5}.$$

Thus the expectation over τ in (7.22) is larger than the expectation over τ in (7.28), and so $\sigma \notin \widetilde{T}$. □

Proof of Lemma 7.8.15. We can write $X_{i,j} = \widetilde{X}_{i,j} + \widetilde{X}^{(1)}_{i,j} + \widetilde{X}^{(2)}_{i,j}$, where

$$\begin{aligned}\widetilde{X}^{(1)}_{i,j} &= \left\langle \mathbf{1}[|\sigma_i|, |\sigma_j| \le \log N, \sigma \notin \widetilde{T}_{i,j}]\sigma_i\sigma_j \left(e^{H_{N,\sim 2}(\sigma) - N\xi_{\sim 2}(1)/2} - \mathbf{1}[i=j]\right)\right\rangle_2, \\ \widetilde{X}^{(2)}_{i,j} &= \left\langle \mathbf{1}[|\sigma_i| \vee |\sigma_j| > \log N]\sigma_i\sigma_j \left(e^{H_{N,\sim 2}(\sigma) - N\xi_{\sim 2}(1)/2} - \mathbf{1}[i=j]\right)\right\rangle_2.\end{aligned}$$

By using $(a+b)^2 \le 2a^2 + 2b^2$, we deduce that

$$X^2_{i,j} \le 2\widetilde{X}^2_{i,j} + 2(\widetilde{X}^{(1)}_{i,j} + \widetilde{X}^{(2)}_{i,j})^2,$$

so the rest of the proof is dedicated to showing that $|\widetilde{X}^{(1)}_{i,j} + \widetilde{X}^{(2)}_{i,j}| \le e^{-c\log^2 N}$ with probability $1 - e^{-c\log^2 N}$. To do so, we will simply apply Markov to control $|\widetilde{X}^{(1)}_{i,j}|$ and $|\widetilde{X}^{(2)}_{i,j}|$.

By Fact 7.B.4 and using $|\sigma_i|, |\sigma_j| \le \sqrt{N}$,

$$\begin{aligned}|\widetilde{X}^{(1)}_{i,j}| &\le N\left\langle \mathbf{1}[|\sigma_i|, |\sigma_j| \le \log N, \sigma \notin \widetilde{T}_{i,j}]\left(e^{H_{N,\sim 2}(\sigma) - N\xi_{\sim 2}(1)/2} + \mathbf{1}[i=j]\right)\right\rangle_2 \\ &\le N\left\langle \mathbf{1}[\sigma \notin \widetilde{T}]\left(e^{H_{N,\sim 2}(\sigma) - N\xi_{\sim 2}(1)/2} + \mathbf{1}[i=j]\right)\right\rangle_2.\end{aligned}$$

By the first two equations from Corollary 7.8.12, $\mathbb{E}_{\sim 2}|\widetilde{X}^{(1)}_{i,j}| \le e^{-cN^{1/5}}$. Furthermore,

$$\begin{aligned}\mathop{\mathbb{E}}_{\sim 2}|\widetilde{X}^{(2)}_{i,j}| &\le N\mathop{\mathbb{E}}_{\sim 2}\left\langle (\mathbf{1}[|\sigma_i| > \log N] + \mathbf{1}[|\sigma_j| > \log N])\left(e^{H_{N,\sim 2}(\sigma) - N\xi_{\sim 2}(1)/2} + \mathbf{1}[i=j]\right)\right\rangle_2 \\ &\le 2N\left\langle \mathbf{1}[|\sigma_i| > \log N] + \mathbf{1}[|\sigma_j| > \log N]\right\rangle_2 \le e^{-c\log^2 N}\end{aligned}$$

by Proposition 7.8.8(1). By Markov's inequality, with probability $1 - e^{-c\log^2 N}$, $|\widetilde{X}^{(1)}_{i,j}| + |\widetilde{X}^{(2)}_{i,j}| \le e^{-c\log^2 N}$, after adjusting c as necessary. This concludes the proof. □

Proof of Lemma 7.8.16. We can write $\widetilde{X}^2_{i,j} = \widehat{X}_{i,j} + \widehat{X}^{(1)}_{i,j} - \widehat{X}^{(2)}_{i,j}$, where

$$\begin{aligned}\widehat{X}^{(1)}_{i,j} = \Bigg\langle &\mathbf{1}[|\sigma^1_i|, |\sigma^1_j|, |\sigma^2_i|, |\sigma^2_j| \le \log N, \sigma^1, \sigma^2 \in \widetilde{T}_{i,j}, |R(\sigma^1_{\sim i,j}, \sigma^2_{\sim i,j})| > 2N^{-2/5}] \\ &\sigma^1_i\sigma^1_j\sigma^2_i\sigma^2_j\left(e^{H_{N,\sim 2}(\sigma^1) - N\xi_{\sim 2}(1)/2} - \mathbf{1}[i=j]\right)\left(e^{H_{N,\sim 2}(\sigma^2) - N\xi_{\sim 2}(1)/2} - \mathbf{1}[i=j]\right)\Bigg\rangle_2, \\ \widehat{X}^{(2)}_{i,j} = \Bigg\langle &\mathbf{1}[|\sigma^1_i|, |\sigma^1_j|, |\sigma^2_i|, |\sigma^2_j| \le \log N, (\sigma^1 \notin \widetilde{T}_{i,j} \vee \sigma^2 \notin \widetilde{T}_{i,j}), |R(\sigma^1_{\sim i,j}, \sigma^2_{\sim i,j})| \le 2N^{-2/5}] \\ &\sigma^1_i\sigma^1_j\sigma^2_i\sigma^2_j\left(e^{H_{N,\sim 2}(\sigma^1) - N\xi_{\sim 2}(1)/2} - \mathbf{1}[i=j]\right)\left(e^{H_{N,\sim 2}(\sigma^2) - N\xi_{\sim 2}(1)/2} - \mathbf{1}[i=j]\right)\Bigg\rangle_2.\end{aligned}$$

By the same argument as before, it suffices to control $\mathbb{E}_{\sim 2}|\widehat{X}^{(1)}_{i,j}|$ and $\mathbb{E}_{\sim 2}|\widehat{X}^{(2)}_{i,j}|$. Note that almost surely,

$$\begin{aligned}|\widehat{X}^{(1)}_{i,j}| \le N^2\Bigg\langle &\mathbf{1}[|\sigma^1_i|, |\sigma^1_j|, |\sigma^2_i|, |\sigma^2_j| \le \log N, \sigma^1, \sigma^2 \in \widetilde{T}_{i,j}, |R(\sigma^1_{\sim i,j}, \sigma^2_{\sim i,j})| > 2N^{-2/5}] \\ &\left(e^{H_{N,\sim 2}(\sigma^1) - N\xi_{\sim 2}(1)/2} + 1\right)\left(e^{H_{N,\sim 2}(\sigma^2) - N\xi_{\sim 2}(1)/2} + 1\right)\Bigg\rangle_2 \le N^2(\widehat{X}^{(3)}_{i,j} + \widehat{X}^{(4)}_{i,j})\end{aligned}$$

where

$$\begin{aligned}\widehat{X}^{(3)}_{i,j} &= \left\langle \mathbf{1}[\sigma^1 \in \widetilde{T}_{i,j}, |R(\sigma^1_{\sim i,j}, \sigma^2_{\sim i,j})| > 2N^{-2/5}]e^{H_{N,\sim 2}(\sigma^1) + H_{N,\sim 2}(\sigma^2) - N\xi_{\sim 2}(1)}\right\rangle_2, \\ \widehat{X}^{(4)}_{i,j} &= \Bigg\langle \mathbf{1}[|\sigma^1_i|, |\sigma^1_j|, |\sigma^2_i|, |\sigma^2_j| \le \log N, |R(\sigma^1_{\sim i,j}, \sigma^2_{\sim i,j})| > 2N^{-2/5}] \\ &\qquad (e^{H_{N,\sim 2}(\sigma^1) - N\xi_{\sim 2}(1)/2} + e^{H_{N,\sim 2}(\sigma^2) - N\xi_{\sim 2}(1)/2} + 1)\Bigg\rangle_2.\end{aligned}$$

By definition of $\widetilde{T}_{i,j}$,
$$\widehat{X}^{(3)}_{i,j} \le \left\langle \mathbf{1}[\sigma^1 \in \widetilde{T}_{i,j}] e^{H_{N,\sim 2}(\sigma^1) - N\xi_{\sim 2}(1)/2 - cN^{1/5}} \right\rangle_2,$$
and thus $\mathbb{E}_{\sim 2} \widehat{X}^{(3)}_{i,j} \le e^{-cN^{1/5}}$. Furthermore,
$$\widehat{X}^{(4)}_{i,j} \le \left\langle \mathbf{1}[|R(\sigma^1,\sigma^2)| > N^{-2/5}](e^{H_{N,\sim 2}(\sigma^1) - N\xi_{\sim 2}(1)/2} + e^{H_{N,\sim 2}(\sigma^2) - N\xi_{\sim 2}(1)/2} + 1) \right\rangle_2,$$
and thus
$$\mathop{\mathbb{E}}_{\sim 2} \widehat{X}^{(4)}_{i,j} \le 3 \left\langle \mathbf{1}[|R(\sigma^1,\sigma^2)| > N^{-2/5}] \right\rangle_2 \le e^{-cN^{1/5}}$$
by Proposition 7.8.8(1). Combining shows $\mathbb{E}_{\sim 2} |\widehat{X}^{(1)}_{i,j}| \le e^{-cN^{1/5}}$. Similarly
$$\begin{aligned}|\widehat{X}^{(2)}_{i,j}| \le N^2 \Big\langle & \mathbf{1}[|\sigma^1_i|, |\sigma^1_j|, |\sigma^2_i|, |\sigma^2_j| \le \log N, (\sigma^1 \not\in \widetilde{T}_{i,j} \vee \sigma^2 \not\in \widetilde{T}_{i,j}), \\ & |R(\sigma^1_{\sim i,j}, \sigma^2_{\sim i,j})| \le 2N^{-2/5}] \left(e^{H_{N,\sim 2}(\sigma^1) - N\xi_{\sim 2}(1)/2} + 1\right) \left(e^{H_{N,\sim 2}(\sigma^2) - N\xi_{\sim 2}(1)/2} + 1\right) \Big\rangle_2\end{aligned}$$
By Fact 7.B.4, on the indicator in this expectation, $\sigma^1, \sigma^2 \in \widetilde{T}$. Moreover
$$|R(\sigma^1,\sigma^2)| \le |R(\sigma^1_{\sim i,j}, \sigma^2_{\sim i,j})| + \frac{|\sigma^1_i||\sigma^2_i| + |\sigma^1_j||\sigma^1_j|}{N} \le 3N^{-2/5}.$$
Thus
$$\begin{aligned}|\widehat{X}^{(2)}_{i,j}| \le 2N^2 \Big\langle & \mathbf{1}[\sigma^1 \not\in \widetilde{T}, |R(\sigma^1,\sigma^2)| \le 3N^{-2/5}] \\ & \left(e^{H_{N,\sim 2}(\sigma^1) - N\xi_{\sim 2}(1)/2} + 1\right) \left(e^{H_{N,\sim 2}(\sigma^2) - N\xi_{\sim 2}(1)/2} + 1\right) \Big\rangle_2\end{aligned}$$
and Corollary 7.8.12 implies $\mathbb{E}_{\sim 2} |\widehat{X}^{(2)}_{i,j}| \le e^{-cN^{1/5}}$. □

Bibliography

[AB13] Antonio Auffinger and Gérard Ben Arous. Complexity of random smooth functions on the high-dimensional sphere. *Ann. Probab.*, 41(6):4214–4247, 2013.

[Abb18] Emmanuel Abbe. Community detection and stochastic block models: recent developments. *J. Mach. Learn. Res.*, 18(177):1–86, 2018.

[ABČ13] Antonio Auffinger, Gérard Ben Arous, and Jiří Černỳ. Random matrices and complexity of spin glasses. *Comm. Pure Appl. Math.*, 66(2):165–201, 2013.

[ABDM23] Erik Aurell, Jean Barbier, Aurélien Decelle, and Roberto Mulet. The mighty force: statistical inference and high-dimensional statistics. In *Spin Glass Theory and Far Beyond: Replica Symmetry Breaking After 40 Years*, pages 423–451. 2023.

[ABG+12] Elena Agliari, Adriano Barra, Andrea Galluzzi, Francesco Guerra, and Francesco Moauro. Multitasking associative networks. *Phys. Rev. Lett.*, 109(26):268101, 2012.

[ABL22] Antonio Auffinger, Gérard Ben Arous, and Zhehua Li. Sharp complexity asymptotics and topological trivialization for the (p, k) spiked tensor model. *J. Math. Phys.*, 63(4):043303, 2022.

[ABXY24] Arka Adhikari, Christian Brennecke, Changji Xu, and Horng-Tzer Yau. Spectral gap estimates for mixed p-spin models at high temperature. *Probab. Theory Rel. Fields*, pages 1–29, 2024.

[AC08] Dimitris Achlioptas and Amin Coja-Oghlan. Algorithmic barriers from phase transitions. In *Proc. 49th FOCS*, pages 793–802, 2008.

[AC15] Antonio Auffinger and Wei-Kuo Chen. The Parisi formula has a unique minimizer. *Comm. Math. Phys.*, 335(3):1429–1444, 2015.

[AC17] Antonio Auffinger and Wei-Kuo Chen. Parisi formula for the ground state energy in the mixed p-spin model. *Ann. Probab.*, 45(6b):4617–4631, 2017.

[AC18] Antonio Auffinger and Wei-Kuo Chen. On the energy landscape of spherical spin glasses. *Adv. Math.*, 330:553–588, 2018.

[AEGP23] Ahmed El Alaoui, Ronen Eldan, Reza Gheissari, and Arianna Piana. Fast relaxation of the random field Ising dynamics. *arXiv:2311.06171*, 2023.

[AEK17a] Oskari Ajanki, László Erdős, and Torben Krüger. Singularities of solutions to quadratic vector equations on the complex upper half-plane. *Comm. Pure. Appl. Math.*, 70(9):1672–1705, 2017.

[AEK17b] Oskari H Ajanki, László Erdős, and Torben Krüger. Universality for general Wigner-type matrices. *Probab. Theory Rel. Fields*, 169:667–727, 2017.

[AEK19a] Oskari Ajanki, László Erdős, and Torben Krüger. *Quadratic Vector Equations on Complex Upper Half-Plane*, volume 261. American Mathematical Society, 2019.

[AEK19b] Oskari H Ajanki, László Erdős, and Torben Krüger. Stability of the matrix Dyson equation and random matrices with correlations. *Probab. Theory Rel. Fields*, 173:293–373, 2019.

[AEK20] Johannes Alt, László Erdős, and Torben Krüger. The Dyson equation with linear self-energy: spectral bands, edges and cusps. *Doc. Math.*, 25:1421–1539, 2020.

[AEKN19] Johannes Alt, László Erdös, Torben H Krüger, and Yuriy Nemish. Location of the spectrum of Kronecker random matrices. *Ann. Inst. Henri Poincaré Probab. Stat.*, 55(2), 2019.

[AG20] Antonio Auffinger and Julian Gold. The number of saddles of the spherical p-spin model. *arXiv:2007.09269*, 2020.

[AG24] Ahmed El Alaoui and Jason Gaitonde. Bounds on the covariance matrix of the Sherrington–Kirkpatrick model. *Electron. Comm. Probab.*, 29:1–13, 2024.

[AGZ10] Greg W Anderson, Alice Guionnet, and Ofer Zeitouni. *An Introduction to Random Matrices*, volume 118. Cambridge University Press, 2010.

[Aid98] Shigeki Aida. Uniform positivity improving property, Sobolev inequalities, and spectral gaps. *J. Func. Anal.*, 158(1):152–185, 1998.

[AJK+21a] Nima Anari, Vishesh Jain, Frederic Koehler, Huy Tuan Pham, and Thuy-Duong Vuong. Entropic independence I: modified log-Sobolev inequalities for fractionally log-concave distributions and high-temperature Ising models. *arXiv:2106.04105*, 2021.

[AJK+21b] Nima Anari, Vishesh Jain, Frederic Koehler, Huy Tuan Pham, and Thuy-Duong Vuong. Entropic independence II: optimal sampling and concentration via restricted modified log-Sobolev inequalities. *arXiv:2111.03247*, 2021.

[AJK+22] Nima Anari, Vishesh Jain, Frederic Koehler, Huy Tuan Pham, and Thuy-Duong Vuong. Entropic independence: optimal mixing of down-up random walks. In *Proc. 54th STOC*, pages 1418–1430, 2022.

[AJK+24] Nima Anari, Vishesh Jain, Frederic Koehler, Huy Tuan Pham, and Thuy-Duong Vuong. Universality of spectral independence with applications to fast mixing in spin glasses. In *Proc. 35th SODA*, pages 5029–5056, 2024.

[AKV24] Nima Anari, Frederic Koehler, and Thuy-Duong Vuong. Trickle-down in localization schemes and applications. In *Proc. 56th STOC*, pages 1094–1105, 2024.

[AL20] Vedat Levi Alev and Lap Chi Lau. Improved analysis of higher order random walks and applications. In *Proc. 52nd STOC*, pages 1198–1211, 2020.

[ALO21] Nima Anari, Kuikui Liu, and Shayan Oveis Gharan. Spectral independence in high-dimensional expanders and applications to the hardcore model. *SIAM J. Comp.*, pages FOCS20–1, 2021.

[ALOV24] Nima Anari, Kuikui Liu, Shayan Oveis Gharan, and Cynthia Vinzant. Log-concave polynomials II: high-dimensional walks and an FPRAS for counting bases of a matroid. *Ann. Math.*, 199(1):259–299, 2024.

[ALR87] Michael Aizenman, Joel L Lebowitz, and David Ruelle. Some rigorous results on the Sherrington–Kirkpatrick spin glass model. *Comm. Math. Phys.*, 112:3–20, 1987.

[ALS22a] Emmanuel Abbe, Shuangping Li, and Allan Sly. Binary perceptron: efficient algorithms can find solutions in a rare well-connected cluster. In *Proc. 54th STOC*, pages 860–873, 2022.

[ALS22b] Emmanuel Abbe, Shuangping Li, and Allan Sly. Proof of the contiguity conjecture and lognormal limit for the symmetric perceptron. In *Proc. 62nd FOCS*, pages 327–338, 2022.

[Alt23] Dylan J Altschuler. Critical window of the symmetric perceptron. *Electron. J. Probab.*, 28:1–28, 2023.

[AM22] Ahmed El Alaoui and Andrea Montanari. An information-theoretic view of stochastic localization. *IEEE Trans. Inform. Theory*, 68(11):7423–7426, 2022.

[AMS21] Ahmed El Alaoui, Andrea Montanari, and Mark Sellke. Optimization of mean-field spin glasses. *Ann. Probab.*, 49(6):2922–2960, 2021.

[AMS22] Ahmed El Alaoui, Andrea Montanari, and Mark Sellke. Sampling from the Sherrington-Kirkpatrick Gibbs measure via algorithmic stochastic localization. In *Proc. 63rd FOCS*, pages 323–334, 2022.

[AMS23a] Ahmed El Alaoui, Andrea Montanari, and Mark Sellke. Local algorithms for maximum cut and minimum bisection on locally treelike regular graphs of large degree. *Rand. Struct. Alg.*, 2023.

[AMS23b] Ahmed El Alaoui, Andrea Montanari, and Mark Sellke. Sampling from mean-field Gibbs measures via diffusion processes. *arXiv:2310.08912*, 2023.

[AMS23c] Antonio Auffinger, Andrea Montanari, and Eliran Subag. Optimization of random high-dimensional functions: structure and algorithms. In *Spin Glass Theory and Far Beyond: Replica Symmetry Breaking After 40 Years*, pages 609–633. 2023.

[AMS25] Ahmed El Alaoui, Andrea Montanari, and Mark Sellke. Shattering in pure spherical spin glasses. *Comm. Math. Phys.*, 406(5):1–36, 2025.

[APZ19] Benjamin Aubin, Will Perkins, and Lenka Zdeborová. Storage capacity in symmetric binary perceptrons. *J. Phys. A*, 52(29):294003, 2019.

[AS22] Ahmed El Alaoui and Mark Sellke. Algorithmic pure states for the negative spherical perceptron. *J. Stat. Phys.*, 189(2):27, 2022.

[ASS03] Michael Aizenman, Robert Sims, and Shannon L Starr. Extended variational principle for the Sherrington–Kirkpatrick spin-glass model. *Phys. Rev. B*, 68(21):214403, 2003.

[AT09] Robert J Adler and Jonathan E Taylor. *Random Fields and Geometry.* Springer, 2009.

[AT24] Dylan J Altschuler and Konstantin Tikhomirov. A note on the capacity of the binary perceptron. *arXiv:2401.15092*, 2024.

[BAKZ24] Damien Barbier, Ahmed El Alaoui, Florent Krzakala, and Lenka Zdeborová. On the atypical solutions of the symmetric binary perceptron. *J. Phys. A*, 57(19):195202, 2024.

[BB19] Roland Bauerschmidt and Thierry Bodineau. A very simple proof of the LSI for high temperature spin systems. *J. Func. Anal.*, 276(8):2582–2588, 2019.

[BBCG08] Dominique Bakry, Franck Barthe, Patrick Cattiaux, and Arnaud Guillin. A simple proof of the Poincaré inequality for a large class of probability measures including the log-concave case. *Electron. Comm. Probab.*, 13:60–66, 2008.

[BBM23] Gérard Ben Arous, Paul Bourgade, and Benjamin McKenna. Exponential growth of random determinants beyond invariance. *Probab. Math. Phys.*, 3(4):731–789, 2023.

[BBM24] Gérard Ben Arous, Paul Bourgade, and Benjamin McKenna. Landscape complexity beyond invariance and the elastic manifold. *Comm. Pure Appl. Math.*, 77(2):1302–1352, 2024.

[BBvH23] Afonso S Bandeira, March T Boedihardjo, and Ramon van Handel. Matrix concentration inequalities and free probability. *Invent. Math.*, 234(1):419–487, 2023.

[BC16] Victor Bapst and Amin Coja-Oghlan. The condensation phase transition in the regular k-SAT model. *Approx. Rand. Combin. Optim. Algor. Techniques*, 2016.

[BCE+22] Krishna Balasubramanian, Sinho Chewi, Murat A Erdogdu, Adil Salim, and Shunshi Zhang. Towards a theory of non-log-concave sampling: first-order stationarity guarantees for Langevin Monte Carlo. In *Proc. 2022 COLT*, pages 2896–2923, 2022.

[BCG+12] William Bialek, Andrea Cavagna, Irene Giardina, Thierry Mora, Edmondo Silvestri, Massimiliano Viale, and Aleksandra M Walczak. Statistical mechanics for natural flocks of birds. *Proc. Natl. Acad. Sci.*, 109(13):4786–4791, 2012.

[BCH+16] Victor Bapst, Amin Coja-Oghlan, Samuel Hetterich, Felicia Raßmann, and Dan Vilenchik. The condensation phase transition in random graph coloring. *Comm. Math. Phys.*, 341:543–606, 2016.

[BCKM98] Jean-Philippe Bouchaud, Leticia F Cugliandolo, Jorge Kurchan, and Marc Mézard. Out of equilibrium dynamics in spin-glasses and other glassy systems. *Spin Glass. Rand. Fields*, pages 161–223, 1998.

[BCMT15] Adriano Barra, Pierluigi Contucci, Emanuele Mingione, and Daniele Tantari. Multi-species mean field spin glasses. Rigorous results. In *Ann. Henri Poincaré*, volume 16, pages 691–708, 2015.

[BČNS22] David Belius, Jiří Černỳ, Shuta Nakajima, and Marius A Schmidt. Triviality of the geometry of mixed p-spin spherical Hamiltonians with external field. *J. Stat. Phys.*, 186(1):12, 2022.

[BD00] Rajendra Bhatia and Chandler Davis. A better bound on the variance. *Amer. Math. Monthly*, 107(4):353–357, 2000.

[BDG01] Gérard Ben Arous, Amir Dembo, and Alice Guionnet. Aging of spherical spin glasses. *Probab. Theory Rel. Fields*, 120:1–67, 2001.

[BDG06] Gérard Ben Arous, Amir Dembo, and Alice Guionnet. Cugliandolo-Kurchan equations for dynamics of spin-glasses. *Probab. Theory Rel. Fields*, 136(4):619–660, 2006.

[BFK21] Gérard Ben Arous, Yan V Fyodorov, and Boris A Khoruzhenko. Counting equilibria of large complex systems by instability index. *Proc. Natl. Acad. Sci.*, 118(34):e2023719118, 2021.

[BFP97] Alain Barrat, Silvio Franz, and Giorgio Parisi. Temperature evolution and bifurcations of metastable states in mean-field spin glasses, with connections with structural glasses. *J. Phys. A*, 30(16):5593, 1997.

[BG95] Gérard Ben Arous and Alice Guionnet. Large deviations for Langevin spin glass dynamics. *Probab. Theory Rel. Fields*, 102:455–509, 1995.

[BG97a] Gérard Ben Arous and Alice Guionnet. Large deviations for Wigner's law and Voiculescu's non-commutative entropy. *Probab. Theory Rel. Fields*, 108(4):517–542, 1997.

[BG97b] Gérard Ben Arous and Alice Guionnet. Symmetric Langevin spin glass dynamics. *Ann. Probab.*, 25(3):1367–1422, 1997.

[BGG10] Adriano Barra, Giuseppe Genovese, and Francesco Guerra. The replica symmetric approximation of the analogical neural network. *J. Stat. Phys.*, 140:784–796, 2010.

[BGJ20] Gérard Ben Arous, Reza Gheissari, and Aukosh Jagannath. Bounding flows for spherical spin glass dynamics. *Comm. Math. Phys.*, 373(3):1011–1048, 2020.

[BGL01] Sergey G Bobkov, Ivan Gentil, and Michel Ledoux. Hypercontractivity of Hamilton–Jacobi equations. *J. Math. Pure. Appl.*, 80(7):669–696, 2001.

[BGT10] Mohsen Bayati, David Gamarnik, and Prasad Tetali. Combinatorial approach to the interpolation method and scaling limits in sparse random graphs. In *Proc. 42nd STOC*, pages 105–114, 2010.

[BGZ25] Antonio Blanca, Reza Gheissari, and Xusheng Zhang. Mean-field Potts and random-cluster dynamics from high-entropy initializations. In *Proc. 36th SODA*, pages 5434–5467, 2025.

[BH22] Guy Bresler and Brice Huang. The algorithmic phase transition of random k-SAT for low degree polynomials. In *Proc. 62nd FOCS*, pages 298–309, 2022.

[BJ24] Gérard Ben Arous and Aukosh Jagannath. Shattering versus metastability in spin glasses. *Comm. Pure Appl. Math.*, 77(1):139–176, 2024.

[BL16] Jinho Baik and Ji Oon Lee. Fluctuations of the free energy of the spherical Sherrington–Kirkpatrick model. *J. Stat. Phys.*, 165:185–224, 2016.

[BL20] Jinho Baik and Ji Oon Lee. Free energy of bipartite spherical Sherrington–Kirkpatrick model. *Ann. Inst. Henri Poincaré Probab. Stat.*, 56(4):2897–2934, 2020.

[BM80] Alan J Bray and Michael A Moore. Metastable states in spin glasses. *J. Phys. C*, 13(19):L469, 1980.

[BM11] Mohsen Bayati and Andrea Montanari. The dynamics of message passing on dense graphs, with applications to compressed sensing. *IEEE Trans. Inform. Theory*, 57(2):764–785, 2011.

[BMMN19] Gérard Ben Arous, Song Mei, Andrea Montanari, and Mihai Nica. The landscape of the spiked tensor model. *Comm. Pure Appl. Math.*, 72(11):2282–2330, 2019.

[BMN20] Raphael Berthier, Andrea Montanari, and Phan-Minh Nguyen. State evolution for approximate message passing with non-separable functions. *Inform. Inference*, 9(1):33–79, 2020.

[BMN23] Jean-Philippe Bouchaud, Matteo Marsili, and Jean-Pierre Nadal. Application of spin glass ideas in social sciences, economics and finance. In *Spin Glass Theory and Far Beyond: Replica Symmetry Breaking After 40 Years*, pages 561–579. 2023.

[BNN24] Guy Bresler, Dheeraj Nagaraj, and Eshaan Nichani. Metastable mixing of Markov chains: efficiently sampling low temperature exponential random graphs. *Ann. Appl. Probab.*, 34(1a):517–554, 2024.

[BNSX22] Erwin Bolthausen, Shuta Nakajima, Nike Sun, and Changji Xu. Gardner formula for Ising perceptron models at small densities. In *Proc. 2022 COLT*, pages 1787–1911, 2022.

[Bol14] Erwin Bolthausen. An iterative construction of solutions of the TAP equations for the Sherrington–Kirkpatrick model. *Comm. Math. Phys.*, 325(1):333–366, 2014.

[Bol19] Erwin Bolthausen. A Morita type proof of the replica-symmetric formula for SK. *Stat. Mech. Class. Disord. Syst.*, page 63, 2019.

[Bor76] Christer Borell. Gaussian Radon measures on locally convex spaces. *Math. Scandin.*, 38(2):265–284, 1976.

[BS98] Zhi-Dong Bai and Jack W Silverstein. No eigenvalues outside the support of the limiting spectral distribution of large-dimensional sample covariance matrices. *Ann. Probab.*, 26(1):316–345, 1998.

[BS20] Nikhil Bansal and Joel H Spencer. On-line balancing of random inputs. *Rand. Struct. Alg.*, 57(4):879–891, 2020.

[BS22a] Erik Bates and Youngtak Sohn. Crisanti-Sommers formula and simultaneous symmetry breaking in multi-species spherical spin glasses. *Comm. Math. Phys.*, 394(3):1101–1152, 2022.

[BS22b] Erik Bates and Youngtak Sohn. Free energy in multi-species mixed p-spin spherical models. *Electron. J. Probab.*, 27:1–75, 2022.

[BSXY24] Christian Brennecke, Adrien Schertzer, Changji Xu, and Horng-Tzer Yau. The two point function of the SK model without external field at high temperature. *Probab. Math. Phys*, 5:131–175, 2024.

[BSZ20] Gérard Ben Arous, Eliran Subag, and Ofer Zeitouni. Geometry and temperature chaos in mixed spherical spin glasses at low temperature: the perturbative regime. *Comm. Pure. Appl. Math.*, 73(8):1732–1828, 2020.

[BW87] Joseph D Bryngelson and Peter G Wolynes. Spin glasses and the statistical mechanics of protein folding. *Proc. Natl. Acad. Sci.*, 84(21):7524–7528, 1987.

[BXY23] Christian Brennecke, Changji Xu, and Horng-Tzer Yau. Operator norm bounds on the correlation matrix of the SK model at high temperature. *arXiv:2307.12535*, 2023.

[BY05] ZD Bai and Jianfeng Yao. On the convergence of the spectral empirical process of Wigner matrices. *Bernoulli*, 11(6):1059–1092, 2005.

[BY22] Christian Brennecke and Horng-Tzer Yau. The replica symmetric formula for the SK model revisited. *J. Math. Phys.*, 63(7), 2022.

[CCL+23] Sitan Chen, Sinho Chewi, Jerry Li, Yuanzhi Li, Adil Salim, and Anru R Zhang. Sampling is as easy as learning the score: theory for diffusion models with minimal data assumptions. In *Proc. 11th ICLR*, 2023.

[CCM21] Michael Celentano, Chen Cheng, and Andrea Montanari. The high-dimensional asymptotics of first order methods with random data. *arXiv:2112.07572*, 2021.

[CDL+12] Paul Cuff, Jian Ding, Oren Louidor, Eyal Lubetzky, Yuval Peres, and Allan Sly. Glauber dynamics for the mean-field Potts model. *J. Stat. Phys.*, 149:432–477, 2012.

[CE22] Yuansi Chen and Ronen Eldan. Localization schemes: a framework for proving mixing bounds for Markov chains. In *Proc. 63rd FOCS*, pages 110–122, 2022.

[CEJ+18] Amin Coja-Oghlan, Charilaos Efthymiou, Nor Jaafari, Mihyun Kang, and Tobias Kapetanopoulos. Charting the replica symmetric phase. *Comm. Math. Phys.*, 359:603–698, 2018.

[Cel24] Michael Celentano. Sudakov-Fernique post-AMP, and a new proof of the local convexity of the TAP free energy. *Ann. Probab.*, 52(3):923–954, 2024.

[Ces12] Lamberto Cesari. *Optimization—Theory and Applications: Problems with Ordinary Differential Equations*, volume 17. Springer, 2012.

[CFLM23] Michael Celentano, Zhou Fan, Licong Lin, and Song Mei. Mean-field variational inference with the TAP free energy: geometric and statistical properties in linear models. *arXiv:2311.08442*, 2023.

[CFM23] Michael Celentano, Zhou Fan, and Song Mei. Local convexity of the TAP free energy and AMP convergence for $\mathbb{Z}_2$-synchronization. *Ann. Stat.*, 51(2):519–546, 2023.

[CGG07] Patrick Cattiaux, Ivan Gentil, and Arnaud Guillin. Weak logarithmic Sobolev inequalities and entropic convergence. *Probab. Theory Rel. Fields*, 139:563–603, 2007.

[CGG+16] Jin-Yi Cai, Andreas Galanis, Leslie Ann Goldberg, Heng Guo, Mark Jerrum, Daniel Štefankovič, and Eric Vigoda. #BIS-hardness for 2-spin systems on bipartite bounded degree graphs in the tree non-uniqueness region. *J. Comp. Sys. Sci.*, 82(5):690–711, 2016.

[CGM12] Prasad Chebolu, Leslie Ann Goldberg, and Russell Martin. The complexity of approximately counting stable matchings. *Theor. Comp. Sci.*, 437:35–68, 2012.

[CGPR19] Wei-Kuo Chen, David Gamarnik, Dmitry Panchenko, and Mustazee Rahman. Suboptimality of local algorithms for a class of max-cut problems. *Ann. Probab.*, 47(3):1587–1618, 2019.

[Cha14] Sourav Chatterjee. *Superconcentration and Related Topics*, volume 15. Springer, 2014.

[Che13] Wei-Kuo Chen. The Aizenman–Sims–Starr scheme and Parisi formula for mixed p-spin spherical models. *Electron. J. Probab.*, 18, 2013.

[Che23a] Wei-Kuo Chen. A gaussian convexity for logarithmic moment generating functions. *arXiv:2311.08351*, 2023.

[Che23b] Sinho Chewi. *Log-Concave Sampling.* Book draft available at https://chewisinho.github.io, 2023.

[CHKW23] Sergio Caracciolo, Alexander Hartmann, Scott Kirkpatrick, and Martin Weigel. Simulated annealing, optimization, searching for ground states. In *Spin Glass Theory and Far Beyond: Replica Symmetry Breaking After 40 Years*, pages 1–20. 2023.

[CHL18] Wei-Kuo Chen, Madeline Handschy, and Gilad Lerman. On the energy landscape of the mixed even p-spin model. *Probab. Theory Rel. Fields*, 171(1-2):53–95, 2018.

[CHM$^+$15] Anna Choromanska, Mikael Henaff, Michael Mathieu, Gérard Ben Arous, and Yann LeCun. The loss surfaces of multilayer networks. In *Artif. Intel. Stat.*, pages 192–204, 2015.

[CHS93] Andrea Crisanti, Heinz Horner, and H-J Sommers. The spherical p-spin interaction spin-glass model: the dynamics. *Zeit. Phys. B*, 92:257–271, 1993.

[CIS76] Boris S Cirel'son, Ildar A Ibragimov, and Vladimir N Sudakov. Norms of gaussian sample functions. In *Proc. 3rd Japan—USSR Sym. Probab. Theory*, pages 20–41, 1976.

[CK93] Leticia F Cugliandolo and Jorge Kurchan. Analytical solution of the off-equilibrium dynamics of a long-range spin-glass model. *Phys. Rev. Lett.*, 71(1):173, 1993.

[CK94] Leticia F. Cugliandolo and Jorge Kurchan. On the out-of-equilibrium relaxation of the Sherrington–Kirkpatrick model. *J. Phys. A*, 27(17):5749, 1994.

[CKLDP97] Leticia F Cugliandolo, Jorge Kurchan, Pierre Le Doussal, and Luca Peliti. Glassy behaviour in disordered systems with nonrelaxational dynamics. *Phys. Rev. Lett.*, 78(2):350, 1997.

[CKM20] Amin Coja-Oghlan, Tobias Kapetanopoulos, and Noela Müller. The replica symmetric phase of random constraint satisfaction problems. *Combin. Probab. Comp.*, 29(3):346–422, 2020.

[CKPZ17] Amin Coja-Oghlan, Florent Krzakala, Will Perkins, and Lenka Zdeborová. Information-theoretic thresholds from the cavity method. In *Proc. 49th STOC*, pages 146–157, 2017.

[CLA$^+$21] Sinho Chewi, Chen Lu, Kwangjun Ahn, Xiang Cheng, Thibaut Le Gouic, and Philippe Rigollet. Optimal dimension dependence of the Metropolis-adjusted Langevin algorithm. In *Proc. 2021 COLT*, pages 1260–1300, 2021.

[CLR03] Andrea Crisanti, Luca Leuzzi, and Tommaso Rizzo. The complexity of the spherical p-spin spin glass model, revisited. *Eur. Phys. J. B*, 36(1):129–136, 2003.

[CLR05] Andrea Crisanti, Luca Leuzzi, and Tommaso Rizzo. Complexity in mean-field spin-glass models: Ising p-spin. *Phys. Rev. B*, 71(9):094202, 2005.

[CN95] Francis Comets and Jacques Neveu. The Sherrington–Kirkpatrick model of spin glasses and stochastic calculus: the high temperature case. *Comm. Math. Phys.*, 166:549–564, 1995.

[Cov65] Thomas M Cover. Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition. *IEEE Trans. Electron. Comp.*, 3(14):326–334, 1965.

[CS92] Andrea Crisanti and H-J Sommers. The spherical p-spin interaction spin glass model: the statics. *Zeit. Phys. B*, 87(3):341–354, 1992.

[CS95] A Crisanti and H-J Sommers. Thouless–Anderson–Palmer approach to the spherical p-spin spin glass model. *J. Phys. I*, 5(7):805–813, 1995.

[CS17] Wei-Kuo Chen and Arnab Sen. Parisi formula, disorder chaos and fluctuation for the ground state energy in the spherical mixed p-spin models. *Comm. Math. Phys.*, 350(1):129–173, 2017.

[CS21] Sourav Chatterjee and Leila Sloman. Average Gromov hyperbolicity and the Parisi ansatz. *Adv. Math.*, 376:107417, 2021.

[CWZZ24] Xiang Cheng, Bohan Wang, Jingzhao Zhang, and Yusong Zhu. Fast conditional mixing of MCMC algorithms for non-log-concave distributions. *Proc. 2024 NeurIPS*, 36, 2024.

[CZ17] Li-Juan Cheng and Kun Zhang. Reflecting diffusion processes on manifolds carrying geometric flow. *J. Theor. Probab.*, 30:1334–1368, 2017.

[Dal17] Arnak S Dalalyan. Theoretical guarantees for approximate sampling from smooth and log-concave densities. *J. Royal Stat. Soc. B*, 79(3):651–676, 2017.

[dAT78] Jairo RL de Almeida and David J Thouless. Stability of the Sherrington–Kirkpatrick solution of a spin glass model. *J. Phys. A*, 11(5):983, 1978.

[DCWY19] Raaz Dwivedi, Yuansi Chen, Martin J Wainwright, and Bin Yu. Log-concave sampling: Metropolis-Hastings algorithms are fast. *J. Mach. Learn. Res.*, 20(183):1–42, 2019.

[Der81] Bernard Derrida. Random-energy model: an exactly solvable model of disordered systems. *Phys. Rev. B*, 24(5):2613, 1981.

[DEZ15] Jian Ding, Ronen Eldan, and Alex Zhai. On multiple peaks and moderate deviations for the supremum of a gaussian field. *Ann. Probab.*, 43(6):3468–3493, 2015.

[DF91] Martin Dyer and Alan Frieze. Computing the volume of convex bodies: a case where randomness provably helps. *Probab. Combin. Appl.*, 44(123-170):0754–68052, 1991.

[DFK91] Martin Dyer, Alan Frieze, and Ravi Kannan. A random polynomial-time algorithm for approximating the volume of convex bodies. *J. ACM*, 38(1):1–17, 1991.

[DFM94] Vik Dotsenko, Silvio Franz, and Marc Mezard. Partial annealing and overfrustration in disordered systems. *J. Phys. A*, 27(7):2351, 1994.

[DG21] Amir Dembo and Reza Gheissari. Diffusions interacting through a random matrix: universality via stochastic Taylor expansion. *Probab. Theory Rel. Fields*, 180:1057–1097, 2021.

[DGM07] Amir Dembo, Alice Guionnet, and Christian Mazza. Limiting dynamics for spherical models of spin glasses at high temperature. *J. Stat. Phys.*, 126:781–815, 2007.

[DKMZ11] Aurelien Decelle, Florent Krzakala, Cristopher Moore, and Lenka Zdeborová. Asymptotic analysis of the stochastic block model for modular networks and its algorithmic applications. *Phys. Rev. E*, 84(6):066106, 2011.

[DLZ21] Amir Dembo, Eyal Lubetzky, and Ofer Zeitouni. Universality for Langevin-like spin glass dynamics. *Ann. Appl. Probab.*, 31(6):2864–2880, 2021.

[DM17] Alain Durmus and Eric Moulines. Nonasymptotic convergence analysis for the unadjusted Langevin algorithm. *Ann. Appl. Probab.*, 27(3):1551–1587, 2017.

[DMS17] Amir Dembo, Andrea Montanari, and Subhabrata Sen. Extremal cuts of sparse random graphs. *Ann. Probab.*, 45(2):1190–1217, 2017.

[DS18] Jian Ding and Nike Sun. Capacity lower bound for the ising perceptron. *arXiv preprint arXiv:1809.07742*, 2018.

[DS20] Amir Dembo and Eliran Subag. Dynamics for spherical spin glasses: disorder dependent initial conditions. *J. Stat. Phys.*, 181:465–514, 2020.

[DSC96] Persi Diaconis and Laurent Saloff-Coste. Logarithmic Sobolev inequalities for finite Markov chains. *Ann. Appl. Probab.*, 6(3), 1996.

[DZ95] Jean-Dominique Deuschel and Ofer Zeitouni. Limiting curves for IID records. *Ann. Probab.*, pages 852–878, 1995.

[EA75] Samuel F Edwards and Phil W Anderson. Theory of spin glasses. *J. Phys. F*, 5(5):965, 1975.

[EKS19] László Erdős, Torben Krüger, and Dominik Schröder. Random matrices with slow correlation decay. In *Forum Math. Sigma*, volume 7, page e8, 2019.

[EKZ22] Ronen Eldan, Frederic Koehler, and Ofer Zeitouni. A spectral condition for spectral gap: fast mixing in high-temperature Ising models. *Probab. Theory Rel. Fields*, 182(3):1035–1051, 2022.

[Eld13] Ronen Eldan. Thin shell implies spectral gap up to polylog via a stochastic localization scheme. *Geom. Func. Anal.*, 23(2):532–569, 2013.

[Eld20a] Ronen Eldan. A simple approach to chaos for p-spin models. *J. Stat. Phys.*, 181(4):1266–1276, 2020.

[Eld20b] Ronen Eldan. Taming correlations through entropy-efficient measure decompositions with applications to mean-field approximation. *Probab. Theory Rel. Fields*, 176(3-4):737–755, 2020.

[Eld22] Ronen Eldan. Analysis of high-dimensional distributions using pathwise methods. In *Proc. Int. Cong. Math*, volume 6, pages 4246–4270, 2022.

[Fer75] Xavier Fernique. Regularite des trajectoires des fonctions aleatoires gaussiennes. *Ecole Probab. Saint-Flour IV-1974*, pages 1–96, 1975.

[FFRT21] Giampaolo Folena, Silvio Franz, and Federico Ricci-Tersenghi. Gradient descent dynamics in the mixed p-spin spherical model: finite-size simulations and comparison with mean-field integration. *J. Stat. Mech. Theor. Exper.*, 2021(3):033302, 2021.

[FK16] Yan V Fyodorov and Boris A Khoruzhenko. Nonlinear analogue of the May–Wigner instability transition. *Proc. Natl. Acad. Sci.*, 113(25):6827–6832, 2016.

[FKS87a] Ya V Fyodorov, I Ya Korenblit, and EF Shender. Antiferromagnetic Ising spin glass. *J. Phys. C*, 20(12):1835, 1987.

[FKS87b] Ya V Fyodorov, I Ya Korenblit, and EF Shender. Phase transitions in frustrated metamagnets. *Europhys. Lett.*, 4(7):827, 1987.

[FLD14] Yan V Fyodorov and Pierre Le Doussal. Topology trivialization and large deviations for the minimum in the simplest random optimization. *J. Stat. Phys.*, 154(1-2):466–490, 2014.

[FLDRT18] Yan V Fyodorov, Pierre Le Doussal, Alberto Rosso, and Christophe Texier. Exponential number of equilibria and depinning threshold for a directed polymer in a random potential. *Ann. Phys.*, 397:1–64, 2018.

[FMM21] Zhou Fan, Song Mei, and Andrea Montanari. TAP free energy, spin glasses and variational inference. *Ann. Probab.*, 49(1), 2021.

[FSU21] Silvio Franz, Antonio Sclocchi, and Pierfrancesco Urbani. Surfing on minima of isostatic landscapes: avalanches and unjamming transition. *J. Stat. Mech. Theor. Exper.*, 2021(2):023208, 2021.

[Fyo04] Yan V Fyodorov. Complexity of random energy landscapes, glass transition, and absolute value of the spectral determinant of random matrices. *Phys. Rev. Lett.*, 92(24):240601, 2004.

[Fyo15] Yan V Fyodorov. High-dimensional random fields and random matrix theory. *Markov Proc. and Rel. Fields*, 21(3):483–518, 2015.

[Fyo16] Yan V Fyodorov. Topology trivialization transition in random non-gradient autonomous ODEs on a sphere. *J. Stat. Mech. Theor. Exper.*, 2016(12):124003, 2016.

[Gam21] David Gamarnik. The overlap gap property: a topological barrier to optimizing over random structures. *Proc. Natl. Acad. Sci.*, 118(41), 2021.

[Gar87] Elizabeth Gardner. Maximum storage capacity in neural networks. *Europhys. Lett.*, 4(4):481, 1987.

[Gar88] Elizabeth Gardner. The space of interactions in neural network models. *J. Phys. A*, 21(1):257, 1988.

[Gar07] John Garnett. *Bounded Analytic Functions*, volume 236. Springer, 2007.

[Gar17] Xavier Garcia. On the number of equilibria with a given number of unstable directions. *arXiv:1709.04021*, 2017.

[GD88] Elizabeth Gardner and Bernard Derrida. Optimal storage properties of neural network models. *J. Phys. A*, 21(1):271, 1988.

[GDKZ24] Davide Ghio, Yatin Dandi, Florent Krzakala, and Lenka Zdeborová. Sampling with flows, diffusion, and autoregressive neural networks from a spin-glass perspective. *Proc. Natl. Acad. Sci.*, 121(27):e2311810121, 2024.

[GG98] Stefano Ghirlanda and Francesco Guerra. General properties of overlap probability distributions in disordered spin systems. Towards Parisi ultrametricity. *J. Phys. A*, 31(46):9149–9155, 1998.

[GGLZ23] Marylou Gabrié, Surya Ganguli, Carlo Lucibello, and Riccardo Zecchina. Neural networks: From the perceptron to deep nets. In *Spin Glass Theory and Far Beyond: Replica Symmetry Breaking After 40 Years*, pages 477–497. 2023.

[GJ12] Leslie Ann Goldberg and Mark Jerrum. Approximating the partition function of the ferromagnetic Potts model. *J. ACM*, 59(5):1–31, 2012.

[GJ19] Reza Gheissari and Aukosh Jagannath. On the spectral gap of spherical spin glass dynamics. *Ann. Inst. Henri Poincaré Probab. Stat.*, 55(2):756–776, 2019.

[GJ21] David Gamarnik and Aukosh Jagannath. The overlap gap property and approximate message passing algorithms for p-spin models. *Ann. Probab.*, 49(1):180–205, 2021.

[GJK23] David Gamarnik, Aukosh Jagannath, and Eren C Kızıldağ. Shattering in the Ising pure p-spin model. *arXiv:2307.07461*, 2023.

[GJW20] David Gamarnik, Aukosh Jagannath, and Alexander S. Wein. Low-degree hardness of random optimization problems. In *Proc. 61st FOCS*, pages 131–140, 2020.

[GK23] David Gamarnik and Eren C Kızıldağ. Algorithmic obstructions in the random number partitioning problem. *Ann. Appl. Probab.*, 33(6B):5497–5563, 2023.

[GKPX22] David Gamarnik, Eren C Kızıldağ, Will Perkins, and Changji Xu. Algorithms and barriers in the symmetric binary perceptron model. In *Proc. 63rd FOCS*, pages 576–587, 2022.

[GKPX23] David Gamarnik, Eren C Kızıldağ, Will Perkins, and Changji Xu. Geometric barriers for stable and online algorithms for discrepancy minimization. In *Proc. 2023 COLT*, pages 3231–3263, 2023.

[GLR18] Rong Ge, Holden Lee, and Andrej Risteski. Simulated tempering Langevin Monte Carlo II: an improved proof using soft Markov chain decomposition. *arXiv:1812.00793*, 2018.

[GLZ+18] Ruiqi Gao, Yang Lu, Junpei Zhou, Song-Chun Zhu, and Ying Nian Wu. Learning generative convnets via multi-grid modeling and sampling. In *Proc. 2018 CVPR*, pages 9155–9164, 2018.

[GMT06] Sharad Goel, Ravi Montenegro, and Prasad Tetali. Mixing time bounds via the spectral profile. *Electron. J. Probab.*, 11, 2006.

[Gor85] Yehoram Gordon. Some inequalities for Gaussian processes and applications. *Israel J. Math.*, 50:265–289, 1985.

[Gor88] Yehoram Gordon. On Milman's inequality and random subspaces which escape through a mesh in $\mathbb{R}^n$. In *Geo. Func. Anal. Israel Sem.*, pages 84–106, 1988.

[GS17a] David Gamarnik and Madhu Sudan. Limits of local algorithms over sparse random graphs. *Ann. Probab.*, 45(4):2353–2376, 2017.

[GS17b] David Gamarnik and Madhu Sudan. Performance of sequential local algorithms for the random NAE-k-SAT problem. *SIAM J. Comp.*, 46(2):590–619, 2017.

[GS22] Reza Gheissari and Alistair Sinclair. Low-temperature Ising dynamics with random initializations. In *Proc. 54th STOC*, pages 1445–1458, 2022.

[GS24] Reza Gheissari and Alistair Sinclair. Spatial mixing and the random-cluster dynamics on lattices. *Rand. Struct. Alg.*, 64(2):490–534, 2024.

[GSVY16] Andreas Galanis, Daniel Stefankovic, Eric Vigoda, and Linji Yang. Ferromagnetic Potts model: refined #BIS-hardness and related results. *SIAM J. Comp.*, 45(6):2004–2065, 2016.

[GT02] Francesco Guerra and Fabio Lucio Toninelli. The thermodynamic limit in mean field spin glass models. *Comm. Math. Phys.*, 230(1):71–79, 2002.

[GTC24] Wei Guo, Molei Tao, and Yongxin Chen. Provable benefit of annealed Langevin Monte Carlo for non-log-concave sampling. *arXiv:2407.16936*, 2024.

[Gue03] Francesco Guerra. Broken replica symmetry bounds in the mean field spin glass model. *Comm. Math. Phys.*, 233(1):1–12, 2003.

[GW95] Michel X Goemans and David P Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *J. ACM*, 42(6):1115–1145, 1995.

[GZ00] A Guionnet and O Zeitouni. Concentration of the spectral measure for large matrices. *Elec. Comm. Probab.*, 5:119–136, 2000.

[HFS07] J William Helton, Reza Rashidi Far, and Roland Speicher. Operator-valued semicircular elements: solving a quadratic matrix equation with positivity constraints. *Intl. Math. Res. Not.*, 2007(9), 2007.

[Hin10] Geoffrey Hinton. A practical guide to training restricted Boltzmann machines. *Momentum*, 9(1):926, 2010.

[HJA20] Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *Proc. 2020 NeurIPS*, 33:6840–6851, 2020.

[HK23] Jun-Ting Hsieh and Pravesh K Kothari. Approximating max-cut on bounded degree graphs: tighter analysis of the FKL algorithm. In *Proc. 50th ICALP*, volume 261, page 77, 2023.

[HMP24] Brice Huang, Andrea Montanari, and Huy Tuan Pham. Sampling from spherical spin glasses in total variation via algorithmic stochastic localization. *arXiv:2404.15651*, 2024.

[HMRW25] Brice Huang, Sidhanth Mohanty, Amit Rajaraman, and David X Wu. Weak Poincaré inequalities, simulated annealing, and sampling from spherical spin glasses. In *Proc. 57th STOC (to appear)*, 2025.

[HPG18] Gavin S Hartnett, Edward Parker, and Edward Geist. Replica symmetry breaking in bipartite spin glasses and neural networks. *Phys. Rev. E*, 98(2):022116, 2018.

[HPR19] Tyler Helmuth, Will Perkins, and Guus Regts. Algorithmic Pirogov–Sinai theory. In *Proc. 51st STOC*, pages 1009–1020, 2019.

[HS23a] Brice Huang and Mark Sellke. Algorithmic threshold for multi-species spherical spin glasses. *arXiv:2303.12172*, 2023.

[HS23b] Brice Huang and Mark Sellke. A constructive proof of the spherical Parisi formula. *arXiv:2311.15495*, 2023.

[HS23c] Brice Huang and Mark Sellke. Strong topological trivialization of multi-species spherical spin glasses. *arXiv:2308.09677*, 2023.

[HS24] Brice Huang and Mark Sellke. Optimization algorithms for multi-species spherical spin glasses. *J. Stat. Phys.*, 191(2):29, 2024.

[HS25] Brice Huang and Mark Sellke. Tight Lipschitz hardness for optimizing mean field spin glasses. *Comm. Pure Appl. Math.*, 78(1):60–119, 2025.

[HSS25] Brice Huang, Mark Sellke, and Nike Sun. Optimal Lipschitz algorithms for high-dimensional projection pursuit. Manuscript in preparation, 2025.

[Hsu02] Elton P Hsu. *Stochastic Analysis on Manifolds*, volume 38. American Mathematical Society, 2002.

[Hua24] Brice Huang. Capacity threshold for the Ising perceptron. In *Proc. 65th FOCS*, pages 1126–1136, 2024.

[Jag17] Aukosh Jagannath. Approximate ultrametricity for random measures and applications to spin glasses. *Comm. Pure. Appl. Math.*, 70(4):611–664, 2017.

[JKP20] Matthew Jenssen, Peter Keevash, and Will Perkins. Algorithms for #BIS-hard problems on expander graphs. *SIAM J. Comp.*, 49(4):681–710, 2020.

[JM13] Adel Javanmard and Andrea Montanari. State evolution for general approximate message passing algorithms, with applications to spatial coupling. *Inform. Inference*, 2(2):115–144, 2013.

[JMSS23] Chris Jones, Kunal Marwaha, Juspreet Singh Sandhu, and Jonathan Shi. Random max-CSPs inherit algorithmic hardness from spin glasses. *Proc. 14th ITCS*, 2023.

[Joe92] Harry Joe. Generalized majorization orderings and applications. *Lecture Notes-Monograph Series*, pages 145–158, 1992.

[JR76] Robert E Jamison and William H Ruckle. Factoring absolutely convergent series. *Math. Annalen*, 224:143–148, 1976.

[JSV04] Mark Jerrum, Alistair Sinclair, and Eric Vigoda. A polynomial-time approximation algorithm for the permanent of a matrix with nonnegative entries. *J. ACM*, 51(4):671–697, 2004.

[JT16] Aukosh Jagannath and Ian Tobasco. A dynamic programming approach to the Parisi functional. *Proc. Amer. Math. Soc.*, 144(7):3135–3150, 2016.

[JT17] Aukosh Jagannath and Ian Tobasco. Low temperature asymptotics of spherical mean field spin glasses. *Comm. Math. Phys.*, 352(3):979–1017, 2017.

[JT18] Aukosh Jagannath and Ian Tobasco. Bounds on the complexity of replica symmetry breaking for spherical spin glasses. *Proc. Amer. Math Soc.*, 146(7):3127–3142, 2018.

[Kac48] M Kac. On the average number of real roots of a random algebraic equation (II). *Proc. London Math. Soc.*, 2(1):390–408, 1948.

[Kar09] Roman Karasev. KKM-type theorems for products of simplicegues and cutting sets and measures by straight lines. *arXiv:0909.0604*, 2009.

[KC75] John M Kincaid and Ezechiel Godert David Cohen. Phase diagrams of liquid helium mixtures and metamagnets: experiment and mean field theory. *Phys. Rep.*, 22(2):57–143, 1975.

[Kiv23] Pax Kivimae. The ground state energy and concentration of complexity in spherical bipartite models. *Comm. Math. Phys.*, 403(1):37–81, 2023.

[Kiv24] Pax Kivimae. Concentration of equilibria and relative instability in disordered non-relaxational dynamics. *Comm. Math. Phys.*, 405(12):289, 2024.

[KKMO07] Subhash Khot, Guy Kindler, Elchanan Mossel, and Ryan O'Donnell. Optimal inapproximability results for MAX-CUT and other 2-variable CSPs? *SIAM J. Comp.*, 37(1):319–357, 2007.

[KL24] Bo'az Klartag and Joseph Lehec. Isoperimetric inequalities in high-dimensional convex sets. *arXiv:2406.01324*, 2024.

[KLS97] Ravi Kannan, László Lovász, and Miklós Simonovits. Random walks and an $O^*(n^5)$ volume algorithm for convex bodies. *Rand. Struct. Alg.*, 11(1):1–50, 1997.

[KLV24] Frederic Koehler, Holden Lee, and June Vuong. Efficiently learning and sampling multimodal distributions with data-based initialization. *arXiv:2411.09117*, 2024.

[KM89] Werner Krauth and Marc Mézard. Storage capacity of memory networks with binary couplings. *J. Physique*, 50(20):3057–3066, 1989.

[KMRT$^+$07] Florent Krzakala, Andrea Montanari, Federico Ricci-Tersenghi, Guilhem Semerjian, and Lenka Zdeborová. Gibbs states and the set of solutions of random constraint satisfaction problems. *Proc. Natl. Acad. Sci.*, 104(25):10318–10323, 2007.

[KO20] Tali Kaufman and Izhar Oppenheim. High order random walks: beyond spectral gap. *Combinatorica*, 40:245–281, 2020.

[KR98] Jeong Han Kim and James R Roche. Covering cubes by random half cubes, with applications to binary neural networks. *J. Comput. Syst. Sci.*, 56(2):223–252, 1998.

[Kro79] Stephen Kronwith. Convex manifolds of nonnegative curvature. *J. Diff. Geom.*, 14(4):621–628, 1979.

[KS85] I Ya Korenblit and E.F. Shender. Spin glass in an Ising two-sublattice magnet. *Zh. Eksp. Teor. Fiz*, 89:1785–1795, 1985.

[KT87] Theodore R Kirkpatrick and Devarajan Thirumalai. p-spin-interaction spin-glass models: connections with the structural glass problem. *Phys. Rev. B*, 36(10):5388, 1987.

[KV24] Frederic Koehler and Thuy-Duong Vuong. Sampling multimodal distributions with the vanilla score: Benefits of data-based initialization. In *Proc. 12th ICLR*, 2024.

[LACTFLD24] Bertrand Lacroix-A-Chez-Toine, Yan V Fyodorov, and Pierre Le Doussal. Replica-symmetry breaking transitions in the large deviations of the ground-state of a spherical spin-glass. *J. Stat. Phys.*, 191(2):11, 2024.

[LE23] Mufan Li and Murat A Erdogdu. Riemannian Langevin algorithm for solving semidefinite programs. *Bernoulli*, 29(4):3093–3113, 2023.

[LG16] Jean-François Le Gall. *Brownian Motion, Martingales, and Stochastic Calculus*. Springer, 2016.

[Liu23] Kuikui Liu. *Spectral Independence: A New Tool to Analyze Markov Chains*. PhD thesis, University of Washington, 2023.

[LMR+24] Kuikui Liu, Sidhanth Mohanty, Prasad Raghavendra, Amit Rajaraman, and David X Wu. Locally stationary distributions: a framework for analyzing slow-mixing Markov chains. In *Proc. 65th FOCS*, pages 203–215, 2024.

[LMRW24] Kuikui Liu, Sidhanth Mohanty, Amit Rajaraman, and David X Wu. Fast mixing in sparse random Ising models. In *Proc. 65th FOCS*, pages 120–128, 2024.

[LRG18] Holden Lee, Andrej Risteski, and Rong Ge. Beyond log-concavity: provable guarantees for sampling multi-modal distributions using simulated tempering Langevin Monte Carlo. *Proc. 2018 NeurIPS*, 31, 2018.

[LS90] László Lovász and Miklós Simonovits. The mixing rate of Markov chains, an isoperimetric inequality, and computing the volume. In *Proc. 31st FOCS*, pages 346–354, 1990.

[LS93] László Lovász and Miklós Simonovits. Random walks in a convex body and an improved volume algorithm. *Rand. Struct. Alg.*, 4(4):359–412, 1993.

[LS16] Eyal Lubetzky and Allan Sly. Information percolation and cutoff for the stochastic Ising model. *J. Amer. Math. Soc.*, 29(3):729–774, 2016.

[LS17] Eyal Lubetzky and Allan Sly. Universality of cutoff for the Ising model. *Ann. Probab.*, 45(6):3664–3696, 2017.

[LSZ25] Shuangping Li, Tselil Schramm, and Kangjie Zhou. Discrepancy algorithms for the binary perceptron. 2025.

[LV24] Yin Tat Lee and Santosh S Vempala. Eldan's stochastic localization and the kls conjecture: isoperimetry, concentration and mixing. *Ann. Math.*, 199(3):1043–1092, 2024.

[Mat06] P Mathieu. Quand l'inegalite log-Sobolev implique l'inegalite de trou spectral. In *Sém. Probab. XXXII*, pages 30–35. 2006.

[MBB20] Antoine Maillard, Gérard Ben Arous, and Giulio Biroli. Landscape complexity for the empirical risk of generalized linear models. In *Math. Sci. Mach. Learn.*, pages 287–327, 2020.

[McK24] Benjamin McKenna. Complexity of bipartite spherical spin glasses. *Ann. Inst. Henri Poincaré Probab. Stat.*, 60(1):636–657, 2024.

[Méz89] Marc Mézard. The space of interactions in neural networks: Gardner's computation with the cavity method. *J. Phys. A*, 22(12):2181, 1989.

[Méz17] Marc Mézard. Mean-field message-passing equations in the Hopfield model and its generalizations. *Phys. Rev. E*, 95(2):022117, 2017.

[MGA23] M. Cristina Marchetti, Irene Giardina, and Ada Altieri. Complex dynamics in ecological systems and animal behavior. In *Spin Glass Theory and Far Beyond: Replica Symmetry Breaking After 40 Years*, pages 581–608. 2023.

[Mil63] John Milnor. *Morse Theory*, volume 51. Princeton University Press, 1963.

[MM09] Marc Mézard and Andrea Montanari. *Information, Physics, and Computation.* Oxford University Press, 2009.

[Mon21] Andrea Montanari. Optimization of the Sherrington–Kirkpatrick Hamiltonian. *SIAM J. Comp.*, pages FOCS19–1, 2021.

[Mon23a] Andrea Montanari. Optimization of random cost functions and statistical physics. In *Solvay Conference on Physics on "The Structure and Dynamics of Disordered Systems"*, 2023.

[Mon23b] Andrea Montanari. Sampling, diffusions, and stochastic localization. *arXiv:2305.10690*, 2023.

[Moo17] Cristopher Moore. The computer science and physics of community detection: Landscapes, phase transitions, and hardness. *Bull. EATCS*, 1(121), 2017.

[Mou21] Jean-Christophe Mourrat. Nonconvex interactions in mean-field spin glasses. *Probab. Math. Phys.*, 2(2):281–339, 2021.

[Mou23] Jean-Christophe Mourrat. Free energy upper bound for mean-field vector spin glasses. *Ann. Inst. Henri Poincaré Probab. Stat.*, 59(3):1143–1182, 2023.

[MPS+84a] Marc Mézard, Giorgio Parisi, Nicolas Sourlas, G Toulouse, and Miguel Virasoro. Nature of the spin-glass phase. *Phys. Rev. Lett.*, 52(13):1156, 1984.

[MPS+84b] Marc Mézard, Giorgio Parisi, Nicolas Sourlas, Gérard Toulouse, and Miguel Virasoro. Replica symmetry breaking and the nature of the spin glass phase. *J. Phys.*, 45(5):843–854, 1984.

[MPV87] Marc Mézard, Giorgio Parisi, and Miguel A Virasoro. *Spin Glass Theory and Beyond: An Introduction to the Replica Method and Its Applications*, volume 9. World Scientific, 1987.

[MR08] N.I. Muskhelishvili and J.R.M. Radok. *Singular Integral Equations: Boundary Problems of Function Theory and their Application to Mathematical Physics.* Courier Corporation, 2008.

[MRT04] Andrea Montanari and Federico Ricci-Tersenghi. Cooling-schedule dependence of the dynamics of mean-field glasses. *Phys. Rev. B*, 70(13):134406, 2004.

[MRTS07] Andrea Montanari, Federico Ricci-Tersenghi, and Guilhem Semerjian. Solving constraint satisfaction problems through belief propagation-guided decimation. *arXiv:0709.1667*, 2007.

[MS23] Andrea Montanari and Eliran Subag. Solving overparametrized systems of random equations: I. model and algorithms for approximate solutions. *arXiv:2306.13326*, 2023.

[MV85] Marc Mézard and Miguel Angel Virasoro. The microstructure of ultrametricity. *J. Phys.*, 46(8):1293–1307, 1985.

[MW23] Andrea Montanari and Yuchen Wu. Posterior sampling from the spiked models via diffusion processes. *arXiv:2304.11449*, 2023.

[MW24] Andrea Montanari and Yuchen Wu. Provably efficient posterior sampling for sparse linear regression via measure decomposition. *arXiv:2406.19550*, 2024.

[MZZ24] Andrea Montanari, Yiqiao Zhong, and Kangjie Zhou. Tractability from overparametrization: the example of the negative perceptron. *Probab. Theory Rel. Fields*, 188(3):805–910, 2024.

[Nes03] Yurii Nesterov. *Introductory Lectures on Convex Optimization: A Basic Course*, volume 87. Springer, 2003.

[NHH+20] Erik Nijkamp, Mitch Hill, Tian Han, Song-Chun Zhu, and Ying Nian Wu. On the anatomy of MCMC-based maximum likelihood learning of energy-based models. In *Proc. 2020 AAAI*, volume 34, pages 5272–5280, 2020.

[NS23] Shuta Nakajima and Nike Sun. Sharp threshold sequence and universality for Ising perceptron models. In *Proc. 34th SODA*, pages 638–674, 2023.

[NV14] Hoi H Nguyen and Van Vu. Random matrices: law of the determinant. *Ann. Probab.*, 42(1):146–167, 2014.

[Pan07] Dmitriy Panchenko. A note on Talagrand's positivity principle. *Elec. Comm. Probab.*, 12:401–410, 2007.

[Pan13a] Dmitry Panchenko. The Parisi ultrametricity conjecture. *Ann. Math.*, pages 383–393, 2013.

[Pan13b] Dmitry Panchenko. *The Sherrington–Kirkpatrick model.* Springer, 2013.

[Pan14] Dmitry Panchenko. The Parisi formula for mixed p-spin models. *Ann. Probab.*, 42(3):946–958, 2014.

[Pan15] Dmitry Panchenko. The free energy in a multi-species Sherrington–Kirkpatrick model. *Ann. Probab.*, 43(6):3494–3513, 2015.

[Pan18] Dmitry Panchenko. On the k-SAT model with large number of clauses. *Rand. Struct. Alg.*, 52(3):536–542, 2018.

[Par79] Giorgio Parisi. Infinite number of order parameters for spin-glasses. *Phys. Rev. Lett.*, 43(23):1754, 1979.

[Par80] Giorgio Parisi. A sequence of approximated solutions to the SK model for spin glasses. *J. Phys. A*, 13(4):L115, 1980.

[Par83] Giorgio Parisi. Order parameter for spin-glasses. *Phys. Rev. Lett.*, 50(24):1946, 1983.

[Par06] Giorgio Parisi. Computing the number of metastable states in infinite-range models. *arXiv:cond-mat/0602349*, 2006.

[PP95] Giorgio Parisi and Marc Potters. Mean-field equations for spin models with orthogonal interaction matrices. *J. Phys. A*, 28(18):5267, 1995.

[PS19] Yury Polyanskiy and Alex Samorodnitsky. Improved log-Sobolev inequalities, hypercontractivity and uncertainty principle on the hypercube. *J. Func. Anal.*, 277(11):108280, 2019.

[PX21] Will Perkins and Changji Xu. Frozen 1-RSB structure of the symmetric Ising perceptron. In *Proc. 53rd STOC*, pages 1579–1588, 2021.

[Rag08] Prasad Raghavendra. Optimal algorithms and inapproximability results for every CSP? In *Proc. 40th STOC*, pages 245–254, 2008.

[RBBC19] Valentina Ros, Gérard Ben Arous, Giulio Biroli, and Chiara Cammarota. Complex energy landscapes in spiked-tensor and simple glassy models: ruggedness, arrangements of local minima, and phase transitions. *Phys. Rev. X*, 9(1):011003, 2019.

[RF23] Valentina Ros and Yan V Fyodorov. The high-dimensional landscape paradigm: spin-glasses, and beyond. In *Spin Glass Theory and Far Beyond: Replica Symmetry Breaking After 40 Years*, pages 95–114. 2023.

[Ric44] Stephen O Rice. Mathematical analysis of random noise. *Bell Syst. Tech. J.*, 23(3):282–332, 1944.

[RMS+21] Alexander Rives, Joshua Meier, Tom Sercu, Siddharth Goyal, Zeming Lin, Jason Liu, Demi Guo, Myle Ott, C Lawrence Zitnick, Jerry Ma, et al. Biological structure and function emerge from scaling unsupervised learning to 250 million protein sequences. *Proc. Natl. Acad. Sci.*, 118(15):e2016239118, 2021.

[Rou13] Tomáš Roubíček. *Nonlinear Partial Differential Equations with Applications*, volume 153. Springer, 2013.

[RU08] Tom Richardson and Ruediger Urbanke. *Modern Coding Theory.* Cambridge University Press, 2008.

[Rud76] Walter Rudin. *Principles of Mathematical Analysis*, volume 3. McGraw-Hill, 1976.

[Rud87] Walter Rudin. *Real and Complex Analysis*, volume 156. McGraw-Hill, 1987.

[Rue87] David Ruelle. A mathematical reformulation of Derrida's REM and GREM. *Comm. Math. Phys.*, 108(2):225–239, 1987.

[RV17] Mustazee Rahman and Bálint Virág. Local algorithms for independent sets are half-optimal. *Ann. Probab.*, 45(3):1543–1577, 2017.

[RW01] Michael Röckner and Feng-Yu Wang. Weak Poincaré inequalities and l^2-convergence rates of Markov semigroups. *J. Func. Anal.*, 185(2):564–603, 2001.

[SCK+12] YiFan Sun, Andrea Crisanti, Florent Krzakala, Luca Leuzzi, and Lenka Zdeborová. Following states in temperature in the spherical $s+p$-spin glass model. *J. Stat. Mech. Theor. Exper.*, 2012(07):P07002, 2012.

[SDWMG15] Jascha Sohl-Dickstein, Eric Weiss, Niru Maheswaranathan, and Surya Ganguli. Deep unsupervised learning using nonequilibrium thermodynamics. In *Proc. 3rd ICLR*, pages 2256–2265, 2015.

[Sel24a] Mark Sellke. Optimizing mean field spin glasses with external field. *Electron. J. Probab.*, 29:1–47, 2024.

[Sel24b] Mark Sellke. The threshold energy of low temperature Langevin dynamics for pure spherical spin glasses. *Comm. Pure. Appl. Math.*, 77(11):4065–4099, 2024.

[SFL19] Ganlin Song, Zhou Fan, and John Lafferty. Surfing: iterative optimization over incrementally trained deep networks. *Proc. 2019 NeurIPS*, 32, 2019.

[SK75] David Sherrington and Scott Kirkpatrick. Solvable model of a spin-glass. *Phys. Rev. Lett.*, 35(26):1792, 1975.

[SS23] Ashwin Sah and Mehtaab Sawhney. Distribution of the threshold for the symmetric perceptron. In *Proc. 64th FOCS*, 2023.

[SSDK+21] Yang Song, Jascha Sohl-Dickstein, Diederik P Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations. In *Proc. 9th ICLR*, 2021.

[SSZ22] Allan Sly, Nike Sun, and Yumeng Zhang. The number of solutions for random regular NAE-SAT. *Probab. Theory Rel. Fields*, pages 1–109, 2022.

[ST03] Mariya Shcherbina and Brunello Tirozzi. Rigorous solution of the Gardner problem. *Comm. Math. Phys.*, 234:383–422, 2003.

[Sto13a] Mihailo Stojnic. Another look at the Gardner problem. *arXiv:1306.3979*, 2013.

[Sto13b] Mihailo Stojnic. Negative spherical perceptron. *arXiv:1306.3980*, 2013.

[Sub17a] Eliran Subag. The complexity of spherical p-spin models — a second moment approach. *Ann. Probab.*, 45(5):3385–3450, 2017.

[Sub17b] Eliran Subag. The geometry of the Gibbs measure of pure spherical spin glasses. *Invent. Math.*, 210(1):135–209, 2017.

[Sub21a] Eliran Subag. Following the ground states of full-RSB spherical spin glasses. *Comm. Pure Appl. Math.*, 74(5):1021–1044, 2021.

[Sub21b] Eliran Subag. TAP approach for multi-species spherical spin glasses I: general theory. *arXiv:2111.07132*, 2021.

[Sub23a] Eliran Subag. Concentration for the zero set of random polynomial systems. *arXiv:2303.11924*, 2023.

[Sub23b] Eliran Subag. TAP approach for multispecies spherical spin glasses II: the free energy of the pure models. *Ann. Probab.*, 51(3):1004–1024, 2023.

[Sub24] Eliran Subag. Free energy landscapes in spherical spin glasses. *Duke Math. J.*, 173(7):1291–1357, 2024.

[Sud71] Vladimir N Sudakov. Gaussian random processes and measures of solid angles in Hilbert space. *Dokl. Akad. Nauk SSSR*, 197(1):43–45, 1971.

[Sud79] Vladimir N Sudakov. *Geometric Problems in the Theory of Infinite-Dimensional Probability Distributions*, volume 141. American Mathematical Society, 1979.

[SW71] Elias M Stein and Guido Weiss. *Introduction to Fourier Analysis on Euclidean Spaces*, volume 1. Princeton University Press, 1971.

[SZ81] Haim Sompolinsky and Annette Zippelius. Dynamic theory of the spin-glass phase. *Phys. Rev. Lett.*, 47(5):359, 1981.

[SZ17] Eliran Subag and Ofer Zeitouni. The extremal process of critical points of the pure p-spin spherical spin glass model. *Probab. Theory Rel. Fields*, 168(3-4):773–820, 2017.

[SZ21] Eliran Subag and Ofer Zeitouni. Concentration of the complexity of spherical pure p-spin models at arbitrary energies. *J. Math. Phys.*, 62(12):123301, 2021.

[TA03] Jonathan E Taylor and Robert J Adler. Euler characteristics for gaussian fields on manifolds. *Ann. Probab.*, 31(2):533–563, 2003.

[Tal99] Michel Talagrand. Intersecting random half cubes. *Rand. Struct. Alg.*, 15(3-4):436–449, 1999.

[Tal00] Michel Talagrand. Intersecting random half-spaces: toward the Gardner-Derrida formula. *Ann. Probab.*, 28(2):725–758, 2000.

[Tal06a] Michel Talagrand. Free energy of the spherical mean field model. *Probab. Theory Rel. Fields*, 134:339–382, 2006.

[Tal06b] Michel Talagrand. The Parisi formula. *Ann. Math.*, pages 221–263, 2006.

[Tal07] Michel Talagrand. Large deviations, Guerra's and ASS schemes, and the Parisi hypothesis. *J. Stat. Phys.*, 126, 2007.

[Tal10] Michel Talagrand. *Mean Field Models for Spin Glasses: Volume I: Basic Examples*, volume 54. Springer, 2010.

[Tal11] Michel Talagrand. *Mean Field Models for Spin Glasses. Volume II: Advanced Replica-Symmetry and Low Temperature*, volume 55. Springer, 2011.

[TAP77] David J Thouless, Philip W Anderson, and Robert G Palmer. Solution of 'solvable model of a spin glass'. *Philos. Mag.*, 35(3):593–601, 1977.

[Tau49] Olga Taussky. A recurring theorem on determinants. *Amer. Math. Monthly*, 56(10p1):672–676, 1949.

[Tre01] Luca Trevisan. Non-approximability results for optimization problems on bounded degree instances. In *Proc. 33rd STOC*, pages 453–461, 2001.

[TV12] Terence Tao and Van Vu. A central limit theorem for the determinant of a Wigner matrix. *Adv. Math.*, 231(1):74–101, 2012.

[Voi91] Dan Voiculescu. Limit laws for random matrices and free products. *Invent. Math.*, 104(1):201–220, 1991.

[Wan06] Feng-Yu Wang. *Functional Inequalities, Markov Semigroups and Spectral Theory.* Elsevier, 2006.

[Wan14] Feng-Yu Wang. *Analysis for Diffusion Processes on Riemannian Manifolds*, volume 18. World Scientific, 2014.

[Wei22] Alexander S Wein. Optimal low-degree hardness of maximum independent set. *Math. Stat. Learn.*, 4(3):221–251, 2022.

[Wen62] James G Wendel. A problem in geometric probability. *Math. Scandin.*, 11(1):109–111, 1962.

[XLZW16] Jianwen Xie, Yang Lu, Song-Chun Zhu, and Yingnian Wu. A theory of generative convnet. In *Proc. 2016 ICML*, pages 2635–2644, 2016.

[Xu21] Changji Xu. Sharp threshold for the Ising perceptron model. *Ann. Probab.*, 49(5):2399–2415, 2021.

[Yeo18] Dominic Yeo. Frozen percolation on inhomogeneous random graphs. *arXiv:1810.02750*, 2018.

[Zaa86] AC Zaanen. Continuity of measurable functions. *Amer. Math. Monthly*, 93(2):128–130, 1986.

[Zie12] William P Ziemer. *Weakly Differentiable Functions: Sobolev Spaces and Functions of Bounded Variation*, volume 120. Springer, 2012.

[ZK07] Lenka Zdeborová and Florent Krzakala. Phase transitions in the coloring of random graphs. *Phys. Rev. E*, 76(3):031131, 2007.

[ZK10] Lenka Zdeborová and Florent Krzakala. Generalization of the cavity method for adiabatic evolution of Gibbs states. *Phys. Rev. B*, 81(22):224205, 2010.

[ZK16] Lenka Zdeborová and Florent Krzakala. Statistical physics of inference: thresholds and algorithms. *Adv. Phys.*, 65(5):453–552, 2016.