Strong Topological Trivialization for Multi-Species Spin Glasses

Brice Huang (MIT)

Harvard Probabilitas Seminar

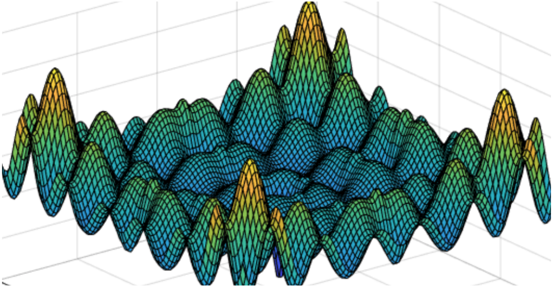
Joint work with Mark Sellke (Harvard)

Plan for this talk

1. Introduction and background
    - Landscape complexity of random functions
    - Topological trivialization
2. Strong topological trivialization for spin glasses
3. Multi-species spin glasses

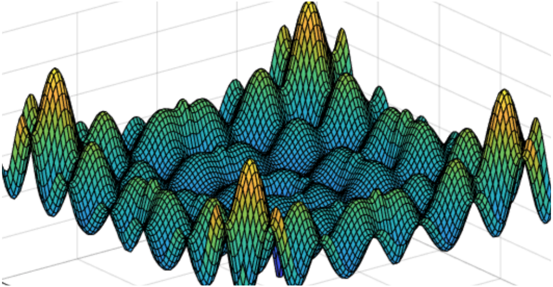
Random Landscapes

Landscapes of random, high-dimensional functions



Random Landscapes

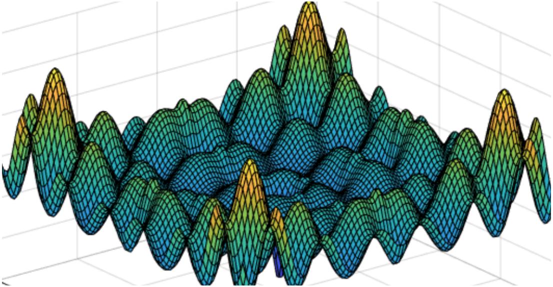
Landscapes of random, high-dimensional functions



- Connections to optimization, questions about algorithmic tractability

Random Landscapes

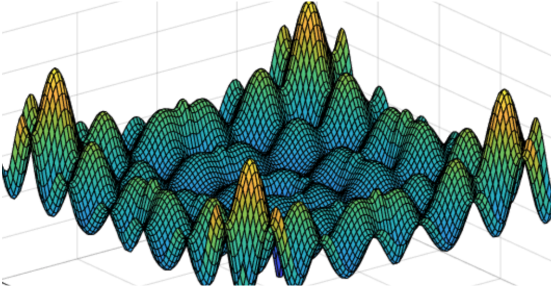
Landscapes of random, high-dimensional functions



- Connections to optimization, questions about algorithmic tractability
- Example: loss function over random data in learning applications

Random Landscapes

Landscapes of random, high-dimensional functions



- Connections to optimization, questions about algorithmic tractability
- Example: loss function over random data in learning applications

How complicated is the landscape of a random function?

Mean Field Spin Glasses

Polynomials $H_N : \mathbb{R}^N \to \mathbb{R}$ with **random** coefficients, e.g. random cubic

$$H_N(\boldsymbol{\sigma}) = \frac{1}{N} \sum_{i_1, i_2, i_3 = 1}^{N} g_{i_1, i_2, i_3} \cdot \sigma_{i_1} \sigma_{i_2} \sigma_{i_3} \qquad g_{i_1, i_2, i_3} \underset{\text{i.i.d.}}{\sim} \mathcal{N}(0, 1)$$

Mean Field Spin Glasses

Polynomials $H_N : \mathbb{R}^N \to \mathbb{R}$ with **random** coefficients, e.g. random cubic

$$H_N(\boldsymbol{\sigma}) = \frac{1}{N} \sum_{i_1, i_2, i_3 = 1}^{N} g_{i_1, i_2, i_3} \cdot \sigma_{i_1} \sigma_{i_2} \sigma_{i_3} = \frac{1}{N} \langle \boldsymbol{G}^{(3)}, \boldsymbol{\sigma}^{\otimes 3} \rangle \qquad g_{i_1, i_2, i_3} \underset{\text{i.i.d.}}{\sim} \mathcal{N}(0, 1)$$

Mean Field Spin Glasses

Polynomials $H_N : \mathbb{R}^N \to \mathbb{R}$ with **random** coefficients, e.g. random cubic

$$H_N(\boldsymbol{\sigma}) = \frac{1}{N} \sum_{i_1,i_2,i_3=1}^{N} g_{i_1,i_2,i_3} \cdot \sigma_{i_1}\sigma_{i_2}\sigma_{i_3} = \frac{1}{N}\langle \boldsymbol{G}^{(3)}, \boldsymbol{\sigma}^{\otimes 3}\rangle \qquad g_{i_1,i_2,i_3} \underset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$$

More generally, mix different degrees. For $\gamma_1, \gamma_2, \ldots \geq 0$,

$$H_N(\boldsymbol{\sigma}) = \sum_{p=1}^{P} \frac{\gamma_p}{N^{(p-1)/2}} \langle \boldsymbol{G}^{(p)}, \boldsymbol{\sigma}^{\otimes p}\rangle \qquad \boldsymbol{G}^{(p)} \in (\mathbb{R}^N)^{\otimes p} \text{ i.i.d. } \mathcal{N}(0,1)\text{s}$$

Mean Field Spin Glasses

Polynomials $H_N : \mathbb{R}^N \to \mathbb{R}$ with **random** coefficients, e.g. random cubic

$$H_N(\boldsymbol{\sigma}) = \frac{1}{N} \sum_{i_1,i_2,i_3=1}^{N} g_{i_1,i_2,i_3} \cdot \sigma_{i_1}\sigma_{i_2}\sigma_{i_3} = \frac{1}{N}\langle \boldsymbol{G}^{(3)}, \boldsymbol{\sigma}^{\otimes 3}\rangle \qquad g_{i_1,i_2,i_3} \underset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$$

More generally, mix different degrees. For $\gamma_1, \gamma_2, \ldots \geq 0$,

$$H_N(\boldsymbol{\sigma}) = \sum_{p=1}^{P} \frac{\gamma_p}{N^{(p-1)/2}} \langle \boldsymbol{G}^{(p)}, \boldsymbol{\sigma}^{\otimes p}\rangle \qquad \boldsymbol{G}^{(p)} \in (\mathbb{R}^N)^{\otimes p} \text{ i.i.d. } \mathcal{N}(0,1)\text{s}$$

Inputs on sphere $S_N = \{\boldsymbol{\sigma} \in \mathbb{R}^N : \|\boldsymbol{\sigma}\|_2 = \sqrt{N}\}$

Mean Field Spin Glasses

Polynomials $H_N : \mathbb{R}^N \to \mathbb{R}$ with **random** coefficients, e.g. random cubic

$$H_N(\boldsymbol{\sigma}) = \frac{1}{N} \sum_{i_1,i_2,i_3=1}^{N} g_{i_1,i_2,i_3} \cdot \sigma_{i_1}\sigma_{i_2}\sigma_{i_3} = \frac{1}{N}\langle \boldsymbol{G}^{(3)}, \boldsymbol{\sigma}^{\otimes 3}\rangle \qquad g_{i_1,i_2,i_3} \underset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$$

More generally, mix different degrees. For $\gamma_1, \gamma_2, \ldots \geq 0$,

$$H_N(\boldsymbol{\sigma}) = \sum_{p=1}^{P} \frac{\gamma_p}{N^{(p-1)/2}} \langle \boldsymbol{G}^{(p)}, \boldsymbol{\sigma}^{\otimes p}\rangle \qquad \boldsymbol{G}^{(p)} \in (\mathbb{R}^N)^{\otimes p} \text{ i.i.d. } \mathcal{N}(0,1)\text{s}$$

Inputs on sphere $S_N = \{\boldsymbol{\sigma} \in \mathbb{R}^N : \|\boldsymbol{\sigma}\|_2 = \sqrt{N}\}$

Rotationally invariant Gaussian process with covariance

$$\mathbb{E}[H_N(\boldsymbol{\sigma})H_N(\boldsymbol{\rho})] = N\xi(\langle \boldsymbol{\sigma}, \boldsymbol{\rho}\rangle / N) \qquad \xi(q) = \sum_{p=1}^{P} \gamma_p^2 q^p$$

Mean Field Spin Glasses

Polynomials $H_N : \mathbb{R}^N \to \mathbb{R}$ with **random** coefficients, e.g. random cubic

$$H_N(\boldsymbol{\sigma}) = \frac{1}{N} \sum_{i_1,i_2,i_3=1}^{N} g_{i_1,i_2,i_3} \cdot \sigma_{i_1}\sigma_{i_2}\sigma_{i_3} = \frac{1}{N}\langle \boldsymbol{G}^{(3)}, \boldsymbol{\sigma}^{\otimes 3}\rangle \qquad g_{i_1,i_2,i_3} \underset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$$

More generally, mix different degrees. For $\gamma_1, \gamma_2, \ldots \geq 0$,

$$H_N(\boldsymbol{\sigma}) = \sum_{p=1}^{P} \frac{\gamma_p}{N^{(p-1)/2}} \langle \boldsymbol{G}^{(p)}, \boldsymbol{\sigma}^{\otimes p}\rangle \qquad \boldsymbol{G}^{(p)} \in (\mathbb{R}^N)^{\otimes p} \text{ i.i.d. } \mathcal{N}(0,1)\text{s}$$

Inputs on sphere $S_N = \{\boldsymbol{\sigma} \in \mathbb{R}^N : \|\boldsymbol{\sigma}\|_2 = \sqrt{N}\}$

Rotationally invariant Gaussian process with covariance

$$\mathbb{E}[H_N(\boldsymbol{\sigma})H_N(\boldsymbol{\rho})] = N\xi(\langle \boldsymbol{\sigma}, \boldsymbol{\rho}\rangle / N) \qquad \xi(q) = \sum_{p=1}^{P} \gamma_p^2 q^p$$

ξ **mixture function**. Cubic above: $\xi(q) = q^3$.

Mean Field Spin Glasses

Polynomials $H_N : \mathbb{R}^N \to \mathbb{R}$ with **random** coefficients, e.g. random cubic

$$H_N(\boldsymbol{\sigma}) = \frac{1}{N} \sum_{i_1,i_2,i_3=1}^{N} g_{i_1,i_2,i_3} \cdot \sigma_{i_1}\sigma_{i_2}\sigma_{i_3} = \frac{1}{N}\langle \boldsymbol{G}^{(3)}, \boldsymbol{\sigma}^{\otimes 3}\rangle \qquad g_{i_1,i_2,i_3} \underset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$$

More generally, mix different degrees. For $\gamma_1, \gamma_2, \ldots \geq 0$,

$$H_N(\boldsymbol{\sigma}) = \sum_{p=1}^{P} \frac{\gamma_p}{N^{(p-1)/2}} \langle \boldsymbol{G}^{(p)}, \boldsymbol{\sigma}^{\otimes p}\rangle \qquad \boldsymbol{G}^{(p)} \in (\mathbb{R}^N)^{\otimes p} \text{ i.i.d. } \mathcal{N}(0,1)\text{s}$$

Inputs on sphere $S_N = \{\boldsymbol{\sigma} \in \mathbb{R}^N : \|\boldsymbol{\sigma}\|_2 = \sqrt{N}\}$

Rotationally invariant Gaussian process with covariance

$$\mathbb{E}[H_N(\boldsymbol{\sigma})H_N(\boldsymbol{\rho})] = N\xi(\langle \boldsymbol{\sigma}, \boldsymbol{\rho}\rangle / N) \qquad \xi(q) = \sum_{p=1}^{P} \gamma_p^2 q^p$$

ξ **mixture function**. Cubic above: $\xi(q) = q^3$. $p = 1$ term is **external field**.

Kac-Rice Formula

Computes expected number of zeros of random function $f : \Omega \times [0, L] \to \mathbb{R}$

$$\mathbb{E}[\#\text{zeros}(f)] = \int_{[0,L]} \mathbb{E}\left[|f'(x)| \middle| f(x) = 0\right] \varphi_{f(x)}(0)\, \mathrm{d}x.$$

Here $\varphi_{f(x)}$ is density of r.v. $f(x)$.

Kac-Rice Formula

Computes expected number of zeros of random function $f : \Omega \times [0, L] \to \mathbb{R}$

$$\mathbb{E}[\#\text{zeros}(f)] = \int_{[0,L]} \mathbb{E}\left[|f'(x)| \middle| f(x) = 0\right] \varphi_{f(x)}(0)\ \mathrm{d}x.$$

Here $\varphi_{f(x)}$ is density of r.v. $f(x)$. Heuristic derivation:

$$\mathbb{E}[\#\text{zeros}(f)] = \mathbb{E} \int_{f([0,L])} \delta(y)\ \mathrm{d}y$$

Kac-Rice Formula

Computes expected number of zeros of random function $f : \Omega \times [0, L] \to \mathbb{R}$

$$\mathbb{E}[\#\text{zeros}(f)] = \int_{[0,L]} \mathbb{E}\left[|f'(x)| \middle| f(x) = 0\right] \varphi_{f(x)}(0)\ \mathrm{d}x.$$

Here $\varphi_{f(x)}$ is density of r.v. $f(x)$. Heuristic derivation:

$$\mathbb{E}[\#\text{zeros}(f)] = \mathbb{E}\int_{f([0,L])} \delta(y)\ \mathrm{d}y = \mathbb{E}\int_{[0,L]} |f'(x)|\delta(f(x))\ \mathrm{d}x$$

Kac-Rice Formula

Computes expected number of zeros of random function $f : \Omega \times [0, L] \to \mathbb{R}$

$$\mathbb{E}[\#\text{zeros}(f)] = \int_{[0,L]} \mathbb{E}\left[|f'(x)|\middle| f(x) = 0\right] \varphi_{f(x)}(0)\ \mathrm{d}x.$$

Here $\varphi_{f(x)}$ is density of r.v. $f(x)$. Heuristic derivation:

$$\begin{aligned}\mathbb{E}[\#\text{zeros}(f)] &= \mathbb{E}\int_{f([0,L])} \delta(y)\ \mathrm{d}y = \mathbb{E}\int_{[0,L]} |f'(x)|\delta(f(x))\ \mathrm{d}x \\ &= \int_{[0,L]} \mathbb{E}\left[|f'(x)||f(x) = 0\right] \underbrace{\mathbb{E}\delta(f(x))}_{=\varphi_{f(x)}(0)}\ \mathrm{d}x.\end{aligned}$$

Kac-Rice Formula

Computes expected number of zeros of random function $f : \Omega \times [0, L] \to \mathbb{R}$

$$\mathbb{E}[\#\text{zeros}(f)] = \int_{[0,L]} \mathbb{E}\left[|f'(x)|\middle|f(x) = 0\right] \varphi_{f(x)}(0) \,\mathrm{d}x.$$

Here $\varphi_{f(x)}$ is density of r.v. $f(x)$. Heuristic derivation:

$$\begin{aligned}\mathbb{E}[\#\text{zeros}(f)] &= \mathbb{E}\int_{f([0,L])} \delta(y)\,\mathrm{d}y = \mathbb{E}\int_{[0,L]} |f'(x)|\delta(f(x))\,\mathrm{d}x \\ &= \int_{[0,L]} \mathbb{E}\left[|f'(x)||f(x) = 0\right] \underbrace{\mathbb{E}\delta(f(x))}_{=\varphi_{f(x)}(0)} \,\mathrm{d}x.\end{aligned}$$

For multi-dimensional $f : \Omega \times [0, L]^N \to \mathbb{R}^N$:

$$\mathbb{E}[\#\text{zeros}(f)] = \int_{[0,L]^N} \mathbb{E}\left[|\det \nabla f(\boldsymbol{x})|\middle|f(\boldsymbol{x}) = 0\right] \varphi_{f(\boldsymbol{x})}(0) \,\mathrm{d}\boldsymbol{x}.$$

A simple example

How many zeros does this function have on $[-L, L]$, $L < 1$, in expectation?

$$f(x) = g_0 + g_1 x + g_2 x^2 + \cdots, \qquad g_i \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, 1)$$

A simple example

How many zeros does this function have on $[-L, L]$, $L < 1$, in expectation?

$$f(x) = g_0 + g_1 x + g_2 x^2 + \cdots, \qquad g_i \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$$

At any $|x| < 1$, r.v.s $f(x), f'(x)$ jointly Gaussian with covariance

$$\mathbb{E}f(x)^2 = \frac{1}{1-x^2} \qquad \mathbb{E}f'(x)^2 = \frac{1+x^2}{(1-x^2)^3} \qquad \mathbb{E}f(x)f'(x) = \frac{x}{(1-x)^2}$$

A simple example

How many zeros does this function have on $[-L, L]$, $L < 1$, in expectation?

$$f(x) = g_0 + g_1 x + g_2 x^2 + \cdots, \qquad g_i \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$$

At any $|x| < 1$, r.v.s $f(x), f'(x)$ jointly Gaussian with covariance

$$\mathbb{E} f(x)^2 = \frac{1}{1-x^2} \qquad \mathbb{E} f'(x)^2 = \frac{1+x^2}{(1-x^2)^3} \qquad \mathbb{E} f(x) f'(x) = \frac{x}{(1-x)^2}$$

So

$$\mathrm{Var}[f'(x)|f(x)] = \mathbb{E} f'(x)^2 - \frac{(\mathbb{E} f(x) f'(x))^2}{\mathbb{E} f(x)^2} = \frac{1}{(1-x^2)^3}$$

A simple example

How many zeros does this function have on $[-L, L]$, $L < 1$, in expectation?

$$f(x) = g_0 + g_1 x + g_2 x^2 + \cdots, \qquad g_i \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$$

At any $|x| < 1$, r.v.s $f(x), f'(x)$ jointly Gaussian with covariance

$$\mathbb{E} f(x)^2 = \frac{1}{1-x^2} \qquad \mathbb{E} f'(x)^2 = \frac{1+x^2}{(1-x^2)^3} \qquad \mathbb{E} f(x) f'(x) = \frac{x}{(1-x)^2}$$

So

$$\mathrm{Var}[f'(x)|f(x)] = \mathbb{E} f'(x)^2 - \frac{(\mathbb{E} f(x) f'(x))^2}{\mathbb{E} f(x)^2} = \frac{1}{(1-x^2)^3}$$

Then

$$\mathbb{E}[\#\text{zeros}(f)] = \int_{-L}^{L} \mathbb{E}\left[|f'(x)| \middle| f(x) = 0\right] \varphi_{f(x)}(0)\, \mathrm{d}x$$

A simple example

How many zeros does this function have on $[-L, L]$, $L < 1$, in expectation?

$$f(x) = g_0 + g_1 x + g_2 x^2 + \cdots, \qquad g_i \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$$

At any $|x| < 1$, r.v.s $f(x), f'(x)$ jointly Gaussian with covariance

$$\mathbb{E}f(x)^2 = \frac{1}{1-x^2} \qquad \mathbb{E}f'(x)^2 = \frac{1+x^2}{(1-x^2)^3} \qquad \mathbb{E}f(x)f'(x) = \frac{x}{(1-x)^2}$$

So

$$\mathrm{Var}[f'(x)|f(x)] = \mathbb{E}f'(x)^2 - \frac{(\mathbb{E}f(x)f'(x))^2}{\mathbb{E}f(x)^2} = \frac{1}{(1-x^2)^3}$$

Then

$$\begin{aligned}\mathbb{E}[\#\text{zeros}(f)] &= \int_{-L}^{L} \mathbb{E}\left[|f'(x)| \middle| f(x) = 0\right] \varphi_{f(x)}(0)\, \mathrm{d}x \\ &= \int_{-L}^{L} \sqrt{\frac{2}{\pi}} \frac{1}{(1-x^2)^{3/2}} \cdot \sqrt{\frac{1-x^2}{2\pi}}\, \mathrm{d}x\end{aligned}$$

A simple example

How many zeros does this function have on $[-L, L]$, $L < 1$, in expectation?

$$f(x) = g_0 + g_1 x + g_2 x^2 + \cdots, \qquad g_i \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$$

At any $|x| < 1$, r.v.s $f(x), f'(x)$ jointly Gaussian with covariance

$$\mathbb{E}f(x)^2 = \frac{1}{1-x^2} \qquad \mathbb{E}f'(x)^2 = \frac{1+x^2}{(1-x^2)^3} \qquad \mathbb{E}f(x)f'(x) = \frac{x}{(1-x)^2}$$

So

$$\mathrm{Var}[f'(x)|f(x)] = \mathbb{E}f'(x)^2 - \frac{(\mathbb{E}f(x)f'(x))^2}{\mathbb{E}f(x)^2} = \frac{1}{(1-x^2)^3}$$

Then

$$\begin{aligned}
\mathbb{E}[\#\text{zeros}(f)] &= \int_{-L}^{L} \mathbb{E}\left[|f'(x)| \middle| f(x) = 0\right] \varphi_{f(x)}(0)\, \mathrm{d}x \\
&= \int_{-L}^{L} \sqrt{\frac{2}{\pi}} \frac{1}{(1-x^2)^{3/2}} \cdot \sqrt{\frac{1-x^2}{2\pi}}\, \mathrm{d}x \\
&= \frac{1}{\pi} \int_{-L}^{L} \frac{1}{1-x^2}\, \mathrm{d}x = \frac{2}{\pi} \operatorname{arctanh} L
\end{aligned}$$

Kac-Rice to High-Dimensional Landscape Complexity

- Quantify complexity of landscape by # critical points

Kac-Rice to High-Dimensional Landscape Complexity

- Quantify complexity of landscape by # critical points
- (Fyodorov 04, Auffinger-Ben Arous-Černý 13): Kac-Rice on ∇H gives $\mathbb{E}|\mathrm{Crt}|$

Kac-Rice to High-Dimensional Landscape Complexity

- Quantify complexity of landscape by # critical points

- (Fyodorov 04, Auffinger-Ben Arous-Černý 13): Kac-Rice on ∇H gives $\mathbb{E}|\mathrm{Crt}|$
- In high dimensions, reduces to random determinant:

$$\mathbb{E}|\mathrm{Crt}| = (\text{simple term}) \cdot \mathbb{E}|\det(\text{random matrix})|.$$

Kac-Rice to High-Dimensional Landscape Complexity

- Quantify complexity of landscape by # critical points
- (Fyodorov 04, Auffinger-Ben Arous-Černý 13): Kac-Rice on ∇H gives $\mathbb{E}|\mathrm{Crt}|$
- In high dimensions, reduces to random determinant:

$$\mathbb{E}|\mathrm{Crt}| = (\text{simple term}) \cdot \mathbb{E}|\det(\text{random matrix})|.$$

- For spin glasses, random matrix is $\mathrm{GOE}_N + cI_N$, $\mathbb{E}|\det|$ known **exactly**

Kac-Rice to High-Dimensional Landscape Complexity

- Quantify complexity of landscape by # critical points
- (Fyodorov 04, Auffinger-Ben Arous-Černý 13): Kac-Rice on ∇H gives $\mathbb{E}|\mathrm{Crt}|$
- In high dimensions, reduces to random determinant:

$$\mathbb{E}|\mathrm{Crt}| = (\text{simple term}) \cdot \mathbb{E}|\det(\text{random matrix})|.$$

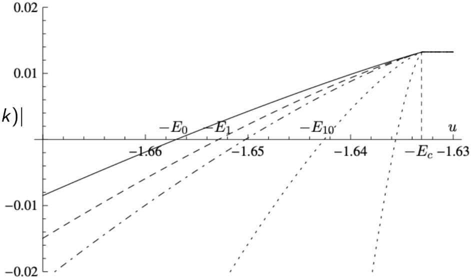
- For spin glasses, random matrix is $\mathrm{GOE}_N + cI_N$, $\mathbb{E}|\det|$ known **exactly**
- Other applications:
  - (Sagun-Guney-Ben Arous-LeCun 14) neural networks
  - (Ben Arous-Mei-Montanari-Nica 19) spiked tensor models
  - (Ben Arous-Fyodorov-Khoruzhenko 21, Subag 23) non-gradient vector fields
  - (Ben Arous-Bourgarde-McKenna 23) elastic manifold

Critical Point Complexity in Pure Spin Glasses

- (Auffinger-Ben Arous-Černý 13): $\mathbb{E}\#$ critical points at given (energy, index)

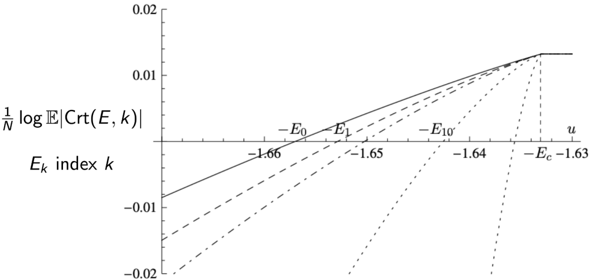
$\frac{1}{N} \log \mathbb{E}|\mathrm{Crt}(E,k)|$

E_k index k



Critical Point Complexity in Pure Spin Glasses

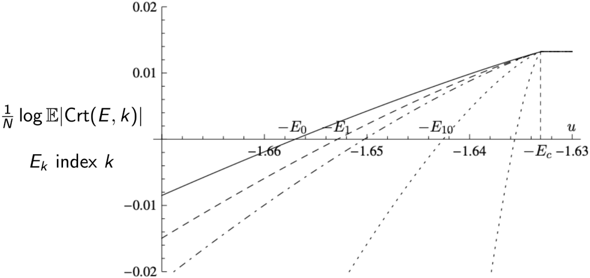
- (Auffinger-Ben Arous-Černý 13): $\mathbb{E}\#$ critical points at given (energy, index)



- (Subag 17): this matches **typical** counts by 2nd moment method

Critical Point Complexity in Pure Spin Glasses

- (Auffinger-Ben Arous-Černý 13): $\mathbb{E}\#$ critical points at given (energy, index)



- (Subag 17): this matches **typical** counts by 2nd moment method
- Consequence: ground state energy matching Parisi formula

Topological Trivialization

- If external field large, only 2 critical points are global max and min ...

Topological Trivialization

- If external field large, only 2 critical points are global max and min . . .
- At how large external field does this occur? Is there a phase transition?

Topological Trivialization

- If external field large, only 2 critical points are global max and min ...
- At how large external field does this occur? Is there a phase transition?

Theorem (Fyodorov 15, Belius-Černý-Nakajima-Schmidt 22)

1. *If* $\xi'(1) > \xi''(1)$, $\mathbb{E}|\mathrm{Crt}| = 2 + o(1)$ *(so w.h.p. 2 critical points)*

Topological Trivialization

- If external field large, only 2 critical points are global max and min ...
- At how large external field does this occur? Is there a phase transition?

Theorem (Fyodorov 15, Belius-Černý-Nakajima-Schmidt 22)

1. *If* $\xi'(1) > \xi''(1)$, $\mathbb{E}|\mathrm{Crt}| = 2 + o(1)$ *(so w.h.p. 2 critical points)*
2. *If* $\xi'(1) < \xi''(1)$, $\mathbb{E}|\mathrm{Crt}| \geq e^{cN}$

Topological Trivialization

- If external field large, only 2 critical points are global max and min ...
- At how large external field does this occur? Is there a phase transition?

Theorem (Fyodorov 15, Belius-Černý-Nakajima-Schmidt 22)

1. *If* $\xi'(1) > \xi''(1)$, $\mathbb{E}|\mathrm{Crt}| = 2 + o(1)$ *(so w.h.p. 2 critical points)*
2. *If* $\xi'(1) < \xi''(1)$, $\mathbb{E}|\mathrm{Crt}| \geq e^{cN}$

- Phase boundary for **annealed topological trivialization**

Topological Trivialization

- If external field large, only 2 critical points are global max and min ...
- At how large external field does this occur? Is there a phase transition?

Theorem (Fyodorov 15, Belius-Černý-Nakajima-Schmidt 22)

1. *If* $\xi'(1) > \xi''(1)$, $\mathbb{E}|\mathrm{Crt}| = 2 + o(1)$ *(so w.h.p. 2 critical points)*
2. *If* $\xi'(1) < \xi''(1)$, $\mathbb{E}|\mathrm{Crt}| \geq e^{cN}$

- Phase boundary for **annealed topological trivialization**
- $\xi'(1) > \xi''(1)$ equivalent to $\gamma_1^2 > \sum_{p \geq 3} p(p-2)\gamma_p^2$

Topological Trivialization

- If external field large, only 2 critical points are global max and min ...
- At how large external field does this occur? Is there a phase transition?

Theorem (Fyodorov 15, Belius-Černý-Nakajima-Schmidt 22)

1. *If* $\xi'(1) > \xi''(1)$, $\mathbb{E}|\mathrm{Crt}| = 2 + o(1)$ *(so w.h.p. 2 critical points)*
2. *If* $\xi'(1) < \xi''(1)$, $\mathbb{E}|\mathrm{Crt}| \geq e^{cN}$

- Phase boundary for **annealed topological trivialization**
- $\xi'(1) > \xi''(1)$ equivalent to $\gamma_1^2 > \sum_{p\geq 3} p(p-2)\gamma_p^2$
- $\mathbb{E}|\mathrm{Crt}| = 2 + o(1)$ achieved using exact formula for GOE
    - If Kac-Rice random matrix not GOE, current tools only achieve $\mathbb{E}|\mathrm{Crt}| = e^{o(N)}$

Guiding Questions

- Do phase boundaries of annealed and quenched trivialization coincide?
  - Or can $|\mathrm{Crt}| = 2$ w.h.p. in the regime where $\mathbb{E}|\mathrm{Crt}| \geq e^{cN}$?

Guiding Questions

- Do phase boundaries of annealed and quenched trivialization coincide?
  - Or can $|\mathrm{Crt}| = 2$ w.h.p. in the regime where $\mathbb{E}|\mathrm{Crt}| \geq e^{cN}$?
- Does trivialization have algorithmic implications, e.g. fast convergence of Langevin dynamics?
  - Or can regions of small, nonzero gradient still obstruct algorithms?

Guiding Questions

- Do phase boundaries of annealed and quenched trivialization coincide?
  - Or can $|\mathrm{Crt}| = 2$ w.h.p. in the regime where $\mathbb{E}|\mathrm{Crt}| \geq e^{cN}$?
- Does trivialization have algorithmic implications, e.g. fast convergence of Langevin dynamics?
  - Or can regions of small, nonzero gradient still obstruct algorithms?
- Can we boost $\mathbb{E}|\mathrm{Crt}| = e^{o(N)}$ to $|\mathrm{Crt}| = 2$ w.h.p. without exact formulas?

Strong Topological Trivialization

H is **strongly topologically trivial** (STT) if w.h.p.:

Strong Topological Trivialization

H is **strongly topologically trivial** (STT) if w.h.p.:

- $|\mathrm{Crt}| = O(1)$

Strong Topological Trivialization

H is **strongly topologically trivial** (STT) if w.h.p.:

- $|\mathrm{Crt}| = O(1)$
- Each ε-approximate critical point ($\|\nabla_{\mathrm{sp}} H(\boldsymbol{\sigma})\|_2 \le \varepsilon\sqrt{N}$) within $O(\varepsilon)\sqrt{N}$ of a true critical point

Strong Topological Trivialization

H is **strongly topologically trivial** (STT) if w.h.p.:

- $|\text{Crt}| = O(1)$
- Each ε-approximate critical point ($\|\nabla_{\text{sp}} H(\boldsymbol{\sigma})\|_2 \le \varepsilon\sqrt{N}$) within $O(\varepsilon)\sqrt{N}$ of a true critical point
- All critical points well-conditioned, i.e. $\text{spec}(\nabla^2_{\text{sp}} H(\boldsymbol{\sigma})) \cap [-c, c] = \emptyset$

Strong Topological Trivialization

H is **strongly topologically trivial** (STT) if w.h.p.:

- $|\mathrm{Crt}| = O(1)$
- Each ε-approximate critical point ($\|\nabla_{\mathrm{sp}} H(\boldsymbol{\sigma})\|_2 \le \varepsilon\sqrt{N}$) within $O(\varepsilon)\sqrt{N}$ of a true critical point
- All critical points well-conditioned, i.e. $\mathrm{spec}(\nabla^2_{\mathrm{sp}} H(\boldsymbol{\sigma})) \cap [-c, c] = \emptyset$
- At each critical point except the global maximum, $\lambda_{cN}(\nabla^2_{\mathrm{sp}} H(\boldsymbol{\sigma})) > 0$

Strong Topological Trivialization

H is **strongly topologically trivial** (STT) if w.h.p.:

- $|\mathrm{Crt}| = O(1)$
- Each ε-approximate critical point ($\|\nabla_{\mathrm{sp}} H(\boldsymbol{\sigma})\|_2 \le \varepsilon\sqrt{N}$) within $O(\varepsilon)\sqrt{N}$ of a true critical point
- All critical points well-conditioned, i.e. $\mathrm{spec}(\nabla^2_{\mathrm{sp}} H(\boldsymbol{\sigma})) \cap [-c, c] = \emptyset$
- At each critical point except the global maximum, $\lambda_{cN}(\nabla^2_{\mathrm{sp}} H(\boldsymbol{\sigma})) > 0$

Theorem (H.-Sellke 23)

If H is STT, then Langevin dynamics at low enough temperature mixes in $O(\log N)$ time in TV.

Main Result

Theorem (H.-Sellke 23)

1. *If* $\xi'(1) > \xi''(1)$*, STT holds (with 2 critical points).*

Main Result

Theorem (H.-Sellke 23)

1. *If* $\xi'(1) > \xi''(1)$*, STT holds (with 2 critical points).*
2. *If* $\xi'(1) < \xi''(1)$*, STT fails.*
   - *There are* e^{cN} *well-separated approximate critical points w.h.p.*

Main Result

Theorem (H.-Sellke 23)

1. *If* $\xi'(1) > \xi''(1)$*, STT holds (with 2 critical points).*
2. *If* $\xi'(1) < \xi''(1)$*, STT fails.*
   - *There are* e^{cN} *well-separated approximate critical points w.h.p.*

New method to boost $\mathbb{E}|\mathrm{Crt}| = e^{o(N)}$ to STT (and in particular $|\mathrm{Crt}| = 2$ w.h.p.) without exact RMT

Main Result

Theorem (H.-Sellke 23)

1. *If $\xi'(1) > \xi''(1)$, STT holds (with 2 critical points).*
2. *If $\xi'(1) < \xi''(1)$, STT fails.*
   - *There are e^{cN} well-separated approximate critical points w.h.p.*

New method to boost $\mathbb{E}|\mathrm{Crt}| = e^{o(N)}$ to STT (and in particular $|\mathrm{Crt}| = 2$ w.h.p.) without exact RMT

More general threshold for multi-species models (later)

Main Result

Theorem (H.-Sellke 23)

1. *If $\xi'(1) > \xi''(1)$, STT holds (with 2 critical points).*
2. *If $\xi'(1) < \xi''(1)$, STT fails.*
   - *There are e^{cN} well-separated approximate critical points w.h.p.*

New method to boost $\mathbb{E}|\mathrm{Crt}| = e^{o(N)}$ to STT (and in particular $|\mathrm{Crt}| = 2$ w.h.p.) without exact RMT

More general threshold for multi-species models (later)

- In these models, phase boundary for *annealed* trivialization also new

Annealed Critical Point Complexity

Kac-Rice on $\nabla_{\mathsf{sp}} H$:

$$\mathbb{E}|\mathrm{Crt}| = \int_{S_N} \mathbb{E}\left[|\det \nabla_{\mathsf{sp}}^2 H(\boldsymbol{\sigma})| \middle| \nabla_{\mathsf{sp}} H(\boldsymbol{\sigma}) = 0 \right] \varphi_{\nabla_{\mathsf{sp}} H(\boldsymbol{\sigma})}(0) \, \mathrm{d}\boldsymbol{\sigma}.$$

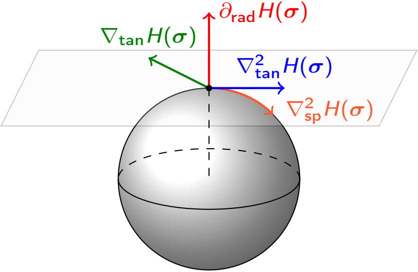
Annealed Critical Point Complexity

Kac-Rice on $\nabla_{\mathrm{sp}} H$:

$$\mathbb{E}|\mathrm{Crt}| = \int_{S_N} \mathbb{E}\left[|\det \nabla^2_{\mathrm{sp}} H(\boldsymbol{\sigma})| \,\middle|\, \nabla_{\mathrm{sp}} H(\boldsymbol{\sigma}) = 0 \right] \varphi_{\nabla_{\mathrm{sp}} H(\boldsymbol{\sigma})}(0)\, \mathrm{d}\boldsymbol{\sigma}.$$

Here ∇_{sp}, ∇^2_{sp} are **Riemannian gradient, Hessian**

$$\nabla_{\mathrm{sp}} H(\boldsymbol{\sigma}) = \nabla_{\tan} H(\boldsymbol{\sigma}), \qquad \nabla^2_{\mathrm{sp}} H(\boldsymbol{\sigma}) = \nabla^2_{\tan} H(\boldsymbol{\sigma}) - \underbrace{\partial_{\mathrm{rad}} H(\boldsymbol{\sigma}) I_{N-1}}_{\text{curvature term}}.$$



Annealed Critical Point Complexity

Kac-Rice on $\nabla_{\mathsf{sp}} H$:

$$\mathbb{E}|\mathsf{Crt}| = \int_{S_N} \mathbb{E}\left[|\det \nabla^2_{\mathsf{sp}} H(\boldsymbol{\sigma})| \middle| \nabla_{\mathsf{sp}} H(\boldsymbol{\sigma}) = 0 \right] \varphi_{\nabla_{\mathsf{sp}} H(\boldsymbol{\sigma})}(0) \, \mathsf{d}\boldsymbol{\sigma}.$$

Here ∇_{sp}, ∇^2_{sp} are **Riemannian gradient, Hessian**

$$\nabla_{\mathsf{sp}} H(\boldsymbol{\sigma}) = \nabla_{\mathsf{tan}} H(\boldsymbol{\sigma}), \qquad \nabla^2_{\mathsf{sp}} H(\boldsymbol{\sigma}) = \nabla^2_{\mathsf{tan}} H(\boldsymbol{\sigma}) - \partial_{\mathsf{rad}} H(\boldsymbol{\sigma}) I_{N-1}\,.$$

Annealed Critical Point Complexity

Kac-Rice on $\nabla_{\mathrm{sp}} H$:

$$\mathbb{E}|\mathrm{Crt}| = \mathbb{E}\left[|\det \nabla^2_{\mathrm{sp}} H(\boldsymbol{\sigma})| \middle| \nabla_{\mathrm{sp}} H(\boldsymbol{\sigma}) = 0 \right] \cdot \varphi_{\nabla_{\mathrm{sp}} H(\boldsymbol{\sigma})}(0) \cdot \mathrm{Vol}(S_N).$$

Here ∇_{sp}, ∇^2_{sp} are **Riemannian gradient, Hessian**

$$\nabla_{\mathrm{sp}} H(\boldsymbol{\sigma}) = \nabla_{\tan} H(\boldsymbol{\sigma}), \qquad \nabla^2_{\mathrm{sp}} H(\boldsymbol{\sigma}) = \nabla^2_{\tan} H(\boldsymbol{\sigma}) - \partial_{\mathrm{rad}} H(\boldsymbol{\sigma}) I_{N-1}\,.$$

Annealed Critical Point Complexity

Kac-Rice on $\nabla_{\mathrm{sp}} H$:

$$\mathbb{E}|\mathrm{Crt}| = \mathbb{E}\left[|\det \nabla^2_{\mathrm{sp}} H(\boldsymbol{\sigma})| \middle| \nabla_{\mathrm{sp}} H(\boldsymbol{\sigma}) = 0\right] \cdot \varphi_{\nabla_{\mathrm{sp}} H(\boldsymbol{\sigma})}(0) \cdot \mathrm{Vol}(S_N).$$

Here ∇_{sp}, ∇^2_{sp} are **Riemannian gradient, Hessian**

$$\nabla_{\mathrm{sp}} H(\boldsymbol{\sigma}) = \nabla_{\tan} H(\boldsymbol{\sigma}), \qquad \nabla^2_{\mathrm{sp}} H(\boldsymbol{\sigma}) = \nabla^2_{\tan} H(\boldsymbol{\sigma}) - \partial_{\mathrm{rad}} H(\boldsymbol{\sigma}) I_{N-1}\,.$$

Fact: $\partial_{\mathrm{rad}} H(\boldsymbol{\sigma})$, $\nabla^2_{\tan} H(\boldsymbol{\sigma})$, $\nabla_{\tan} H(\boldsymbol{\sigma})$ independent.

Annealed Critical Point Complexity

Kac-Rice on $\nabla_{\mathrm{sp}} H$:

$$\mathbb{E}|\mathrm{Crt}| = \mathbb{E}\left|\det(\nabla^2_{\tan} H(\boldsymbol{\sigma}) - \partial_{\mathrm{rad}} H(\boldsymbol{\sigma}) I_{N-1})\right| \cdot \varphi_{\nabla_{\tan} H(\boldsymbol{\sigma})}(0) \cdot \mathrm{Vol}(S_N).$$

Here ∇_{sp}, ∇^2_{sp} are **Riemannian gradient, Hessian**

$$\nabla_{\mathrm{sp}} H(\boldsymbol{\sigma}) = \nabla_{\tan} H(\boldsymbol{\sigma}), \qquad \nabla^2_{\mathrm{sp}} H(\boldsymbol{\sigma}) = \nabla^2_{\tan} H(\boldsymbol{\sigma}) - \partial_{\mathrm{rad}} H(\boldsymbol{\sigma}) I_{N-1}\,.$$

Fact: $\partial_{\mathrm{rad}} H(\boldsymbol{\sigma})$, $\nabla^2_{\tan} H(\boldsymbol{\sigma})$, $\nabla_{\tan} H(\boldsymbol{\sigma})$ independent.

Annealed Critical Point Complexity

Kac-Rice on $\nabla_{\mathrm{sp}} H$:

$$\mathbb{E}|\mathrm{Crt}| = \mathbb{E}\left|\det(\nabla^2_{\tan} H(\boldsymbol{\sigma}) - \partial_{\mathrm{rad}} H(\boldsymbol{\sigma}) I_{N-1})\right| \cdot \varphi_{\nabla_{\tan} H(\boldsymbol{\sigma})}(0) \cdot \mathrm{Vol}(S_N).$$

Here ∇_{sp}, ∇^2_{sp} are **Riemannian gradient, Hessian**

$$\nabla_{\mathrm{sp}} H(\boldsymbol{\sigma}) = \nabla_{\tan} H(\boldsymbol{\sigma}), \qquad \nabla^2_{\mathrm{sp}} H(\boldsymbol{\sigma}) = \nabla^2_{\tan} H(\boldsymbol{\sigma}) - \partial_{\mathrm{rad}} H(\boldsymbol{\sigma}) I_{N-1}\,.$$

Fact: $\partial_{\mathrm{rad}} H(\boldsymbol{\sigma})$, $\nabla^2_{\tan} H(\boldsymbol{\sigma})$, $\nabla_{\tan} H(\boldsymbol{\sigma})$ independent. They have laws

$$\partial_{\mathrm{rad}} H(\boldsymbol{\sigma}) =_d c_1 \mathcal{N}(0, 1/N) \qquad \nabla^2_{\tan} H(\boldsymbol{\sigma}) =_d c_2 \mathrm{GOE}_{N-1}$$

$$\nabla_{\tan} H(\boldsymbol{\sigma}) =_d c_3 \mathcal{N}(0, I_{N-1})$$

Annealed Critical Point Complexity

Kac-Rice on $\nabla_{\mathrm{sp}} H$:

$$\mathbb{E}|\mathrm{Crt}| = \mathbb{E}\left|\det(\nabla^2_{\tan} H(\boldsymbol{\sigma}) - \partial_{\mathrm{rad}} H(\boldsymbol{\sigma}) I_{N-1})\right| \cdot \varphi_{\nabla_{\tan} H(\boldsymbol{\sigma})}(0) \cdot \mathrm{Vol}(S_N).$$

Here ∇_{sp}, ∇^2_{sp} are **Riemannian gradient, Hessian**

$$\nabla_{\mathrm{sp}} H(\boldsymbol{\sigma}) = \nabla_{\tan} H(\boldsymbol{\sigma}), \qquad \nabla^2_{\mathrm{sp}} H(\boldsymbol{\sigma}) = \nabla^2_{\tan} H(\boldsymbol{\sigma}) - \partial_{\mathrm{rad}} H(\boldsymbol{\sigma}) I_{N-1}\,.$$

Fact: $\partial_{\mathrm{rad}} H(\boldsymbol{\sigma})$, $\nabla^2_{\tan} H(\boldsymbol{\sigma})$, $\nabla_{\tan} H(\boldsymbol{\sigma})$ independent. They have laws

$$\partial_{\mathrm{rad}} H(\boldsymbol{\sigma}) =_d c_1 \mathcal{N}(0, 1/N) \qquad \nabla^2_{\tan} H(\boldsymbol{\sigma}) =_d c_2 \mathrm{GOE}_{N-1}$$
$$\nabla_{\tan} H(\boldsymbol{\sigma}) =_d c_3 \mathcal{N}(0, I_{N-1})$$

Integrate out $\partial_{\mathrm{rad}} H(\boldsymbol{\sigma}) = x$:

$$\mathbb{E}|\mathrm{Crt}| = (\text{simple term}) \times \int_{\mathbb{R}} \mathbb{E}\left|\det(c_2 \mathrm{GOE}_{N-1} - xI)\right| \exp\left(-\frac{Nx^2}{2c_1}\right) \,\mathrm{d}x$$

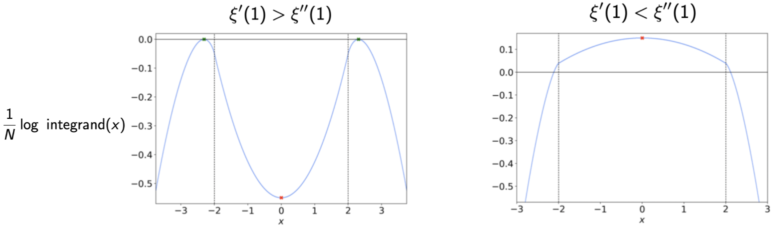
Annealed Critical Point Complexity

$$\mathbb{E}|\mathrm{Crt}| = (\text{simple term}) \times \int_{\mathbb{R}} \mathbb{E}\left|\det(c_2 \mathrm{GOE}_{N-1} - xI)\right| \exp\left(-\frac{Nx^2}{2c_1}\right) \, \mathrm{d}x$$

Annealed Critical Point Complexity

$$\mathbb{E}|\mathrm{Crt}| = (\text{simple term}) \times \int_{\mathbb{R}} \mathbb{E}\left|\det(c_2 \mathrm{GOE}_{N-1} - xI)\right| \exp\left(-\frac{Nx^2}{2c_1}\right) \, \mathrm{d}x$$

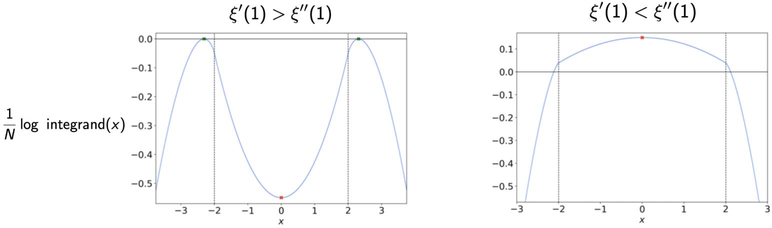
Laplace's principle $\Rightarrow$ maximize exponential rate of integrand



Annealed Critical Point Complexity

$$\mathbb{E}|\mathrm{Crt}| = (\text{simple term}) \times \int_{\mathbb{R}} \mathbb{E}\left|\det(c_2 \mathrm{GOE}_{N-1} - xI)\right| \exp\left(-\frac{Nx^2}{2c_1}\right) \mathrm{d}x$$

Laplace's principle $\Rightarrow$ maximize exponential rate of integrand

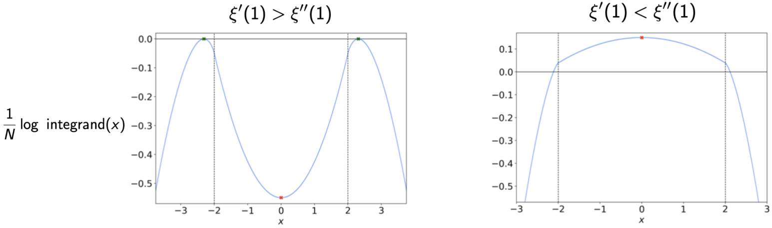


- $\xi'(1) > \xi''(1) \Rightarrow \mathbb{E}|\mathrm{Crt}| = e^{o(N)}$, (weak form of) annealed trivialization

Annealed Critical Point Complexity

$$\mathbb{E}|\mathrm{Crt}| = (\text{simple term}) \times \int_{\mathbb{R}} \mathbb{E}\left|\det(c_2 \mathrm{GOE}_{N-1} - xI)\right| \exp\left(-\frac{Nx^2}{2c_1}\right)\, \mathrm{d}x$$

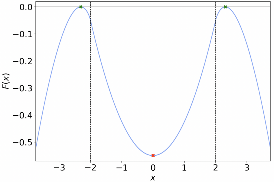
Laplace's principle $\Rightarrow$ maximize exponential rate of integrand



- $\xi'(1) > \xi''(1) \Rightarrow \mathbb{E}|\mathrm{Crt}| = e^{o(N)}$, (weak form of) annealed trivialization
- $\xi'(1) < \xi''(1) \Rightarrow \mathbb{E}|\mathrm{Crt}| \geq e^{cN}$, failure of annealed trivialization

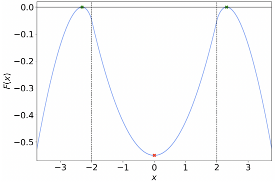
Statistics of Critical Points

Suppose we are in trivial regime $\xi'(1) > \xi''(1)$, so $\mathbb{E}|\mathrm{Crt}| = e^{o(N)}$



Statistics of Critical Points

Suppose we are in trivial regime $\xi'(1) > \xi''(1)$, so $\mathbb{E}|\mathrm{Crt}| = e^{o(N)}$

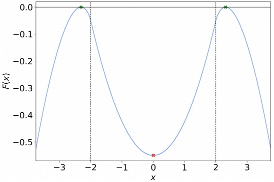


For all $\boldsymbol{\sigma} \in \mathrm{Crt}$,

$$\partial_{\mathrm{rad}} H(\boldsymbol{\sigma}) = \pm x_{\mathrm{OPT}} + o(1)$$

Statistics of Critical Points

Suppose we are in trivial regime $\xi'(1) > \xi''(1)$, so $\mathbb{E}|\mathrm{Crt}| = e^{o(N)}$



For all $\boldsymbol{\sigma} \in \mathrm{Crt}$,

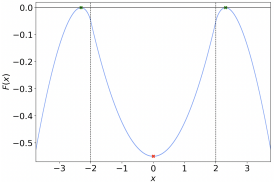
$$\partial_{\mathrm{rad}} H(\boldsymbol{\sigma}) = \pm x_{\mathrm{OPT}} + o(1)$$

In fact,

$$\left(\partial_{\mathrm{rad}} H(\boldsymbol{\sigma}), \tfrac{1}{N} H(\boldsymbol{\sigma}), \tfrac{1}{N} \langle \boldsymbol{G}^{(1)}, \boldsymbol{\sigma} \rangle \right) = \pm (x_{\mathrm{OPT}}, E_{\mathrm{OPT}}, R_{\mathrm{OPT}}) + o(1)$$

Statistics of Critical Points

Suppose we are in trivial regime $\xi'(1) > \xi''(1)$, so $\mathbb{E}|\mathrm{Crt}| = e^{o(N)}$



For all $\boldsymbol{\sigma} \in \mathrm{Crt}$,

$$\partial_{\mathrm{rad}} H(\boldsymbol{\sigma}) = \pm x_{\mathrm{OPT}} + o(1)$$

In fact,

$$\left(\partial_{\mathrm{rad}} H(\boldsymbol{\sigma}), \tfrac{1}{N} H(\boldsymbol{\sigma}), \tfrac{1}{N}\langle \boldsymbol{G}^{(1)}, \boldsymbol{\sigma}\rangle\right) = \pm(x_{\mathrm{OPT}}, E_{\mathrm{OPT}}, R_{\mathrm{OPT}}) + o(1)$$

x_{OPT} outside spectral bulk $[-2c_2, 2c_2]$ of $\nabla^2_{\mathrm{tan}} H(\boldsymbol{\sigma})$ so

$$\nabla^2_{\mathrm{sp}} H(\boldsymbol{\sigma}) = \nabla^2_{\mathrm{tan}} H(\boldsymbol{\sigma}) - \partial_{\mathrm{rad}} H(\boldsymbol{\sigma}) I_N \qquad \text{well-conditioned}$$

Attempt: from $e^{o(N)}$ Critical Points to 2

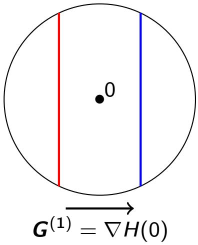
For all $\boldsymbol{\sigma} \in \mathrm{Crt}$, $\frac{1}{N}\langle \boldsymbol{G}^{(1)}, \boldsymbol{\sigma} \rangle = \pm R_{\mathsf{OPT}} + o(1)$.

Attempt: from $e^{o(N)}$ Critical Points to 2

For all $\boldsymbol{\sigma} \in \mathrm{Crt}$, $\frac{1}{N}\langle \boldsymbol{G}^{(1)}, \boldsymbol{\sigma} \rangle = \pm R_{\mathsf{OPT}} + o(1)$. **Suppose** $\frac{1}{N}\langle \boldsymbol{G}^{(1)}, \boldsymbol{\sigma} \rangle = \pm R_{\mathsf{OPT}}$

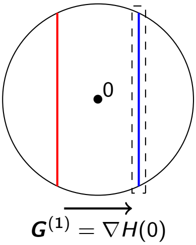
Attempt: from $e^{o(N)}$ Critical Points to 2

For all $\boldsymbol{\sigma} \in \text{Crt}$, $\frac{1}{N}\langle \boldsymbol{G}^{(1)}, \boldsymbol{\sigma} \rangle = \pm R_{\text{OPT}} + o(1)$. **Suppose** $\frac{1}{N}\langle \boldsymbol{G}^{(1)}, \boldsymbol{\sigma} \rangle = \pm R_{\text{OPT}}$



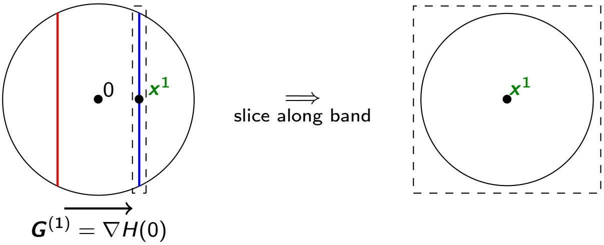
Attempt: from $e^{o(N)}$ Critical Points to 2

For all $\boldsymbol{\sigma} \in \mathrm{Crt}$, $\frac{1}{N}\langle \boldsymbol{G}^{(1)}, \boldsymbol{\sigma} \rangle = \pm R_{\mathrm{OPT}} + o(1)$. **Suppose** $\frac{1}{N}\langle \boldsymbol{G}^{(1)}, \boldsymbol{\sigma} \rangle = \pm R_{\mathrm{OPT}}$



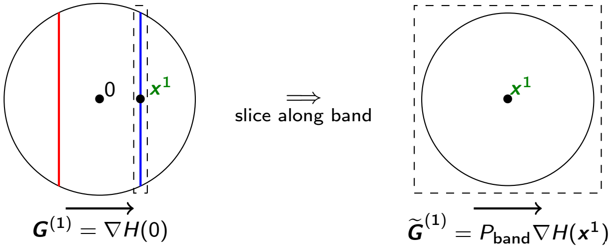
Attempt: from $e^{o(N)}$ Critical Points to 2

For all $\boldsymbol{\sigma} \in \mathrm{Crt}$, $\frac{1}{N}\langle \boldsymbol{G}^{(1)}, \boldsymbol{\sigma} \rangle = \pm R_{\mathsf{OPT}} + o(1)$. **Suppose $\frac{1}{N}\langle \boldsymbol{G}^{(1)}, \boldsymbol{\sigma} \rangle = \pm R_{\mathsf{OPT}}$**



Attempt: from $e^{o(N)}$ Critical Points to 2

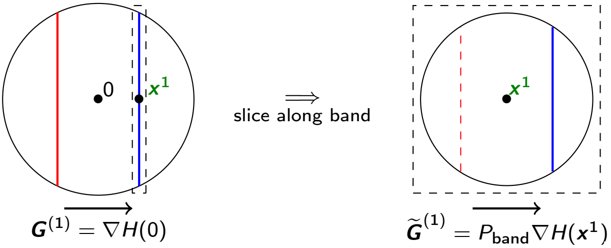
For all $\boldsymbol{\sigma} \in \mathrm{Crt}$, $\frac{1}{N}\langle \boldsymbol{G}^{(1)}, \boldsymbol{\sigma}\rangle = \pm R_{\mathrm{OPT}} + o(1)$. **Suppose** $\frac{1}{N}\langle \boldsymbol{G}^{(1)}, \boldsymbol{\sigma}\rangle = \pm R_{\mathrm{OPT}}$



- Conditional on $\boldsymbol{G}^{(1)}$, restriction to band is $(N-1)$-dimensional spin glass

Attempt: from $e^{o(N)}$ Critical Points to 2

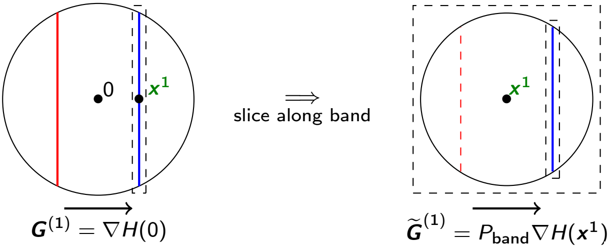
For all $\boldsymbol{\sigma} \in \mathrm{Crt}$, $\frac{1}{N}\langle \boldsymbol{G}^{(1)}, \boldsymbol{\sigma} \rangle = \pm R_{\mathrm{OPT}} + o(1)$. **Suppose** $\frac{1}{N}\langle \boldsymbol{G}^{(1)}, \boldsymbol{\sigma} \rangle = \pm R_{\mathrm{OPT}}$



- Conditional on $\boldsymbol{G}^{(1)}$, restriction to band is $(N-1)$-dimensional spin glass
  - Its critical points have known correlation with field $\widetilde{\boldsymbol{G}}^{(1)}$

Attempt: from $e^{o(N)}$ Critical Points to 2

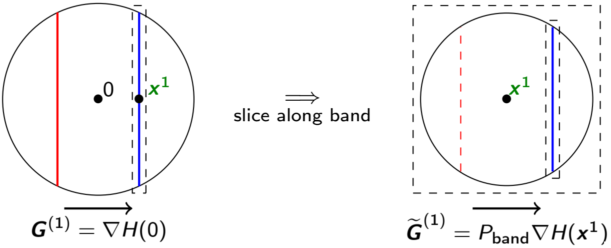
For all $\boldsymbol{\sigma} \in \mathrm{Crt}$, $\frac{1}{N}\langle \boldsymbol{G}^{(1)}, \boldsymbol{\sigma} \rangle = \pm R_{\mathsf{OPT}} + o(1)$. **Suppose** $\frac{1}{N}\langle \boldsymbol{G}^{(1)}, \boldsymbol{\sigma} \rangle = \pm R_{\mathsf{OPT}}$



- Conditional on $\boldsymbol{G}^{(1)}$, restriction to band is $(N-1)$-dimensional spin glass
  - Its critical points have known correlation with field $\widetilde{\boldsymbol{G}}^{(1)}$

Attempt: from $e^{o(N)}$ Critical Points to 2

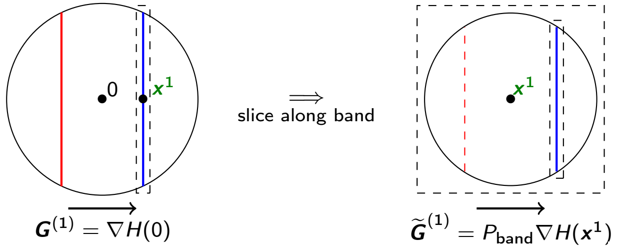
For all $\boldsymbol{\sigma} \in \mathrm{Crt}$, $\frac{1}{N}\langle \boldsymbol{G}^{(1)}, \boldsymbol{\sigma} \rangle = \pm R_{\mathrm{OPT}} + o(1)$. **Suppose** $\frac{1}{N}\langle \boldsymbol{G}^{(1)}, \boldsymbol{\sigma} \rangle = \pm R_{\mathrm{OPT}}$



- Conditional on $\boldsymbol{G}^{(1)}$, restriction to band is $(N-1)$-dimensional spin glass
  - Its critical points have known correlation with field $\widetilde{\boldsymbol{G}}^{(1)}$
- Bands shrink. All critical points localize to 2 regions of diameter $o(\sqrt{N})$

Attempt: from $e^{o(N)}$ Critical Points to 2

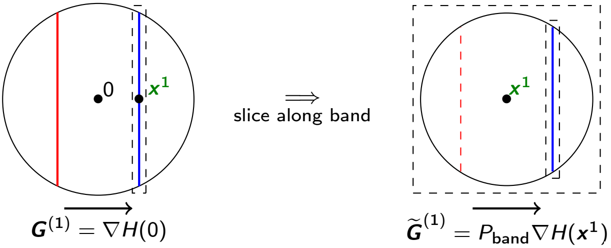
For all $\boldsymbol{\sigma} \in \mathrm{Crt}$, $\frac{1}{N}\langle \boldsymbol{G}^{(1)}, \boldsymbol{\sigma} \rangle = \pm R_{\mathrm{OPT}} + o(1)$. **Suppose $\frac{1}{N}\langle \boldsymbol{G}^{(1)}, \boldsymbol{\sigma} \rangle = \pm R_{\mathrm{OPT}}$**



- Conditional on $\boldsymbol{G}^{(1)}$, restriction to band is $(N-1)$-dimensional spin glass
    - Its critical points have known correlation with field $\widetilde{\boldsymbol{G}}^{(1)}$
- Bands shrink. All critical points localize to 2 regions of diameter $o(\sqrt{N})$
- $\nabla^2_{\mathrm{sp}} H(\boldsymbol{\sigma})$ well-conditioned $\Rightarrow$ each region has ≤ 1 critical point

Attempt: from $e^{o(N)}$ Critical Points to 2

For all $\boldsymbol{\sigma} \in \mathrm{Crt}$, $\frac{1}{N}\langle \boldsymbol{G}^{(1)}, \boldsymbol{\sigma} \rangle = \pm R_{\mathsf{OPT}} + o(1)$. **Suppose** $\frac{1}{N}\langle \boldsymbol{G}^{(1)}, \boldsymbol{\sigma} \rangle = \pm R_{\mathsf{OPT}}$



- Conditional on $\boldsymbol{G}^{(1)}$, restriction to band is $(N-1)$-dimensional spin glass
  - Its critical points have known correlation with field $\widetilde{\boldsymbol{G}}^{(1)}$
- Bands shrink. All critical points localize to 2 regions of diameter $o(\sqrt{N})$
- $\nabla^2_{\mathsf{sp}} H(\boldsymbol{\sigma})$ well-conditioned $\Rightarrow$ each region has ≤ 1 critical point

- **This does not work!** Critical points brittle, cannot tolerate $o(1)$ error ☹

Approximate to Exact Critical Points

Lemma (H.-Sellke 23)

With probability $1 - e^{-cN}$, *all* ε*-approximate critical points* $\boldsymbol{\sigma}$ *of* H *satisfy:*

$$\left\| \left(\partial_{\text{rad}} H(\boldsymbol{\sigma}), \tfrac{1}{N} H(\boldsymbol{\sigma}), \tfrac{1}{N} \langle \boldsymbol{G}^{(1)}, \boldsymbol{\sigma} \rangle \right) \pm (x_{\text{OPT}}, E_{\text{OPT}}, R_{\text{OPT}}) \right\| \le o(1) + \iota(\varepsilon) \qquad (*)$$

and $\nabla^2_{\text{sp}} H(\boldsymbol{\sigma})$ *is well-conditioned.*

Approximate to Exact Critical Points

Lemma (H.-Sellke 23)

With probability $1 - e^{-cN}$, all ε-approximate critical points $\boldsymbol{\sigma}$ of H satisfy:

$$\left\| \left(\partial_{\mathrm{rad}} H(\boldsymbol{\sigma}), \tfrac{1}{N} H(\boldsymbol{\sigma}), \tfrac{1}{N} \langle \boldsymbol{G}^{(1)}, \boldsymbol{\sigma} \rangle \right) \pm (x_{\mathrm{OPT}}, E_{\mathrm{OPT}}, R_{\mathrm{OPT}}) \right\| \le o(1) + \iota(\varepsilon) \qquad (*)$$

and $\nabla^2_{\mathrm{sp}} H(\boldsymbol{\sigma})$ is well-conditioned.

- Proof idea: consider rerandomized Hamiltonian

$$H^{\delta} = \sqrt{1-\delta} H + \sqrt{\delta} H'.$$

Approximate to Exact Critical Points

Lemma (H.-Sellke 23)

With probability $1 - e^{-cN}$, *all* ε*-approximate critical points* $\boldsymbol{\sigma}$ *of* H *satisfy:*

$$\left\| \left(\partial_{\mathrm{rad}} H(\boldsymbol{\sigma}), \tfrac{1}{N} H(\boldsymbol{\sigma}), \tfrac{1}{N} \langle \boldsymbol{G}^{(1)}, \boldsymbol{\sigma} \rangle \right) \pm (x_{\mathrm{OPT}}, E_{\mathrm{OPT}}, R_{\mathrm{OPT}}) \right\| \le o(1) + \iota(\varepsilon) \qquad (*)$$

and $\nabla^2_{\mathrm{sp}} H(\boldsymbol{\sigma})$ *is well-conditioned.*

- Proof idea: consider rerandomized Hamiltonian

$$H^\delta = \sqrt{1-\delta} H + \sqrt{\delta} H'.$$

- Conditional on H having an ε-approx crit violating $(*)$, H^δ has expected $e^{-o_\varepsilon(1)N}$ **exact** crits violating $(*)$

Approximate to Exact Critical Points

Lemma (H.-Sellke 23)

With probability $1 - e^{-cN}$, *all* ε*-approximate critical points* $\boldsymbol{\sigma}$ *of* H *satisfy:*

$$\left\| \left(\partial_{\mathrm{rad}} H(\boldsymbol{\sigma}), \tfrac{1}{N} H(\boldsymbol{\sigma}), \tfrac{1}{N} \langle \boldsymbol{G}^{(1)}, \boldsymbol{\sigma} \rangle \right) \pm (x_{\mathrm{OPT}}, E_{\mathrm{OPT}}, R_{\mathrm{OPT}}) \right\| \le o(1) + \iota(\varepsilon) \qquad (*)$$

and $\nabla^2_{\mathrm{sp}} H(\boldsymbol{\sigma})$ *is well-conditioned.*

- Proof idea: consider rerandomized Hamiltionian

$$H^\delta = \sqrt{1-\delta} H + \sqrt{\delta} H'.$$

- Conditional on H having an ε-approx crit violating $(*)$, H^δ has expected $e^{-o_\varepsilon(1) N}$ **exact** crits violating $(*)$
  - Proved by Kac-Rice **conditionally** on H

Approximate to Exact Critical Points

Lemma (H.-Sellke 23)

With probability $1 - e^{-cN}$, all ε-approximate critical points $\boldsymbol{\sigma}$ of H satisfy:

$$\left\| \left(\partial_{\mathrm{rad}} H(\boldsymbol{\sigma}), \tfrac{1}{N} H(\boldsymbol{\sigma}), \tfrac{1}{N} \langle \boldsymbol{G}^{(1)}, \boldsymbol{\sigma} \rangle \right) \pm (x_{\mathrm{OPT}}, E_{\mathrm{OPT}}, R_{\mathrm{OPT}}) \right\| \le o(1) + \iota(\varepsilon) \qquad (*)$$

and $\nabla^2_{\mathrm{sp}} H(\boldsymbol{\sigma})$ is well-conditioned.

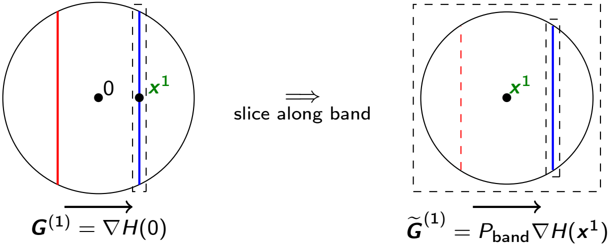
- Proof idea: consider rerandomized Hamiltionian

$$H^\delta = \sqrt{1-\delta} H + \sqrt{\delta} H'.$$

- Conditional on H having an ε-approx crit violating $(*)$, H^δ has expected $e^{-o_\varepsilon(1)N}$ **exact** crits violating $(*)$
  - Proved by Kac-Rice **conditionally** on H
- But there are $\le e^{-c'N}$ such points in (unconditional) expectation!

Band Recursion and Strong Trivialization

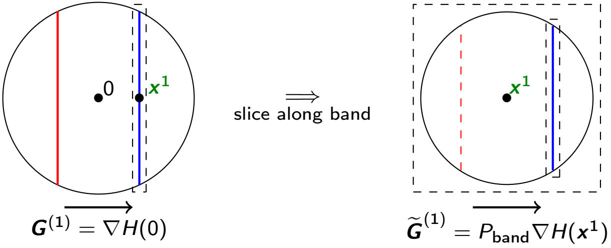
For any ε-approximate critical point $\boldsymbol{\sigma}$, $\frac{1}{N}\langle \boldsymbol{G}^{(1)}, \boldsymbol{\sigma}\rangle = \pm R_{\mathsf{OPT}} + o_\varepsilon(1)$



- Recursion localizes all approx crits to 2 regions of diameter $o(\sqrt{N})$

Band Recursion and Strong Trivialization

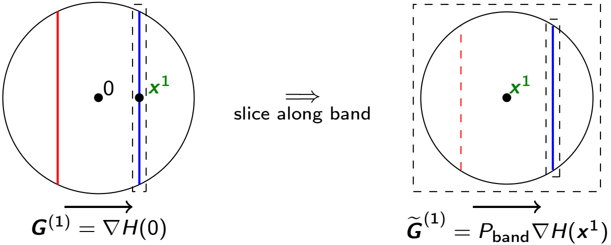
For any ε-approximate critical point $\boldsymbol{\sigma}$, $\frac{1}{N}\langle \boldsymbol{G}^{(1)}, \boldsymbol{\sigma} \rangle = \pm R_{\mathsf{OPT}} + o_\varepsilon(1)$



- Recursion localizes all approx crits to 2 regions of diameter $o(\sqrt{N})$
- $\nabla^2_{\mathsf{sp}} H(\boldsymbol{\sigma})$ well-conditioned $\Rightarrow$ each region has ≤ 1 exact crit

Band Recursion and Strong Trivialization

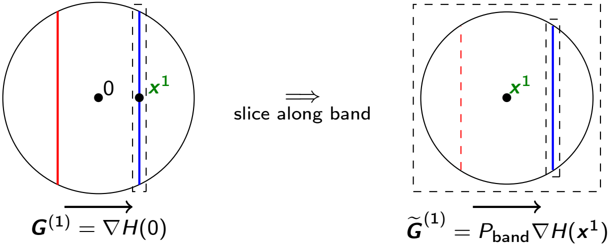
For any ε-approximate critical point $\boldsymbol{\sigma}$, $\frac{1}{N}\langle \boldsymbol{G}^{(1)}, \boldsymbol{\sigma} \rangle = \pm R_{\mathsf{OPT}} + o_\varepsilon(1)$



- Recursion localizes all approx crits to 2 regions of diameter $o(\sqrt{N})$
- $\nabla^2_{\text{sp}} H(\boldsymbol{\sigma})$ well-conditioned $\Rightarrow$ each region has ≤ 1 exact crit
- ≥ 2 exact crits globally $\Rightarrow$ each region has exactly 1

Band Recursion and Strong Trivialization

For any ε-approximate critical point $\boldsymbol{\sigma}$, $\frac{1}{N}\langle \boldsymbol{G}^{(1)}, \boldsymbol{\sigma}\rangle = \pm R_{\mathrm{OPT}} + o_\varepsilon(1)$



- Recursion localizes all approx crits to 2 regions of diameter $o(\sqrt{N})$
- $\nabla^2_{\mathrm{sp}} H(\boldsymbol{\sigma})$ well-conditioned $\Rightarrow$ each region has ≤ 1 exact crit
- ≥ 2 exact crits globally $\Rightarrow$ each region has exactly 1

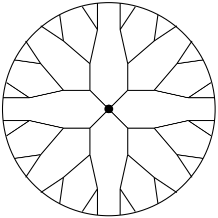
- So 2 exact crits, all approx crits near an exact crit. **Strong trivialization**!

e^{cN} Approximate Critical Points when $\xi'(1) < \xi''(1)$

Branching OGP (H.-Sellke 21, 23): optimization algorithms can reach states forming ultrametric tree with **random** root correlated with $\boldsymbol{G}^{(1)}$

e^{cN} Approximate Critical Points when $\xi'(1) < \xi''(1)$

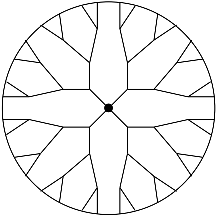
Branching OGP (H.-Sellke 21, 23): optimization algorithms can reach states forming ultrametric tree with **random** root correlated with $\boldsymbol{G}^{(1)}$



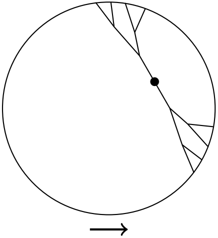
$\boldsymbol{G}^{(1)} = 0$

e^{cN} Approximate Critical Points when $\xi'(1) < \xi''(1)$

Branching OGP (H.-Sellke 21, 23): optimization algorithms can reach states forming ultrametric tree with **random** root correlated with $\boldsymbol{G}^{(1)}$



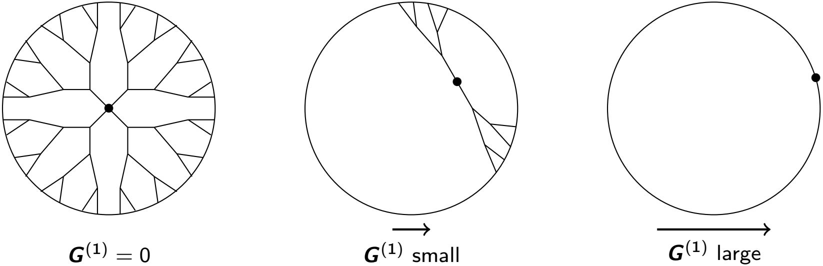
$\boldsymbol{G}^{(1)} = 0$



$\boldsymbol{G}^{(1)}$ small

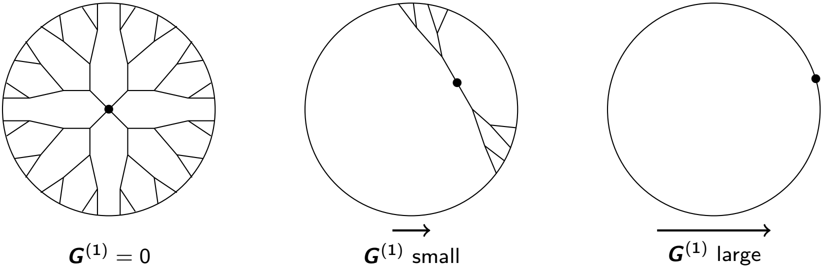
e^{cN} Approximate Critical Points when $\xi'(1) < \xi''(1)$

Branching OGP (H.-Sellke 21, 23): optimization algorithms can reach states forming ultrametric tree with **random** root correlated with $\boldsymbol{G}^{(1)}$



e^{cN} Approximate Critical Points when $\xi'(1) < \xi''(1)$

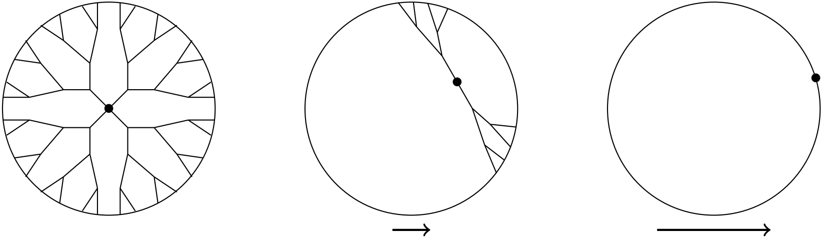
Branching OGP (H.-Sellke 21, 23): optimization algorithms can reach states forming ultrametric tree with **random** root correlated with $\boldsymbol{G}^{(1)}$



- Algorithmic tree is non-degenerate precisely when $\xi'(1) < \xi''(1)$

e^{cN} Approximate Critical Points when $\xi'(1) < \xi''(1)$

Branching OGP (H.-Sellke 21, 23): optimization algorithms can reach states forming ultrametric tree with **random** root correlated with $\boldsymbol{G}^{(1)}$



$\boldsymbol{G}^{(1)} = 0$ $\boldsymbol{G}^{(1)}$ small $\boldsymbol{G}^{(1)}$ large

- Algorithmic tree is non-degenerate precisely when $\xi'(1) < \xi''(1)$
- AMP constructs e^{cN} points in algorithmic tree, all approximate critical points

Multi-Species Models

- Up to now: polynomials in variables $\sigma_1, \ldots, \sigma_N$ that **all look alike**
- Multi-species models: $r = O(1)$ different "variable types"

Multi-Species Models

- Up to now: polynomials in variables $\sigma_1, \ldots, \sigma_N$ that **all look alike**
- Multi-species models: $r = O(1)$ different "variable types"
- Formally, each coordinate part of a **species** $s \in \mathscr{S} = \{1, \ldots, r\}$

$$[N] = \mathcal{I}_1 \cup \cdots \cup \mathcal{I}_r, \qquad |\mathcal{I}_s| = \lambda_s N$$

Multi-Species Models

- Up to now: polynomials in variables $\sigma_1, \ldots, \sigma_N$ that **all look alike**
- Multi-species models: $r = O(1)$ different "variable types"

- Formally, each coordinate part of a **species** $s \in \mathscr{S} = \{1, \ldots, r\}$

$$[N] = \mathcal{I}_1 \cup \cdots \cup \mathcal{I}_r, \qquad |\mathcal{I}_s| = \lambda_s N$$

- Interaction weights $\gamma_2, \gamma_3, \ldots$ now $(\gamma_{s_1,s_2})_{s_1,s_2 \in \mathscr{S}}, (\gamma_{s_1,s_2,s_3})_{s_1,s_2,s_3 \in \mathscr{S}}, \ldots$

Multi-Species Models

- Up to now: polynomials in variables $\sigma_1, \ldots, \sigma_N$ that **all look alike**
- Multi-species models: $r = O(1)$ different "variable types"

- Formally, each coordinate part of a **species** $s \in \mathscr{S} = \{1, \ldots, r\}$

$$[N] = \mathcal{I}_1 \cup \cdots \cup \mathcal{I}_r, \qquad |\mathcal{I}_s| = \lambda_s N$$

- Interaction weights $\gamma_2, \gamma_3, \ldots$ now $(\gamma_{s_1,s_2})_{s_1,s_2 \in \mathscr{S}}, (\gamma_{s_1,s_2,s_3})_{s_1,s_2,s_3 \in \mathscr{S}}, \ldots$
- Input space now **product of spheres**

$$\mathcal{P}_N = \left\{ \boldsymbol{\sigma} \in \mathbb{R}^N : \|\boldsymbol{\sigma}_{|\mathcal{I}_s}\|_2^2 = \lambda_s N \quad \forall s \in \mathscr{S} \right\}$$

Multi-Species Models

- Up to now: polynomials in variables $\sigma_1, \dots, \sigma_N$ that **all look alike**
- Multi-species models: $r = O(1)$ different "variable types"

- Formally, each coordinate part of a **species** $s \in \mathscr{S} = \{1, \dots, r\}$

$$[N] = \mathcal{I}_1 \cup \cdots \cup \mathcal{I}_r, \qquad |\mathcal{I}_s| = \lambda_s N$$

- Interaction weights $\gamma_2, \gamma_3, \dots$ now $(\gamma_{s_1,s_2})_{s_1,s_2 \in \mathscr{S}}, (\gamma_{s_1,s_2,s_3})_{s_1,s_2,s_3 \in \mathscr{S}}, \dots$
- Input space now **product of spheres**

$$\mathcal{P}_N = \left\{ \boldsymbol{\sigma} \in \mathbb{R}^N : \|\boldsymbol{\sigma}_{|\mathcal{I}_s}\|_2^2 = \lambda_s N \quad \forall s \in \mathscr{S} \right\}$$

- Example: **bipartite SK** $H(\boldsymbol{x}, \boldsymbol{y}) = \langle \boldsymbol{x}, \boldsymbol{G}\boldsymbol{y} \rangle$ where $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{S}_{N/2}$

Main Result for Multi-Species Models

$\xi(q_1, \ldots, q_r)$ now r-variate polynomial (so $\nabla\xi \in \mathbb{R}^r$, $\nabla^2\xi \in \mathbb{R}^{r\times r}$)

Main Result for Multi-Species Models

$\xi(q_1, \ldots, q_r)$ now r-variate polynomial (so $\nabla\xi \in \mathbb{R}^r$, $\nabla^2\xi \in \mathbb{R}^{r\times r}$)

Theorem (H.-Sellke 23)

If $\mathrm{diag}(\nabla\xi(\vec{1})) \succ \nabla^2\xi(\vec{1})$, *both annealed and strong trivialization hold.*

Main Result for Multi-Species Models

$\xi(q_1, \ldots, q_r)$ now r-variate polynomial (so $\nabla \xi \in \mathbb{R}^r$, $\nabla^2 \xi \in \mathbb{R}^{r \times r}$)

Theorem (H.-Sellke 23)

If $\mathrm{diag}(\nabla \xi(\vec{1})) \succ \nabla^2 \xi(\vec{1})$, *both annealed and strong trivialization hold.*

- $\mathbb{E}|\mathrm{Crt}| = e^{o(N)}$
- *W.h.p. exactly* 2^r *critical points, all well-conditioned, all approx crits near one*

Main Result for Multi-Species Models

$\xi(q_1, \ldots, q_r)$ now r-variate polynomial (so $\nabla\xi \in \mathbb{R}^r$, $\nabla^2\xi \in \mathbb{R}^{r\times r}$)

Theorem (H.-Sellke 23)

If $\operatorname{diag}(\nabla\xi(\vec{1})) \succ \nabla^2\xi(\vec{1})$, *both annealed and strong trivialization hold.*

- $\mathbb{E}|\mathrm{Crt}| = e^{o(N)}$
- *W.h.p. exactly* 2^r *critical points, all well-conditioned, all approx crits near one*
- *Determine energy,* r *radial derivatives,* r *correlations with external field at each*

Main Result for Multi-Species Models

$\xi(q_1, \ldots, q_r)$ now r-variate polynomial (so $\nabla\xi \in \mathbb{R}^r$, $\nabla^2\xi \in \mathbb{R}^{r\times r}$)

Theorem (H.-Sellke 23)

If $\mathrm{diag}(\nabla\xi(\vec{1})) \succ \nabla^2\xi(\vec{1})$, *both annealed and strong trivialization hold.*

- $\mathbb{E}|\mathrm{Crt}| = e^{o(N)}$
- *W.h.p. exactly* 2^r *critical points, all well-conditioned, all approx crits near one*
- *Determine energy,* r *radial derivatives,* r *correlations with external field at each*

- 2^r critical points minimal for any Morse function on product of r spheres

Main Result for Multi-Species Models

$\xi(q_1, \ldots, q_r)$ now r-variate polynomial (so $\nabla \xi \in \mathbb{R}^r$, $\nabla^2 \xi \in \mathbb{R}^{r \times r}$)

Theorem (H.-Sellke 23)

If $\mathrm{diag}(\nabla \xi(\vec{1})) \succ \nabla^2 \xi(\vec{1})$, *both annealed and strong trivialization hold.*

- $\mathbb{E}|\mathrm{Crt}| = e^{o(N)}$
- *W.h.p. exactly* 2^r *critical points, all well-conditioned, all approx crits near one*
- *Determine energy,* r *radial derivatives,* r *correlations with external field at each*

- 2^r critical points minimal for any Morse function on product of r spheres
- Consequence: ground state energy when $\mathrm{diag}(\nabla \xi(\vec{1})) \succ \nabla^2 \xi(\vec{1})$

Main Result for Multi-Species Models

$\xi(q_1, \ldots, q_r)$ now r-variate polynomial (so $\nabla\xi \in \mathbb{R}^r$, $\nabla^2\xi \in \mathbb{R}^{r\times r}$)

Theorem (H.-Sellke 23)

If $\mathrm{diag}(\nabla\xi(\vec{1})) \succ \nabla^2\xi(\vec{1})$, *both annealed and strong trivialization hold.*

- $\mathbb{E}|\mathrm{Crt}| = e^{o(N)}$
- *W.h.p. exactly* 2^r *critical points, all well-conditioned, all approx crits near one*
- *Determine energy,* r *radial derivatives,* r *correlations with external field at each*

- 2^r critical points minimal for any Morse function on product of r spheres
- Consequence: ground state energy when $\mathrm{diag}(\nabla\xi(\vec{1})) \succ \nabla^2\xi(\vec{1})$

Theorem (H.-Sellke 23)

If $\mathrm{diag}(\nabla\xi(\vec{1})) \not\succ \nabla^2\xi(\vec{1})$, *both annealed and strong trivialization fail.*

- $\mathbb{E}|\mathrm{Crt}| \geq e^{cN}$
- e^{cN} *well-separated approximate critical points w.h.p.*

Annealed Complexity for Multi-Species

$$\begin{aligned}\mathbb{E}|\mathrm{Crt}| &= (\text{simple term}) \times \mathbb{E}|\det \nabla^2_{\mathrm{sp}} H(\boldsymbol{\sigma})| \\ &= (\text{simple term}) \times \mathbb{E}|\det \left(\nabla^2_{\tan} H(\boldsymbol{\sigma}) - \mathrm{Curv}\right)|\end{aligned}$$

Annealed Complexity for Multi-Species

$$\mathbb{E}|\mathrm{Crt}| = (\text{simple term}) \times \mathbb{E}|\det \nabla^2_{\mathrm{sp}} H(\boldsymbol{\sigma})| = (\text{simple term}) \times \mathbb{E}|\det \left(\nabla^2_{\mathrm{tan}} H(\boldsymbol{\sigma}) - \mathrm{Curv}\right)|$$

In this setting (with r blocks of size $\lambda_1 N, \ldots, \lambda_r N$)

$$\nabla^2_{\mathrm{tan}} H(\boldsymbol{\sigma}) = \left[\begin{array}{c|c} c_{11} & c_{12} \\ \hline c_{12} & c_{22} \end{array}\right] \odot \mathrm{GOE}_N, \qquad \mathrm{Curv} = \left[\begin{array}{c|c} \partial_{\mathrm{rad},1} H(\boldsymbol{\sigma}) I & \\ \hline & \partial_{\mathrm{rad},2} H(\boldsymbol{\sigma}) I \end{array}\right],$$

and $\nabla_{\mathrm{rad}} H(\boldsymbol{\sigma}) = (\partial_{\mathrm{rad},1} H(\boldsymbol{\sigma}), \ldots, \partial_{\mathrm{rad},r} H(\boldsymbol{\sigma})) =_d \mathcal{N}(0, \Sigma/N)$

Annealed Complexity for Multi-Species

$$\mathbb{E}|\mathrm{Crt}| = (\text{simple term}) \times \mathbb{E}|\det \nabla^2_{\mathrm{sp}} H(\boldsymbol{\sigma})|$$
$$= (\text{simple term}) \times \mathbb{E}|\det \left(\nabla^2_{\mathrm{tan}} H(\boldsymbol{\sigma}) - \mathrm{Curv}\right)|$$

In this setting (with r blocks of size $\lambda_1 N, \ldots, \lambda_r N$)

$$\nabla^2_{\mathrm{tan}} H(\boldsymbol{\sigma}) = \left[\begin{array}{c|c} c_{11} & c_{12} \\ \hline c_{12} & c_{22} \end{array}\right] \odot \mathrm{GOE}_N, \qquad \mathrm{Curv} = \left[\begin{array}{c|c} \partial_{\mathrm{rad},1} H(\boldsymbol{\sigma}) I & \\ \hline & \partial_{\mathrm{rad},2} H(\boldsymbol{\sigma}) I \end{array}\right],$$

and $\nabla_{\mathrm{rad}} H(\boldsymbol{\sigma}) = (\partial_{\mathrm{rad},1} H(\boldsymbol{\sigma}), \ldots, \partial_{\mathrm{rad},r} H(\boldsymbol{\sigma})) =_d \mathcal{N}(0, \Sigma/N)$

Integrate out $\nabla_{\mathrm{rad}} H(\boldsymbol{\sigma}) = \vec{x}$:

$$\mathbb{E}|\mathrm{Crt}| = (\text{simple term}) \times \int_{\mathbb{R}^r} \mathbb{E}|\det M_N(\vec{x})| \exp\left(-\frac{N}{2}\langle \vec{x}, \Sigma^{-1}\vec{x}\rangle\right) \, \mathrm{d}\vec{x}$$

Annealed Complexity for Multi-Species

$$\begin{aligned}\mathbb{E}|\mathrm{Crt}| &= (\text{simple term}) \times \mathbb{E}|\det \nabla^2_{\mathrm{sp}} H(\boldsymbol{\sigma})| \\ &= (\text{simple term}) \times \mathbb{E}|\det \left(\nabla^2_{\tan} H(\boldsymbol{\sigma}) - \mathrm{Curv}\right)|\end{aligned}$$

In this setting (with r blocks of size $\lambda_1 N, \ldots, \lambda_r N$)

$$\nabla^2_{\tan} H(\boldsymbol{\sigma}) = \left[\begin{array}{c|c} c_{11} & c_{12} \\ \hline c_{12} & c_{22} \end{array}\right] \odot \mathrm{GOE}_N, \qquad \mathrm{Curv} = \left[\begin{array}{c|c} \partial_{\mathrm{rad},1} H(\boldsymbol{\sigma}) I & \\ \hline & \partial_{\mathrm{rad},2} H(\boldsymbol{\sigma}) I \end{array}\right],$$

and $\nabla_{\mathrm{rad}} H(\boldsymbol{\sigma}) = (\partial_{\mathrm{rad},1} H(\boldsymbol{\sigma}), \ldots, \partial_{\mathrm{rad},r} H(\boldsymbol{\sigma})) =_d \mathcal{N}(0, \Sigma/N)$

Integrate out $\nabla_{\mathrm{rad}} H(\boldsymbol{\sigma}) = \vec{x}$:

$$\mathbb{E}|\mathrm{Crt}| = (\text{simple term}) \times \int_{\mathbb{R}^r} \mathbb{E}|\det M_N(\vec{x})| \exp\left(-\frac{N}{2}\langle \vec{x}, \Sigma^{-1}\vec{x}\rangle\right) \, \mathrm{d}\vec{x}$$

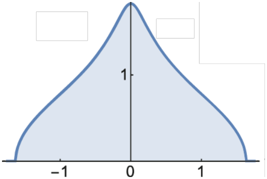
Laplace's principle $\Rightarrow$ maximize exponential rate of integrand over $\vec{x} \in \mathbb{R}^r$

Random Determinants Beyond GOE

- How to calculate $\mathbb{E}|\det M_N(\vec{x})|$ when $M_N(\vec{x})$ is not GOE?

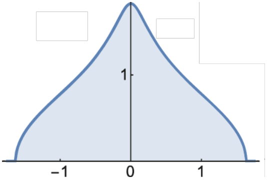
Random Determinants Beyond GOE

- How to calculate $\mathbb{E}|\det M_N(\vec{x})|$ when $M_N(\vec{x})$ is not GOE?
- Free probability $\Rightarrow$ limiting spectral distribution μ of $M_N(\vec{x})$



Random Determinants Beyond GOE

- How to calculate $\mathbb{E}|\det M_N(\vec{x})|$ when $M_N(\vec{x})$ is not GOE?
- Free probability $\Rightarrow$ limiting spectral distribution μ of $M_N(\vec{x})$

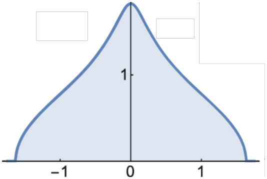


- Heuristically,

$$\frac{1}{N}\log \mathbb{E}|\det(M_N(\vec{x}))| \approx \frac{1}{N}\sum_{i=1}^{N} \log|\lambda_i| \approx \int \log|\lambda| d\mu(\lambda).$$

Random Determinants Beyond GOE

- How to calculate $\mathbb{E}|\det M_N(\vec{x})|$ when $M_N(\vec{x})$ is not GOE?
- Free probability $\Rightarrow$ limiting spectral distribution μ of $M_N(\vec{x})$



- Heuristically,

$$\frac{1}{N}\log \mathbb{E}|\det(M_N(\vec{x}))| \approx \frac{1}{N}\sum_{i=1}^{N}\log|\lambda_i| \approx \int \log|\lambda| \mathrm{d}\mu(\lambda).$$

- (Ben Arous-Bourgade-McKenna 23, McKenna 21): this is correct **to leading exponential order**

Annealed Complexity for Multi-Species: Log Determinant

Unfortunately $\mathbb{E}|\det M_N(\vec{x})|$ is still very non-explicit.

Annealed Complexity for Multi-Species: Log Determinant

Unfortunately $\mathbb{E}|\det M_N(\vec{x})|$ is still very non-explicit.

Let $m_1(z), \ldots, m_r(z) \in \mathbb{H}$ solve **vector Dyson equation**

$$x_s + z = -\frac{1}{m_s(z)} - \sum_{s'=1}^{r} c_{s,s'} m_{s'}(z), \qquad z \in \mathbb{H}$$

and $m(z) = \sum_s \lambda_s m_s(z)$.

Annealed Complexity for Multi-Species: Log Determinant

Unfortunately $\mathbb{E}|\det M_N(\vec{x})|$ is still very non-explicit.

Let $m_1(z), \ldots, m_r(z) \in \mathbb{H}$ solve **vector Dyson equation**

$$x_s + z = -\frac{1}{m_s(z)} - \sum_{s'=1}^{r} c_{s,s'} m_{s'}(z), \qquad z \in \mathbb{H}$$

and $m(z) = \sum_s \lambda_s m_s(z)$.

Limiting spectral measure $\mu_{\vec{x}}$ of $M_N(\vec{x})$ is unique measure on $\mathbb{R}$ satisfying

$$m(z) = \int_{\mathbb{R}} \frac{\mu_{\vec{x}}(\mathrm{d}\lambda)}{\lambda - z}, \qquad z \in \mathbb{H}.$$

Annealed Complexity for Multi-Species: Log Determinant

Unfortunately $\mathbb{E}|\det M_N(\vec{x})|$ is still very non-explicit.

Let $m_1(z), \dots, m_r(z) \in \mathbb{H}$ solve **vector Dyson equation**

$$x_s + z = -\frac{1}{m_s(z)} - \sum_{s'=1}^{r} c_{s,s'} m_{s'}(z), \qquad z \in \mathbb{H}$$

and $m(z) = \sum_s \lambda_s m_s(z)$.

Limiting spectral measure $\mu_{\vec{x}}$ of $M_N(\vec{x})$ is unique measure on $\mathbb{R}$ satisfying

$$m(z) = \int_{\mathbb{R}} \frac{\mu_{\vec{x}}(\mathrm{d}\lambda)}{\lambda - z}, \qquad z \in \mathbb{H}.$$

Then

$$\frac{1}{N} \log \mathbb{E}|\det M_N(\vec{x})| = \int \log |\lambda| \mu_{\vec{x}}(\mathrm{d}\lambda) + o(1).$$

Annealed Trivialization: Solving the Variational Problem

- Wish to maximize $F(\vec{x}) = C + \int \log|\lambda| \mu_{\vec{x}}(\mathrm{d}\lambda) - \frac{1}{2}\langle \vec{x}, \Sigma^{-1}\vec{x}\rangle$

Annealed Trivialization: Solving the Variational Problem

- Wish to maximize $F(\vec{x}) = C + \int \log|\lambda| \mu_{\vec{x}}(\mathrm{d}\lambda) - \frac{1}{2}\langle \vec{x}, \Sigma^{-1}\vec{x}\rangle$
- We find closed form ($\xi' = \nabla\xi(\vec{1})$, $\xi'' = \nabla^2\xi(\vec{1})$)

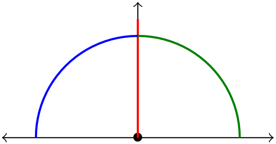
$$\int \log|\lambda| \mu_{\vec{x}}(\mathrm{d}\lambda) = \frac{1}{2}\mathrm{Re}(\langle \vec{m}(0), \xi'' \vec{m}(0)\rangle) - \sum_{s\in[r]} \lambda_s \log|m_s(0)|.$$

Annealed Trivialization: Solving the Variational Problem

- Wish to maximize $F(\vec{x}) = C + \int \log|\lambda| \mu_{\vec{x}}(\mathrm{d}\lambda) - \frac{1}{2}\langle \vec{x}, \Sigma^{-1}\vec{x}\rangle$
- We find closed form ($\xi' = \nabla\xi(\vec{1})$, $\xi'' = \nabla^2\xi(\vec{1})$)

$$\int \log|\lambda| \mu_{\vec{x}}(\mathrm{d}\lambda) = \frac{1}{2}\mathrm{Re}(\langle \vec{m}(0), \xi'' \vec{m}(0)\rangle) - \sum_{s\in[r]} \lambda_s \log|m_s(0)|.$$

- If $\nabla F(\vec{x}) = 0$, then $\mathrm{Re}(m_s(0)) = 0$ or $|m_s(0)| = \sqrt{\lambda_s/\xi'_s}$

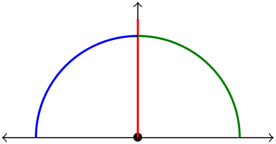


Annealed Trivialization: Solving the Variational Problem

- Wish to maximize $F(\vec{x}) = C + \int \log|\lambda| \mu_{\vec{x}}(\mathrm{d}\lambda) - \frac{1}{2}\langle \vec{x}, \Sigma^{-1}\vec{x}\rangle$
- We find closed form ($\xi' = \nabla\xi(\vec{1})$, $\xi'' = \nabla^2\xi(\vec{1})$)

$$\int \log|\lambda| \mu_{\vec{x}}(\mathrm{d}\lambda) = \frac{1}{2}\mathrm{Re}(\langle \vec{m}(0), \xi'' \vec{m}(0)\rangle) - \sum_{s\in[r]} \lambda_s \log|m_s(0)|.$$

- If $\nabla F(\vec{x}) = 0$, then $\mathrm{Re}(m_s(0)) = 0$ or $|m_s(0)| = \sqrt{\lambda_s/\xi'_s}$



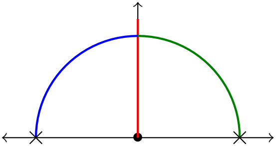
- About 3^r solutions, one for each choice of $\Re(m_s(0)) = 0$, < 0, > 0

Annealed Trivialization: Solving the Variational Problem

- Wish to maximize $F(\vec{x}) = C + \int \log|\lambda| \mu_{\vec{x}}(d\lambda) - \frac{1}{2}\langle \vec{x}, \Sigma^{-1}\vec{x}\rangle$
- We find closed form ($\xi' = \nabla\xi(\vec{1})$, $\xi'' = \nabla^2\xi(\vec{1})$)

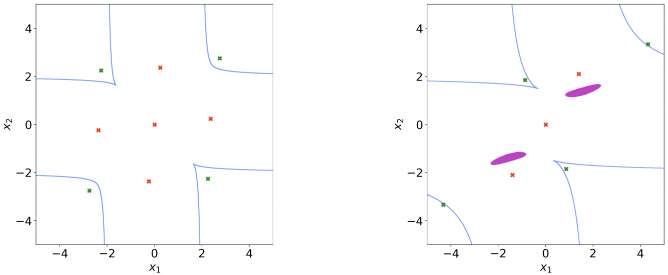
$$\int \log|\lambda| \mu_{\vec{x}}(d\lambda) = \frac{1}{2}\mathrm{Re}(\langle \vec{m}(0), \xi'' \vec{m}(0)\rangle) - \sum_{s\in[r]} \lambda_s \log|m_s(0)|.$$

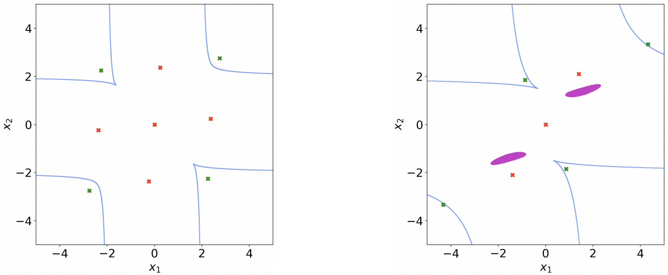
- If $\nabla F(\vec{x}) = 0$, then $\mathrm{Re}(m_s(0)) = 0$ or $|m_s(0)| = \sqrt{\lambda_s/\xi'_s}$



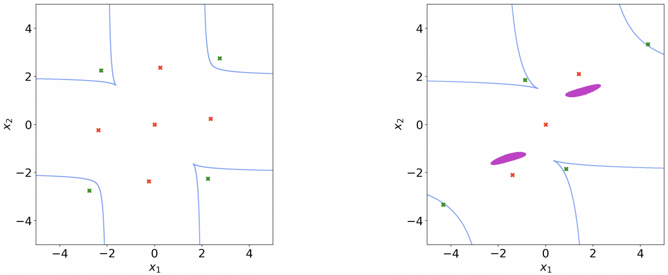
- About 3^r solutions, one for each choice of $\Re(m_s(0)) = 0, < 0, > 0$
- Of these 2^r satisfy $F(\vec{x}) = 0$, where $m_s(0) = \pm\sqrt{\lambda_s/\xi'_s}$

Annealed Trivialization: Solving the Variational Problem



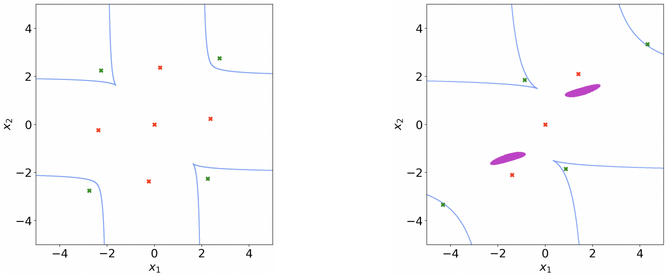


- Need to show the remaining stationary points of F are **not** maxima



- Need to show the remaining stationary points of F are **not** maxima
- At each, construct $\vec{w}$ where $\langle \vec{w}, \nabla^2 F(\vec{x})\vec{w} \rangle > 0$

Annealed Trivialization: Solving the Variational Problem



- Need to show the remaining stationary points of F are **not** maxima
- At each, construct $\vec{w}$ where $\langle \vec{w}, \nabla^2 F(\vec{x})\vec{w}\rangle > 0$
- Thus $\max F = 0 \Rightarrow \mathbb{E}|\mathrm{Crt}| = e^{o(N)}$

From Annealed to Strong Trivialization

- Annealed trivialization $\Rightarrow \mathbb{E}|\mathrm{Crt}| = e^{o(N)}$, and for each $\boldsymbol{\sigma} \in \mathrm{Crt}$:

From Annealed to Strong Trivialization

- Annealed trivialization $\Rightarrow \mathbb{E}|\mathrm{Crt}| = e^{o(N)}$, and for each $\boldsymbol{\sigma} \in \mathrm{Crt}$:
  - $\left(\nabla_{\mathrm{rad}} H(\boldsymbol{\sigma}), \frac{1}{N} H(\boldsymbol{\sigma}), (\frac{1}{N}\langle \boldsymbol{G}_s^{(1)}, \boldsymbol{\sigma}\rangle)_{s\in\mathscr{S}}\right)$ one of 2^r possible values, up to $o(1)$

From Annealed to Strong Trivialization

- Annealed trivialization $\Rightarrow \mathbb{E}|\mathrm{Crt}| = e^{o(N)}$, and for each $\boldsymbol{\sigma} \in \mathrm{Crt}$:
  - $\left(\nabla_{\mathrm{rad}} H(\boldsymbol{\sigma}), \frac{1}{N} H(\boldsymbol{\sigma}), (\frac{1}{N}\langle \boldsymbol{G}_s^{(1)}, \boldsymbol{\sigma}\rangle)_{s\in\mathscr{S}}\right)$ one of 2^r possible values, up to $o(1)$
  - $\nabla_{\mathrm{sp}}^2 H(\boldsymbol{\sigma})$ well-conditioned

From Annealed to Strong Trivialization

- Annealed trivialization $\Rightarrow \mathbb{E}|\mathrm{Crt}| = e^{o(N)}$, and for each $\boldsymbol{\sigma} \in \mathrm{Crt}$:
  - $\left(\nabla_{\mathrm{rad}} H(\boldsymbol{\sigma}), \frac{1}{N} H(\boldsymbol{\sigma}), (\frac{1}{N}\langle \boldsymbol{G}_s^{(1)}, \boldsymbol{\sigma}\rangle)_{s\in\mathscr{S}}\right)$ one of 2^r possible values, up to $o(1)$
  - $\nabla^2_{\mathrm{sp}} H(\boldsymbol{\sigma})$ well-conditioned
- Approximate to exact lemma $\Rightarrow$ all approx crits satisfy this too

From Annealed to Strong Trivialization

- Annealed trivialization $\Rightarrow \mathbb{E}|\mathrm{Crt}| = e^{o(N)}$, and for each $\boldsymbol{\sigma} \in \mathrm{Crt}$:
  - $\left(\nabla_{\mathrm{rad}} H(\boldsymbol{\sigma}), \frac{1}{N} H(\boldsymbol{\sigma}), (\frac{1}{N}\langle \boldsymbol{G}_s^{(1)}, \boldsymbol{\sigma}\rangle)_{s\in\mathscr{S}}\right)$ one of 2^r possible values, up to $o(1)$
  - $\nabla_{\mathrm{sp}}^2 H(\boldsymbol{\sigma})$ well-conditioned
- Approximate to exact lemma $\Rightarrow$ all approx crits satisfy this too
- Band recursion $\Rightarrow$ localize all approx crits to 2^r small regions

From Annealed to Strong Trivialization

- Annealed trivialization $\Rightarrow \mathbb{E}|\mathrm{Crt}| = e^{o(N)}$, and for each $\boldsymbol{\sigma} \in \mathrm{Crt}$:
  - $\left(\nabla_{\mathrm{rad}} H(\boldsymbol{\sigma}), \frac{1}{N} H(\boldsymbol{\sigma}), (\frac{1}{N}\langle \boldsymbol{G}_s^{(1)}, \boldsymbol{\sigma}\rangle)_{s \in \mathscr{S}}\right)$ one of 2^r possible values, up to $o(1)$
  - $\nabla_{\mathrm{sp}}^2 H(\boldsymbol{\sigma})$ well-conditioned
- Approximate to exact lemma $\Rightarrow$ all approx crits satisfy this too

- Band recursion $\Rightarrow$ localize all approx crits to 2^r small regions
- $\nabla_{\mathrm{sp}}^2 H(\boldsymbol{\sigma})$ well-conditioned $\Rightarrow$ each region has ≤ 1 exact crit

From Annealed to Strong Trivialization

- Annealed trivialization $\Rightarrow \mathbb{E}|\mathrm{Crt}| = e^{o(N)}$, and for each $\boldsymbol{\sigma} \in \mathrm{Crt}$:
  - $\left(\nabla_{\mathrm{rad}} H(\boldsymbol{\sigma}), \frac{1}{N} H(\boldsymbol{\sigma}), (\frac{1}{N}\langle \boldsymbol{G}_s^{(1)}, \boldsymbol{\sigma}\rangle)_{s\in\mathscr{S}}\right)$ one of 2^r possible values, up to $o(1)$
  - $\nabla_{\mathrm{sp}}^2 H(\boldsymbol{\sigma})$ well-conditioned
- Approximate to exact lemma $\Rightarrow$ all approx crits satisfy this too

- Band recursion $\Rightarrow$ localize all approx crits to 2^r small regions
- $\nabla_{\mathrm{sp}}^2 H(\boldsymbol{\sigma})$ well-conditioned $\Rightarrow$ each region has ≤ 1 exact crit
- But $\geq 2^r$ exact crits by Morse theory $\Rightarrow$ each region has 1 exact crit, all approx crits near one

From Annealed to Strong Trivialization

- Annealed trivialization $\Rightarrow \mathbb{E}|\mathrm{Crt}| = e^{o(N)}$, and for each $\boldsymbol{\sigma} \in \mathrm{Crt}$:
  - $\left(\nabla_{\mathrm{rad}} H(\boldsymbol{\sigma}), \frac{1}{N} H(\boldsymbol{\sigma}), (\frac{1}{N}\langle \boldsymbol{G}_s^{(1)}, \boldsymbol{\sigma}\rangle)_{s\in\mathscr{S}}\right)$ one of 2^r possible values, up to $o(1)$
  - $\nabla_{\mathrm{sp}}^2 H(\boldsymbol{\sigma})$ well-conditioned
- Approximate to exact lemma $\Rightarrow$ all approx crits satisfy this too

- Band recursion $\Rightarrow$ localize all approx crits to 2^r small regions
- $\nabla_{\mathrm{sp}}^2 H(\boldsymbol{\sigma})$ well-conditioned $\Rightarrow$ each region has ≤ 1 exact crit
- But $\geq 2^r$ exact crits by Morse theory $\Rightarrow$ each region has 1 exact crit, all approx crits near one

- Converse: algorithmically construct e^{cN} approx crits with AMP

Summary

- We determine phase boundary for annealed and strong topological trivialization in multi-species spherical spin glasses

Summary

- We determine phase boundary for annealed and strong topological trivialization in multi-species spherical spin glasses
- Approximate critical points controlled by approximate-to-exact lemma

Summary

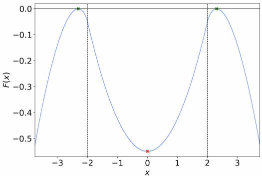
- We determine phase boundary for annealed and strong topological trivialization in multi-species spherical spin glasses

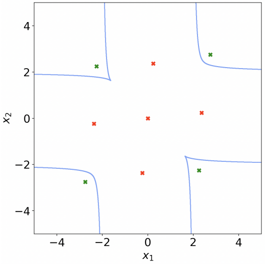
- Approximate critical points controlled by approximate-to-exact lemma
- Band recursion $\Rightarrow$ strong trivialization, given only $e^{o(N)}$-resolution annealed inputs

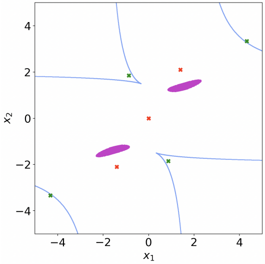
Summary

- We determine phase boundary for annealed and strong topological trivialization in multi-species spherical spin glasses

- Approximate critical points controlled by approximate-to-exact lemma
- Band recursion $\Rightarrow$ strong trivialization, given only $e^{o(N)}$-resolution annealed inputs







Thank you!