Capacity threshold for the Ising perceptron

Brice Huang (MIT)

MIT Combinatorics Seminar

November 14, 2024

High-dimensional disordered systems

Models of random, high-dimensional **objective functions** or **distributions**

High-dimensional disordered systems

Models of random, high-dimensional **objective functions** or **distributions**

- Hardcore model: independent sets $I \subseteq G$ and $\mu(I) \propto \lambda^{|I|}$
- Random k-SAT and $\mu = \text{unif}(\text{satisfying assignments})$

High-dimensional disordered systems

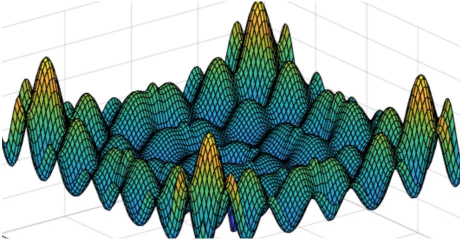
Models of random, high-dimensional **objective functions** or **distributions**

- Hardcore model: independent sets $I \subseteq G$ and $\mu(I) \propto \lambda^{|I|}$
- Random k-SAT and $\mu = \text{unif}(\text{satisfying assignments})$
- Ising model *Hamiltonian* $H : \{\pm 1\}^N \to \mathbb{R}$ and *Gibbs measure* $\mu(\boldsymbol{x}) \propto e^{H(\boldsymbol{x})}$

High-dimensional disordered systems

Models of random, high-dimensional **objective functions** or **distributions**

- Hardcore model: independent sets $I \subseteq G$ and $\mu(I) \propto \lambda^{|I|}$
- Random k-SAT and $\mu = \text{unif}(\text{satisfying assignments})$
- Ising model *Hamiltonian* $H : \{\pm 1\}^N \to \mathbb{R}$ and *Gibbs measure* $\mu(\boldsymbol{x}) \propto e^{H(\boldsymbol{x})}$



Highly non-convex landscapes, often with exponentially many maxima

High-dimensional disordered systems

Applications:

- Statistical inference, e.g. community detection ($\mu =$ posterior distribution)

High-dimensional disordered systems

Applications:

- Statistical inference, e.g. community detection ($\mu =$ posterior distribution)
- Random optimization problems, e.g. training a neural network ($H =$ loss function)

High-dimensional disordered systems

Applications:

- Statistical inference, e.g. community detection ($\mu =$ posterior distribution)
- Random optimization problems, e.g. training a neural network ($H =$ loss function)
- Complex systems & emergent behavior: magnetic materials, protein folding, flocking birds, economics, ...

High-dimensional disordered systems

Applications:

- Statistical inference, e.g. community detection ($\mu =$ posterior distribution)
- Random optimization problems, e.g. training a neural network ($H =$ loss function)
- Complex systems & emergent behavior: magnetic materials, protein folding, flocking birds, economics, . . .

What are the values of **key statistics** of a disordered system, such as the **maximum value**, **number of solutions**, or **satisfiability threshold**?

High-dimensional disordered systems

Applications:

- Statistical inference, e.g. community detection ($\mu =$ posterior distribution)
- Random optimization problems, e.g. training a neural network ($H =$ loss function)
- Complex systems & emergent behavior: magnetic materials, protein folding, flocking birds, economics, ...

What are the values of **key statistics** of a disordered system, such as the **maximum value**, **number of solutions**, or **satisfiability threshold**?

Unified framework of physics **predictions** via replica / cavity methods

High-dimensional disordered systems

Applications:

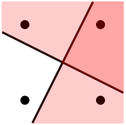
- Statistical inference, e.g. community detection ($\mu =$ posterior distribution)
- Random optimization problems, e.g. training a neural network ($H =$ loss function)
- Complex systems & emergent behavior: magnetic materials, protein folding, flocking birds, economics, ...

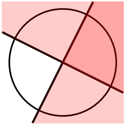
What are the values of **key statistics** of a disordered system, such as the **maximum value**, **number of solutions**, or **satisfiability threshold**?

Unified framework of physics **predictions** via replica / cavity methods

Goal: unified rigorous theory & broadly applicable tools

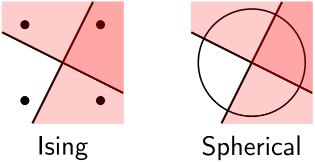
The perceptron model


Ising


Spherical

Intersection of $\Sigma_N = \{-1,1\}^N$ or $\sqrt{N}\mathbb{S}^{N-1}$ with M i.i.d. random half-spaces

The perceptron model

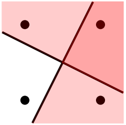


Intersection of $\Sigma_N = \{-1,1\}^N$ or $\sqrt{N}\mathbb{S}^{N-1}$
with M i.i.d. random half-spaces

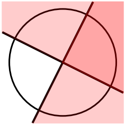
Formally: $\kappa \in \mathbb{R}$ fixed, $\boldsymbol{g}^1, \boldsymbol{g}^2, \ldots \sim \mathcal{N}(0, I_N)$,

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

The perceptron model



Ising



Spherical

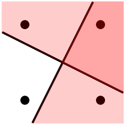
Intersection of $\Sigma_N = \{-1, 1\}^N$ or $\sqrt{N}\mathbb{S}^{N-1}$ with M i.i.d. random half-spaces

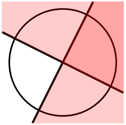
Formally: $\kappa \in \mathbb{R}$ fixed, $\boldsymbol{g}^1, \boldsymbol{g}^2, \ldots \sim \mathcal{N}(0, I_N)$,

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

Capacity: largest M such that $S(M) \neq \emptyset$

The perceptron model





Intersection of $\Sigma_N = \{-1,1\}^N$ or $\sqrt{N}\mathbb{S}^{N-1}$ with M i.i.d. random half-spaces

Formally: $\kappa \in \mathbb{R}$ fixed, $\boldsymbol{g}^1, \boldsymbol{g}^2, \ldots \sim \mathcal{N}(0, I_N)$,

$$S(M) = \left\{ \boldsymbol{x} \in \Sigma_N : \frac{\langle \boldsymbol{g}^a, \boldsymbol{x} \rangle}{\sqrt{N}} \ge \kappa, \quad \forall 1 \le a \le M \right\}$$

Capacity: largest M such that $S(M) \neq \emptyset$

Main question: what is $\alpha_* = \alpha_*(\kappa) = \text{p-lim}_{N\to\infty} M/N$?

Connections to other problems

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

Capacity: largest M such that $S(M) \neq \emptyset$

Satisfiability threshold of constraint satisfaction problem with **global** constraints

Connections to other problems

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

Capacity: largest M such that $S(M) \neq \emptyset$

Satisfiability threshold of constraint satisfaction problem with **global** constraints

Toy model of single-layer **neural network** (Gardner 88):

- $\boldsymbol{x} \leftrightarrow$ synaptic weights. $\boldsymbol{g}^1, \ldots, \boldsymbol{g}^M \leftrightarrow$ random patterns

Connections to other problems

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

Capacity: largest M such that $S(M) \neq \emptyset$

Satisfiability threshold of constraint satisfaction problem with **global** constraints

Toy model of single-layer **neural network** (Gardner 88):

- $\boldsymbol{x} \leftrightarrow$ synaptic weights. $\boldsymbol{g}^1, \ldots, \boldsymbol{g}^M \leftrightarrow$ random patterns
- $S(M) \leftrightarrow$ synaptic weights memorizing all M patterns

Connections to other problems

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

Capacity: largest M such that $S(M) \neq \emptyset$

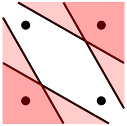
Satisfiability threshold of constraint satisfaction problem with **global** constraints

Toy model of single-layer **neural network** (Gardner 88):

- $\boldsymbol{x} \leftrightarrow$ synaptic weights. $\boldsymbol{g}^1, \ldots, \boldsymbol{g}^M \leftrightarrow$ random patterns
- $S(M) \leftrightarrow$ synaptic weights memorizing all M patterns
- Capacity $\leftrightarrow$ max # patterns neural network can memorize

Connections to other problems

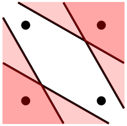
Related model: **symmetric perceptron** with constraints $\langle \boldsymbol{g}^a, \boldsymbol{x} \rangle / \sqrt{N} \in [-\kappa, \kappa]$



(Aubin Perkins Zdeborová 19, Perkins Xu 21, Abbe Li Sly 22, ...)

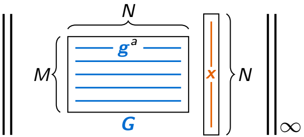
Connections to other problems

Related model: **symmetric perceptron** with constraints $\langle \boldsymbol{g}^a, \boldsymbol{x} \rangle / \sqrt{N} \in [-\kappa, \kappa]$



(Aubin Perkins Zdeborová 19, Perkins Xu 21, Abbe Li Sly 22, . . .)

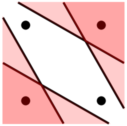
$\leftrightarrow$ **discrepancy minimization**: given $\boldsymbol{G} \in \mathbb{R}^{M \times N}$, find $\boldsymbol{x} \in \{\pm 1\}^N$ minimizing

$\|\boldsymbol{G}\boldsymbol{x}\|_\infty =$ 

(Spencer 85, Bansal 10, Lovett Meka 15, Rothvoss 17, Eldan Singh 18, . . .)

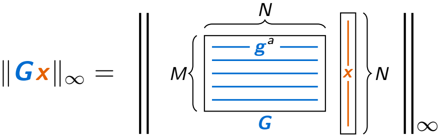
Connections to other problems

Related model: **symmetric perceptron** with constraints $\langle \boldsymbol{g}^a, \boldsymbol{x} \rangle / \sqrt{N} \in [-\kappa, \kappa]$



(Aubin Perkins Zdeborová 19, Perkins Xu 21, Abbe Li Sly 22, . . .)

$\leftrightarrow$ **discrepancy minimization**: given $\boldsymbol{G} \in \mathbb{R}^{M \times N}$, find $\boldsymbol{x} \in \{\pm 1\}^N$ minimizing

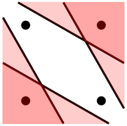


(Spencer 85, Bansal 10, Lovett Meka 15, Rothvoss 17, Eldan Singh 18, . . .)

Applications: randomized control trials, sparsification, differential privacy, . . .

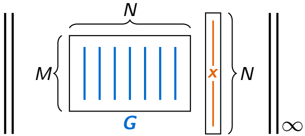
Connections to other problems

Related model: **symmetric perceptron** with constraints $\langle \boldsymbol{g}^a, \boldsymbol{x}\rangle/\sqrt{N} \in [-\kappa, \kappa]$



(Aubin Perkins Zdeborová 19, Perkins Xu 21, Abbe Li Sly 22, . . .)

$\leftrightarrow$ **discrepancy minimization**: given $\boldsymbol{G} \in \mathbb{R}^{M\times N}$, find $\boldsymbol{x} \in \{\pm 1\}^N$ minimizing

$\|\boldsymbol{G}\boldsymbol{x}\|_\infty =$ 

(Spencer 85, Bansal 10, Lovett Meka 15, Rothvoss 17, Eldan Singh 18, . . .)

Applications: randomized control trials, sparsification, differential privacy, . . .

Main result

Conjecture (Krauth Mézard 1989)

For the $\kappa = 0$ Ising perceptron, $\alpha_* = \alpha_{\mathsf{KM}} \approx 0.833$.

Theorem (Ding Sun 2018)

$\alpha_* \geq \alpha_{\mathsf{KM}}$, under condition that an explicit univariate function is ≤ 0.

Main result

Conjecture (Krauth Mézard 1989)

For the $\kappa = 0$ Ising perceptron, $\alpha_* = \alpha_{\mathsf{KM}} \approx 0.833$.

Theorem (Ding Sun 2018)

$\alpha_* \geq \alpha_{\mathsf{KM}}$, under condition that an explicit univariate function is ≤ 0.

Theorem (H. 2024)

$\alpha_* \leq \alpha_{\mathsf{KM}}$, under condition that an explicit 2-variable function (next slide) is ≤ 0.

Main result

Conjecture (Krauth Mézard 1989)

For the $\kappa = 0$ Ising perceptron, $\alpha_* = \alpha_{\mathsf{KM}} \approx 0.833$.

Theorem (Ding Sun 2018)

$\alpha_* \geq \alpha_{\mathsf{KM}}$, under condition that an explicit univariate function is ≤ 0.

Theorem (H. 2024)

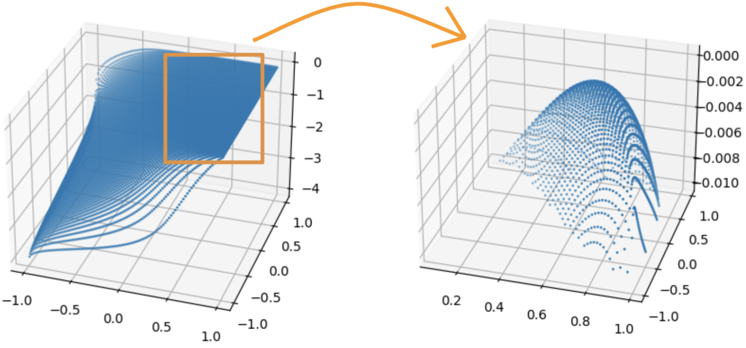
$\alpha_* \leq \alpha_{\mathsf{KM}}$, under condition that an explicit 2-variable function (next slide) is ≤ 0.

(Also for general $\kappa \in \mathbb{R}$, under several numerical conditions depending on κ)

The function $\mathscr{S}_*$ in our numerical condition

$\mathscr{S}_*(1, 0) = 0$ local max, conjecturally unique global max

Plot of $(x, y) \mapsto \mathscr{S}_*(\tanh^{-1}(x), \tanh^{-1}(y))$:



$x, y \in [-1, 1]$ $\qquad$ $\mathscr{S}_* \geq -0.01$

Background: physics predictions

Gardner Derrida 88, Krauth Mézard 89:

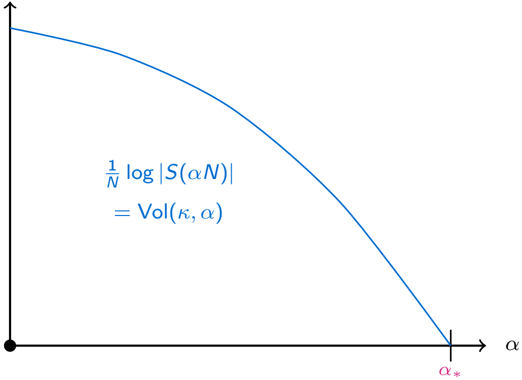
- Volume formula $\frac{1}{N} \log |S(\alpha N)| \to_p \mathrm{Vol}(\kappa, \alpha)$ in terms of fixed point eqn
  (In "replica symmetric" regime of (κ, α), which includes $\kappa = 0$ Ising perceptron)

Background: physics predictions

Gardner Derrida 88, Krauth Mézard 89:

- Volume formula $\frac{1}{N} \log |S(\alpha N)| \to_p \mathrm{Vol}(\kappa, \alpha)$ in terms of fixed point eqn
  (In "replica symmetric" regime of (κ, α), which includes $\kappa = 0$ Ising perceptron)

- Capacity: $\alpha_* = \alpha_*(\kappa)$ solves $\mathrm{Vol}(\kappa, \alpha_*) = 0$



Spherical perceptron, $\kappa \geq 0$

- Shcherbina Tirozzi 03: proof of volume limit $\mathrm{Vol}(\kappa, \alpha)$ (and thus capacity)
- Stojnic 13: simple proof of capacity threshold

Crucial to proofs: $\kappa \geq 0$ spherical perceptron is **convex** problem!

Ising perceptron, $\kappa = 0$

- Talagrand 11, Xu 21, Nakajima Sun 23: sharp threshold **sequence** $\alpha_*(N)$ (non-explicit, doesn't imply $\alpha_* = \lim_{N\to\infty} \alpha_*(N)$ exists)

Ising perceptron, $\kappa = 0$

- Talagrand 11, Xu 21, Nakajima Sun 23: sharp threshold **sequence** $\alpha_*(N)$ (non-explicit, doesn't imply $\alpha_* = \lim_{N\to\infty} \alpha_*(N)$ exists)
- Trivial bound: $\mathbb{E}|S(N\alpha)| \le 2^{(1-\alpha)N}$, so $\alpha_* \le 1$

Ising perceptron, $\kappa = 0$

- Talagrand 11, Xu 21, Nakajima Sun 23: sharp threshold **sequence** $\alpha_*(N)$ (non-explicit, doesn't imply $\alpha_* = \lim_{N\to\infty} \alpha_*(N)$ exists)
- Trivial bound: $\mathbb{E}|S(N\alpha)| \le 2^{(1-\alpha)N}$, so $\alpha_* \le 1$
- Kim Roche 98, Talagrand 99 & 00: $\varepsilon \le \alpha_* \le 1 - \varepsilon$

Ising perceptron, $\kappa = 0$

- Talagrand 11, Xu 21, Nakajima Sun 23: sharp threshold **sequence** $\alpha_*(N)$ (non-explicit, doesn't imply $\alpha_* = \lim_{N\to\infty} \alpha_*(N)$ exists)
- Trivial bound: $\mathbb{E}|S(N\alpha)| \le 2^{(1-\alpha)N}$, so $\alpha_* \le 1$
- Kim Roche 98, Talagrand 99 & 00: $\varepsilon \le \alpha_* \le 1 - \varepsilon$
- Ding Sun 18: $\alpha_* \ge \alpha_{\mathsf{KM}} \approx 0.833$

Ising perceptron, $\kappa = 0$

- Talagrand 11, Xu 21, Nakajima Sun 23: sharp threshold **sequence** $\alpha_*(N)$ (non-explicit, doesn't imply $\alpha_* = \lim_{N\to\infty} \alpha_*(N)$ exists)
- Trivial bound: $\mathbb{E}|S(N\alpha)| \le 2^{(1-\alpha)N}$, so $\alpha_* \le 1$
- Kim Roche 98, Talagrand 99 & 00: $\varepsilon \le \alpha_* \le 1 - \varepsilon$
- Ding Sun 18: $\alpha_* \ge \alpha_{\mathrm{KM}} \approx 0.833$
- Altschuler Tikhomirov 24: $\alpha_* \le 0.847$

Ising perceptron, $\kappa = 0$

- Talagrand 11, Xu 21, Nakajima Sun 23: sharp threshold **sequence** $\alpha_*(N)$ (non-explicit, doesn't imply $\alpha_* = \lim_{N\to\infty} \alpha_*(N)$ exists)
- Trivial bound: $\mathbb{E}|S(N\alpha)| \le 2^{(1-\alpha)N}$, so $\alpha_* \le 1$
- Kim Roche 98, Talagrand 99 & 00: $\varepsilon \le \alpha_* \le 1 - \varepsilon$
- Ding Sun 18: $\alpha_* \ge \alpha_{\mathsf{KM}} \approx 0.833$
- Altschuler Tikhomirov 24: $\alpha_* \le 0.847$
- H. 24: $\alpha_* \le \alpha_{\mathsf{KM}}$

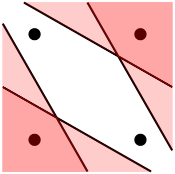
First attempt: 1st/2nd moment method

$\mathbb{E}|S(N\alpha)| \ll 1 \Rightarrow$ no solutions; $\mathbb{E}[|S(N\alpha)|^2] \asymp (\mathbb{E}|S(N\alpha)|)^2 \Rightarrow$ solutions

First attempt: 1st/2nd moment method

$\mathbb{E}|S(N\alpha)| \ll 1 \Rightarrow$ no solutions; $\mathbb{E}[|S(N\alpha)|^2] \asymp (\mathbb{E}|S(N\alpha)|)^2 \Rightarrow$ solutions

Identifies α_* in **symmetric perceptron** with constraints $\langle \boldsymbol{g}^a, \boldsymbol{x} \rangle / \sqrt{N} \in [-\kappa, \kappa]$

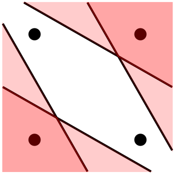


(Aubin Perkins Zdeborová 19, Perkins Xu 21, Abbe Li Sly 22, Altschuler 22, Sah Sawhney 23)

First attempt: 1st/2nd moment method

$\mathbb{E}|S(N\alpha)| \ll 1 \Rightarrow$ no solutions; $\mathbb{E}[|S(N\alpha)|^2] \asymp (\mathbb{E}|S(N\alpha)|)^2 \Rightarrow$ solutions

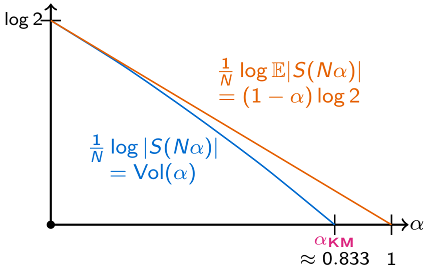
Identifies α_* in **symmetric perceptron** with constraints $\langle \boldsymbol{g}^a, \boldsymbol{x}\rangle/\sqrt{N} \in [-\kappa, \kappa]$



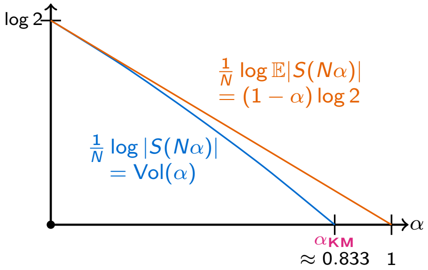
(Aubin Perkins Zdeborová 19, Perkins Xu 21, Abbe Li Sly 22, Altschuler 22, Sah Sawhney 23)

But fails in our model! $\mathbb{E}|S(N\alpha)| = 2^{(1-\alpha)N} \ll 1$ only when $\alpha > 1$

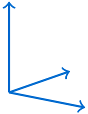
Failure of direct moment method



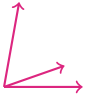
Failure of direct moment method



$\mathbb{E}|S(N\alpha)|$ dominated by events where the $\boldsymbol{g}^a$ are **atypically correlated**



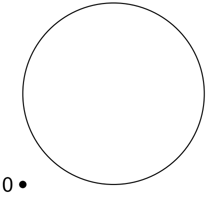
Typically: $\boldsymbol{g}^a$ orthogonal



Atypically: $\boldsymbol{g}^a$ correlated, which inflates # solutions

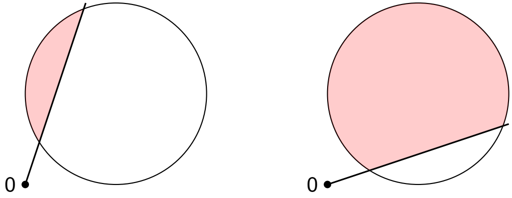
Another perspective on 1st moment failure

Solution set is not centered on origin:



Another perspective on 1st moment failure

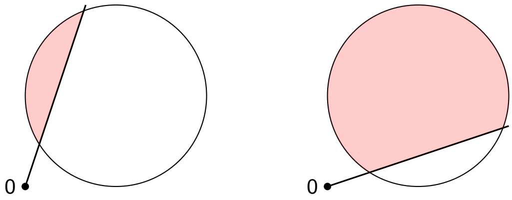
Solution set is not centered on origin:



Each slice deletes half of solutions **on average**, but **genuine fluctuations**

Another perspective on 1st moment failure

Solution set is not centered on origin:

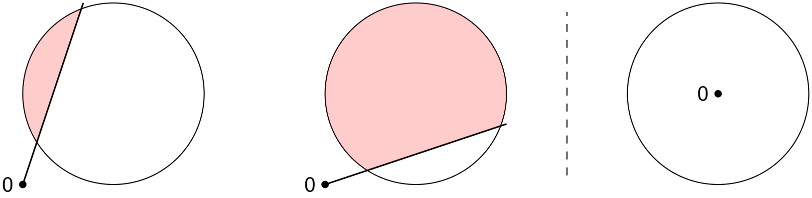


Each slice deletes half of solutions **on average**, but **genuine fluctuations**

Imagine simplified model: each slice deletes 1% or 99%. Then all solutions deleted much faster than 1st moment bound.

Another perspective on 1st moment failure

Solution set is not centered on origin:

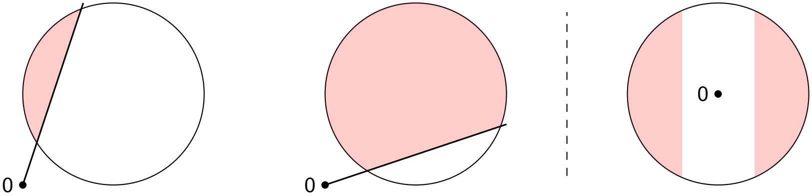


Each slice deletes half of solutions **on average**, but **genuine fluctuations**

Imagine simplified model: each slice deletes 1% or 99%. Then all solutions deleted much faster than 1st moment bound.

Another perspective on 1st moment failure

Solution set is not centered on origin:

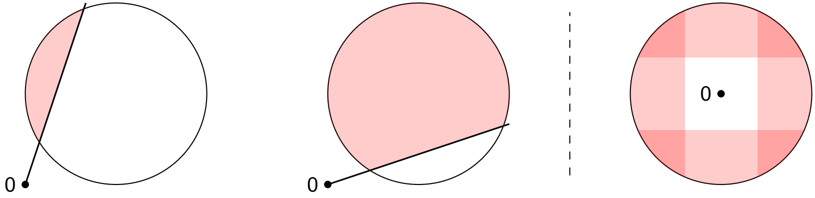


Each slice deletes half of solutions **on average**, but **genuine fluctuations**

Imagine simplified model: each slice deletes 1% or 99%. Then all solutions deleted much faster than 1st moment bound.

Another perspective on 1st moment failure

Solution set is not centered on origin:

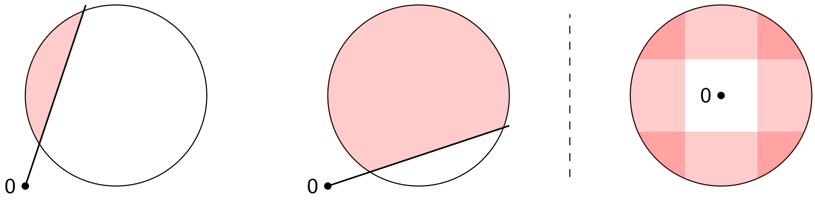


Each slice deletes half of solutions **on average**, but **genuine fluctuations**

Imagine simplified model: each slice deletes 1% or 99%. Then all solutions deleted much faster than 1st moment bound.

Another perspective on 1st moment failure

Solution set is not centered on origin:



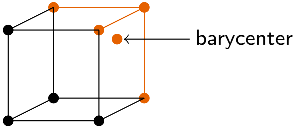
Each slice deletes half of solutions **on average**, but **genuine fluctuations**

Imagine simplified model: each slice deletes 1% or 99%. Then all solutions deleted much faster than 1st moment bound.

Next few slides: non-rigorous physics intuitions on how to remedy this.

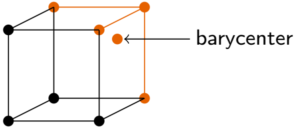
Conditional moment method

Key intuition of Ding Sun 18, Bolthausen 19: 1st mt failure caused by large deviation events in **barycenter** of solution set $S(N\alpha)$



Conditional moment method

Key intuition of Ding Sun 18, Bolthausen 19: 1st mt failure caused by large deviation events in **barycenter** of solution set $S(N\alpha)$

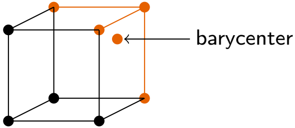


That is, $\mathbb{E}(|S|) \gg$ (typical $|S|$) but we expect

$$(\text{typical } |S|) \asymp \mathbb{E}(|S| \,|\, \text{barycenter})$$

Conditional moment method

Key intuition of Ding Sun 18, Bolthausen 19: 1st mt failure caused by large deviation events in **barycenter** of solution set $S(N\alpha)$

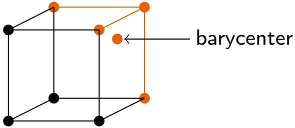


That is, $\mathbb{E}(|S|) \gg (\text{typical } |S|)$ but we expect

$$(\text{typical } |S|) \asymp \mathbb{E}(|S| \,|\, \text{barycenter}) \asymp \mathbb{E}(|S|^2 \,|\, \text{barycenter})^{1/2}$$

Conditional moment method

Key intuition of Ding Sun 18, Bolthausen 19: 1st mt failure caused by large deviation events in **barycenter** of solution set $S(N\alpha)$



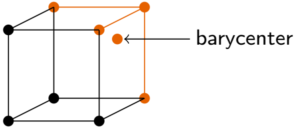
That is, $\mathbb{E}(|S|) \gg (\text{typical } |S|)$ but we expect

$$(\text{typical } |S|) \asymp \mathbb{E}(|S| \,|\, \text{barycenter}) \asymp \mathbb{E}(|S|^2 \,|\, \text{barycenter})^{1/2}$$

This suggests plan: condition on **typical behavior** of barycenter, then 1st/2nd mt

Conditional moment method

Key intuition of Ding Sun 18, Bolthausen 19: 1st mt failure caused by large deviation events in **barycenter** of solution set $S(N\alpha)$



That is, $\mathbb{E}(|S|) \gg (\text{typical } |S|)$ but we expect

$$(\text{typical } |S|) \asymp \mathbb{E}(|S| \,|\, \text{barycenter}) \asymp \mathbb{E}(|S|^2 \,|\, \text{barycenter})^{1/2}$$

This suggests plan: condition on **typical behavior** of barycenter, then 1st/2nd mt

Barycenter is mathematically complicated, but can condition on **physics proxy**

(Heuristic) physics description of barycenter

TAP equation: nonlinear system in

- $\boldsymbol{G} \in \mathbb{R}^{M \times N}$ matrix with rows $\boldsymbol{g}^1, \ldots, \boldsymbol{g}^M$
- $\boldsymbol{m} \in \mathbb{R}^N$ barycenter of $S(M)$
- $\boldsymbol{n} \in \mathbb{R}^M$ average slacks of constraints: $n_a = \text{avg}_{\boldsymbol{x} \in S(M)} \left\{ \frac{\langle \boldsymbol{g}^a, \boldsymbol{x} \rangle}{\sqrt{N}} - \kappa \right\}$

(Heuristic) physics description of barycenter

TAP equation: nonlinear system in

- $\boldsymbol{G} \in \mathbb{R}^{M \times N}$ matrix with rows $\boldsymbol{g}^1, \dots, \boldsymbol{g}^M$
- $\boldsymbol{m} \in \mathbb{R}^N$ barycenter of $S(M)$
- $\boldsymbol{n} \in \mathbb{R}^M$ average slacks of constraints: $n_a = \text{avg}_{\boldsymbol{x} \in S(M)} \left\{ \frac{\langle \boldsymbol{g}^a, \boldsymbol{x} \rangle}{\sqrt{N}} - \kappa \right\}$

For explicit nonlinearities $\dot{F}, \widehat{F}$, constants b, d:

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

(Heuristic) physics description of barycenter

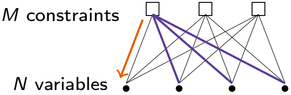
TAP equation: nonlinear system in

- $\boldsymbol{G} \in \mathbb{R}^{M \times N}$ matrix with rows $\boldsymbol{g}^1, \ldots, \boldsymbol{g}^M$
- $\boldsymbol{m} \in \mathbb{R}^N$ barycenter of $S(M)$
- $\boldsymbol{n} \in \mathbb{R}^M$ average slacks of constraints: $n_a = \text{avg}_{\boldsymbol{x} \in S(M)} \left\{ \frac{\langle \boldsymbol{g}^a, \boldsymbol{x} \rangle}{\sqrt{N}} - \kappa \right\}$

For explicit nonlinearities $\dot{F}, \widehat{F}$, constants b, d:

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

($\leftrightarrow$ dense graph limit of **belief propagation**)



$$\boldsymbol{n}_{a \to i} = \hat{f}((\boldsymbol{m}_{j \to a})_{j \neq i})$$
$$\boldsymbol{m}_{i \to a} = \dot{f}((\boldsymbol{n}_{b \to i})_{b \neq a})$$

(Heuristic) physics description of barycenter

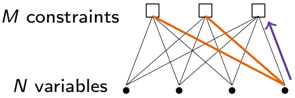
TAP equation: nonlinear system in

- $\boldsymbol{G} \in \mathbb{R}^{M \times N}$ matrix with rows $\boldsymbol{g}^1, \dots, \boldsymbol{g}^M$
- $\boldsymbol{m} \in \mathbb{R}^N$ barycenter of $S(M)$
- $\boldsymbol{n} \in \mathbb{R}^M$ average slacks of constraints: $n_a = \text{avg}_{\boldsymbol{x} \in S(M)} \left\{ \frac{\langle \boldsymbol{g}^a, \boldsymbol{x} \rangle}{\sqrt{N}} - \kappa \right\}$

For explicit nonlinearities $\dot{F}, \widehat{F}$, constants b, d:

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

($\leftrightarrow$ dense graph limit of **belief propagation**)



$$\boldsymbol{n}_{a \to i} = \hat{f}((\boldsymbol{m}_{j \to a})_{j \neq i})$$
$$\boldsymbol{m}_{i \to a} = \dot{f}((\boldsymbol{n}_{b \to i})_{b \neq a})$$

Physics predictions for TAP & planted model

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

Physics prediction: for typical $\boldsymbol{G}$, unique soln $(\boldsymbol{m}, \boldsymbol{n})$. $\boldsymbol{m} \approx$ barycenter, $\boldsymbol{n} \approx$ slacks

---

Physics predictions for TAP & planted model

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

Physics prediction: for typical $\boldsymbol{G}$, unique soln $(\boldsymbol{m}, \boldsymbol{n})$. $\boldsymbol{m} \approx$ barycenter, $\boldsymbol{n} \approx$ slacks

---

Key idea: **planted model**:

1. Sample $(\boldsymbol{m}, \boldsymbol{n})$ from its law (explicit physics prediction)
2. Sample $\boldsymbol{G}$ conditioned on TAP$(\boldsymbol{G}, \boldsymbol{m}, \boldsymbol{n})$

Belief: planted $\approx$ true model;

Physics predictions for TAP & planted model

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

Physics prediction: for typical $\boldsymbol{G}$, unique soln $(\boldsymbol{m}, \boldsymbol{n})$. $\boldsymbol{m} \approx$ barycenter, $\boldsymbol{n} \approx$ slacks

---

Key idea: **planted model**:

1. Sample $(\boldsymbol{m}, \boldsymbol{n})$ from its law (explicit physics prediction)
2. Sample $\boldsymbol{G}$ conditioned on TAP$(\boldsymbol{G}, \boldsymbol{m}, \boldsymbol{n})$

Belief: planted $\approx$ true model; 1st/2nd mt conditional on $(\boldsymbol{m}, \boldsymbol{n})$ locates threshold

---

Physics predictions for TAP & planted model

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

Physics prediction: for typical $\boldsymbol{G}$, unique soln $(\boldsymbol{m}, \boldsymbol{n})$. $\boldsymbol{m} \approx$ barycenter, $\boldsymbol{n} \approx$ slacks

---

Key idea: **planted model**:

1. Sample $(\boldsymbol{m}, \boldsymbol{n})$ from its law (explicit physics prediction)
2. Sample $\boldsymbol{G}$ conditioned on $\mathrm{TAP}(\boldsymbol{G}, \boldsymbol{m}, \boldsymbol{n})$

Belief: planted $\approx$ true model; 1st/2nd mt conditional on $(\boldsymbol{m}, \boldsymbol{n})$ locates threshold

---

! existence/uniqueness of $(\boldsymbol{m}, \boldsymbol{n})$ is **not proven**, so planted $\neq$ true possible

Proof roadmap

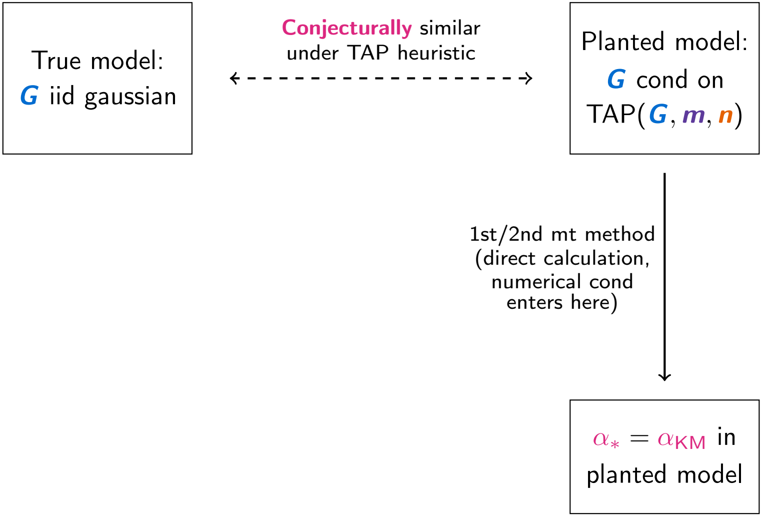
True model:
G iid gaussian

Conjecturally similar
under TAP heuristic

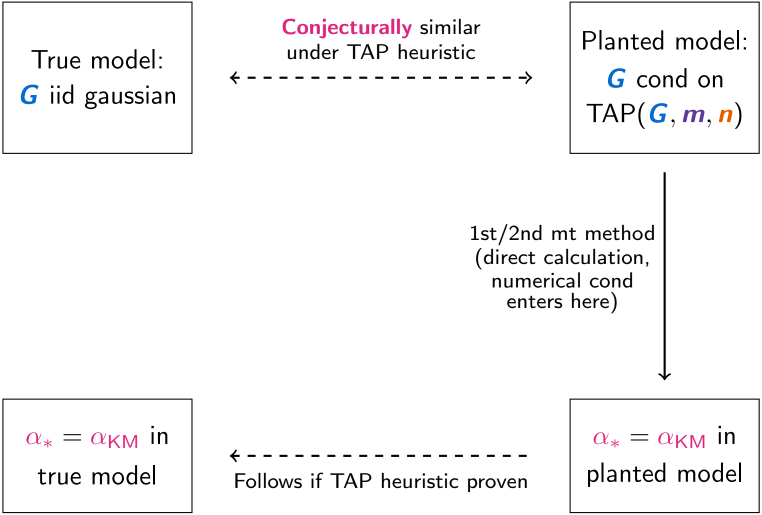
$\leftarrow - - - - - - - - - - - \rightarrow$

Planted model:
G cond on
$\mathrm{TAP}(G, m, n)$

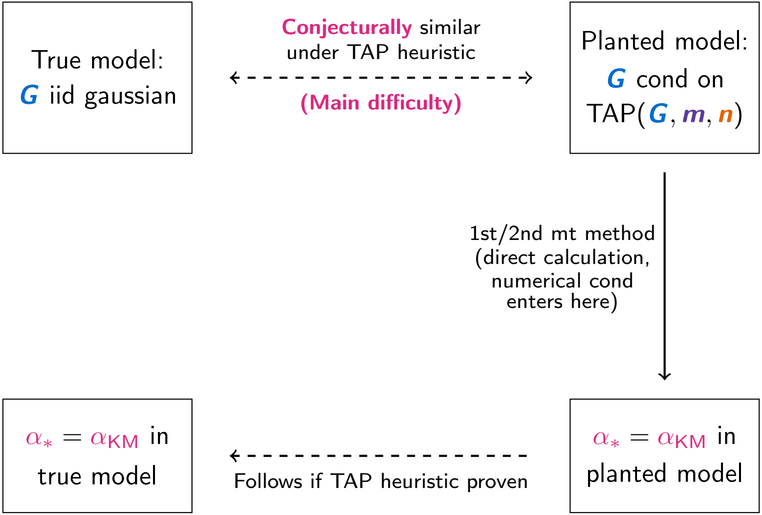
Proof roadmap



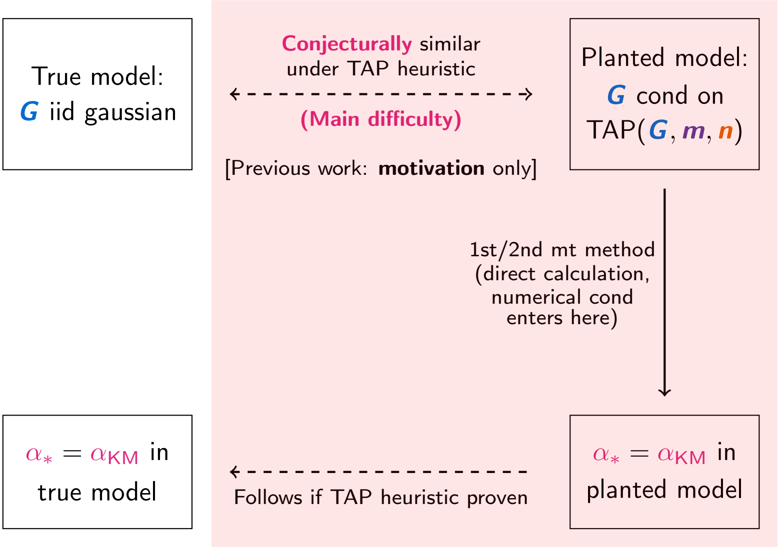
Proof roadmap



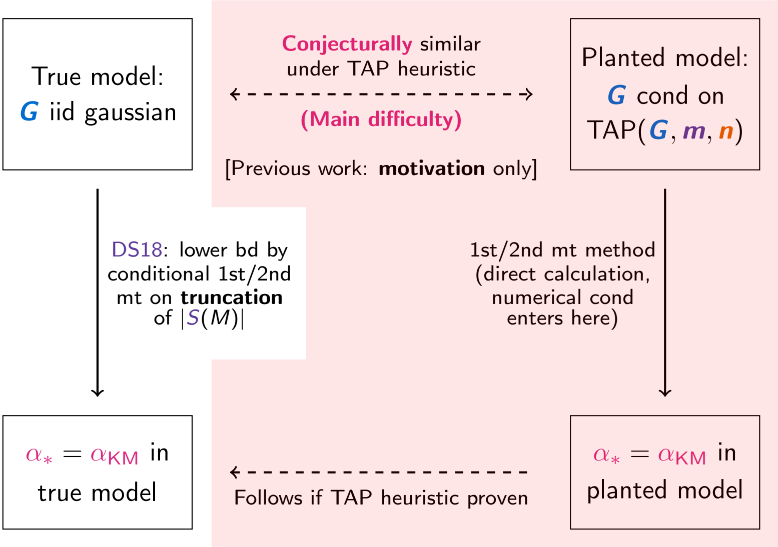
Proof roadmap



Proof roadmap



Proof roadmap



Proof roadmap

True model: $\boldsymbol{G}$ iid gaussian

Conjecturally similar under TAP heuristic

Our work proves this...

Planted model: $\boldsymbol{G}$ cond on $\mathrm{TAP}(\boldsymbol{G}, \boldsymbol{m}, \boldsymbol{n})$

DS18: lower bd by conditional 1st/2nd mt on **truncation** of $|S(M)|$

1st/2nd mt method (direct calculation, numerical cond enters here)

$\alpha_* = \alpha_{\mathsf{KM}}$ in true model

Follows if TAP heuristic proven

$\alpha_* = \alpha_{\mathsf{KM}}$ in planted model

Proof roadmap

True model: G iid gaussian

Conjecturally similar under TAP heuristic ←- - - - - - - - →

Our work proves this...

Planted model: G cond on $\mathrm{TAP}(G, m, n)$

↓ DS18: lower bd by conditional 1st/2nd mt on **truncation** of $|S(M)|$

↓ 1st/2nd mt method (direct calculation, numerical cond enters here)

$\alpha_* = \alpha_{\mathrm{KM}}$ in true model

...justifying this implication

← Follows if TAP heuristic proven

$\alpha_* = \alpha_{\mathrm{KM}}$ in planted model

Properties making the planted model tractable

Recall planted model:

- Sample $(\boldsymbol{m}, \boldsymbol{n})$ from its law
- Sample $\boldsymbol{G}$ conditional on $\mathrm{TAP}(\boldsymbol{G}, \boldsymbol{m}, \boldsymbol{n})$, i.e.

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

Properties making the planted model tractable

Recall planted model:

- Sample $(\boldsymbol{m}, \boldsymbol{n})$ from its law
- Sample $\boldsymbol{G}$ conditional on $\mathrm{TAP}(\boldsymbol{G}, \boldsymbol{m}, \boldsymbol{n})$, i.e.

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

---

1. The **coordinate profiles**

$$\mu(\boldsymbol{m}) = \frac{1}{N}\sum_{i=1}^{N} \delta(m_i) \qquad \nu(\boldsymbol{n}) = \frac{1}{M}\sum_{a=1}^{M} \delta(n_a) \qquad \in \mathcal{P}(\mathbb{R})$$

concentrate around explicit μ_*, ν_*. (Roughly, $m_i \overset{iid}{\sim} \mu_*$ & $n_a \overset{iid}{\sim} \nu_*$)

Properties making the planted model tractable

Recall planted model:

- Sample $(\boldsymbol{m}, \boldsymbol{n})$ from its law
- Sample $\boldsymbol{G}$ conditional on $\mathrm{TAP}(\boldsymbol{G}, \boldsymbol{m}, \boldsymbol{n})$, i.e.

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

---

1. The **coordinate profiles**

$$\mu(\boldsymbol{m}) = \frac{1}{N}\sum_{i=1}^{N}\delta(m_i) \qquad \nu(\boldsymbol{n}) = \frac{1}{M}\sum_{a=1}^{M}\delta(n_a) \qquad \in \mathcal{P}(\mathbb{R})$$

   concentrate around explicit μ_*, ν_*. (Roughly, $m_i \overset{iid}{\sim} \mu_*$ & $n_a \overset{iid}{\sim} \nu_*$)

2. $\mathrm{TAP}(\boldsymbol{G}, \boldsymbol{m}, \boldsymbol{n})$ is **linear** constraint on $\boldsymbol{G} \Rightarrow \boldsymbol{G}$ conditionally gaussian!

Moment calculation in planted model

Plan: 1st/2nd moment method on $|S| = \#(\text{solutions})$ **conditional on** $(\boldsymbol{m}, \boldsymbol{n})$

Moment calculation in planted model

Plan: 1st/2nd moment method on $|S| = \#(\text{solutions})$ **conditional on** $(\boldsymbol{m}, \boldsymbol{n})$

$\boldsymbol{G}$ is **spiked gaussian mtx**

Moment calculation in planted model

Plan: 1st/2nd moment method on $|S| = \#(\text{solutions})$ **conditional on** $(\boldsymbol{m}, \boldsymbol{n})$

$\boldsymbol{G}$ is **spiked gaussian mtx** $\Rightarrow$ $\boldsymbol{G}\boldsymbol{x}$ is **gaussian vector** whose law depends on only

$$a = \langle \boldsymbol{x}, \boldsymbol{m} \rangle \qquad b = \langle \boldsymbol{x}, \boldsymbol{H} \rangle \qquad\qquad (\boldsymbol{H} = \dot{F}^{-1}(\boldsymbol{m}))$$

Moment calculation in planted model

Plan: 1st/2nd moment method on $|S| = \#(\text{solutions})$ **conditional on** $(\boldsymbol{m}, \boldsymbol{n})$

$\boldsymbol{G}$ is **spiked gaussian mtx** $\Rightarrow$ $\boldsymbol{G}\boldsymbol{x}$ is **gaussian vector** whose law depends on only

$$a = \langle \boldsymbol{x}, \boldsymbol{m} \rangle \qquad b = \langle \boldsymbol{x}, \boldsymbol{H} \rangle \qquad\qquad (\boldsymbol{H} = \dot{F}^{-1}(\boldsymbol{m}))$$

$\Rightarrow \mathbb{P}(\boldsymbol{x} \text{ is solution}) = \mathbb{P}(\frac{\boldsymbol{G}\boldsymbol{x}}{\sqrt{N}} \geq \kappa\vec{1})$ is explicit function of a, b

Moment calculation in planted model

Plan: 1st/2nd moment method on $|S| = \#(\text{solutions})$ **conditional on** $(\boldsymbol{m}, \boldsymbol{n})$

$\boldsymbol{G}$ is **spiked gaussian mtx** $\Rightarrow$ $\boldsymbol{G}\boldsymbol{x}$ is **gaussian vector** whose law depends on only

$$a = \langle \boldsymbol{x}, \boldsymbol{m} \rangle \qquad b = \langle \boldsymbol{x}, \boldsymbol{H} \rangle \qquad\qquad (\boldsymbol{H} = \dot{F}^{-1}(\boldsymbol{m}))$$

$\Rightarrow \mathbb{P}(\boldsymbol{x} \text{ is solution}) = \mathbb{P}(\frac{\boldsymbol{G}\boldsymbol{x}}{\sqrt{N}} \geq \kappa \vec{1})$ is explicit function of a, b

$$\mathbb{E}[|S| | \boldsymbol{m}, \boldsymbol{n}] = \sum_{(a,b)} \quad \#(\boldsymbol{x} \in \{\pm 1\}^N \text{ with this } a, b) \quad \times \mathbb{P}(\text{one such } \boldsymbol{x} \text{ is solution})$$

Moment calculation in planted model

Plan: 1st/2nd moment method on $|S| = \#(\text{solutions})$ **conditional on** $(\boldsymbol{m}, \boldsymbol{n})$

$\boldsymbol{G}$ is **spiked gaussian mtx** $\Rightarrow$ $\boldsymbol{G}\boldsymbol{x}$ is **gaussian vector** whose law depends on only

$$a = \langle \boldsymbol{x}, \boldsymbol{m} \rangle \qquad b = \langle \boldsymbol{x}, \boldsymbol{H} \rangle \qquad\qquad (\boldsymbol{H} = \dot{F}^{-1}(\boldsymbol{m}))$$

$\Rightarrow \mathbb{P}(\boldsymbol{x} \text{ is solution}) = \mathbb{P}(\frac{\boldsymbol{G}\boldsymbol{x}}{\sqrt{N}} \geq \kappa \vec{1})$ is explicit function of a, b

$$\mathbb{E}[|S| | \boldsymbol{m}, \boldsymbol{n}] = \sum_{(a,b)} \underbrace{\#(\boldsymbol{x} \in \{\pm 1\}^N \text{ with this } a, b)}_{\text{explicit because } \boldsymbol{m} \text{ coordinate profile known}} \times \mathbb{P}(\text{one such } \boldsymbol{x} \text{ is solution})$$

Moment calculation in planted model

Plan: 1st/2nd moment method on $|S| = \#(\text{solutions})$ **conditional on** $(\boldsymbol{m}, \boldsymbol{n})$

$\boldsymbol{G}$ is **spiked gaussian mtx** $\Rightarrow$ $\boldsymbol{G}\boldsymbol{x}$ is **gaussian vector** whose law depends on only

$$a = \langle \boldsymbol{x}, \boldsymbol{m} \rangle \qquad b = \langle \boldsymbol{x}, \boldsymbol{H} \rangle \qquad (\boldsymbol{H} = \dot{F}^{-1}(\boldsymbol{m}))$$

$\Rightarrow \mathbb{P}(\boldsymbol{x} \text{ is solution}) = \mathbb{P}(\frac{\boldsymbol{G}\boldsymbol{x}}{\sqrt{N}} \geq \kappa \vec{1})$ is explicit function of a, b

$$\mathbb{E}[|S| | \boldsymbol{m}, \boldsymbol{n}] = \sum_{(a,b)} \underbrace{\#(\boldsymbol{x} \in \{\pm 1\}^N \text{ with this } a, b)}_{\text{explicit because } \boldsymbol{m} \text{ coordinate profile known}} \times \mathbb{P}(\text{one such } \boldsymbol{x} \text{ is solution})$$

Summand varies on **exponential scale**: $\text{contribution}(a, b) \approx \exp(N \cdot f(a, b))$

Moment calculation in planted model

Plan: 1st/2nd moment method on $|S| = \#(\text{solutions})$ **conditional on** $(\boldsymbol{m}, \boldsymbol{n})$

$\boldsymbol{G}$ is **spiked gaussian mtx** $\Rightarrow$ $\boldsymbol{G}\boldsymbol{x}$ is **gaussian vector** whose law depends on only

$$a = \langle \boldsymbol{x}, \boldsymbol{m} \rangle \qquad b = \langle \boldsymbol{x}, \boldsymbol{H} \rangle \qquad\qquad (\boldsymbol{H} = \dot{F}^{-1}(\boldsymbol{m}))$$

$\Rightarrow \mathbb{P}(\boldsymbol{x} \text{ is solution}) = \mathbb{P}(\frac{\boldsymbol{G}\boldsymbol{x}}{\sqrt{N}} \geq \kappa \vec{1})$ is explicit function of a, b

$$\mathbb{E}[|S| \,|\, \boldsymbol{m}, \boldsymbol{n}] = \sum_{(a,b)} \underbrace{\#(\boldsymbol{x} \in \{\pm 1\}^N \text{ with this } a, b)}_{\text{explicit because } \boldsymbol{m} \text{ coordinate profile known}} \times \mathbb{P}(\text{one such } \boldsymbol{x} \text{ is solution})$$

Summand varies on **exponential scale**: contribution$(a, b) \approx \exp(N \cdot f(a, b))$

$\Rightarrow \mathbb{E}[|S| \,|\, \boldsymbol{m}, \boldsymbol{n}] \approx \exp(N \cdot \max f(a, b))$ essentially **2 variable maximization**

1st/2nd moment in planted model

1st/2nd moments are both explicit $O(1)$-variable maximizations.

1st/2nd moment in planted model

1st/2nd moments are both explicit $O(1)$-variable maximizations.

1st/2nd moment works!! Conditional on typical $(\boldsymbol{m}, \boldsymbol{n})$,

$$\mathbb{E}[|S(\alpha N)||\boldsymbol{m}, \boldsymbol{n}] \quad \approx \quad \mathbb{E}[|S(\alpha N)|^2|\boldsymbol{m}, \boldsymbol{n}]^{1/2} \quad \approx \quad \exp(N \,\mathrm{Vol}(\alpha))$$

under our + DS18's numerical conditions

1st/2nd moment in planted model

1st/2nd moments are both explicit $O(1)$-variable maximizations.

1st/2nd moment works!! Conditional on typical $(\boldsymbol{m}, \boldsymbol{n})$,

$$\mathbb{E}[|S(\alpha N)||\boldsymbol{m}, \boldsymbol{n}] \quad \approx \quad \mathbb{E}[|S(\alpha N)|^2|\boldsymbol{m}, \boldsymbol{n}]^{1/2} \quad \approx \quad \exp(N\,\mathrm{Vol}(\alpha))$$

under our + DS18's numerical conditions $\Rightarrow$ typically $|S(\alpha N)| \approx \exp(N\,\mathrm{Vol}(\alpha))$

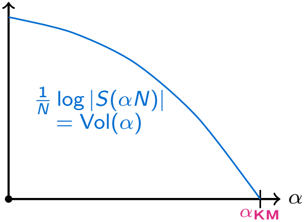
1st/2nd moment in planted model

1st/2nd moments are both explicit $O(1)$-variable maximizations.

1st/2nd moment works!! Conditional on typical $(\boldsymbol{m}, \boldsymbol{n})$,

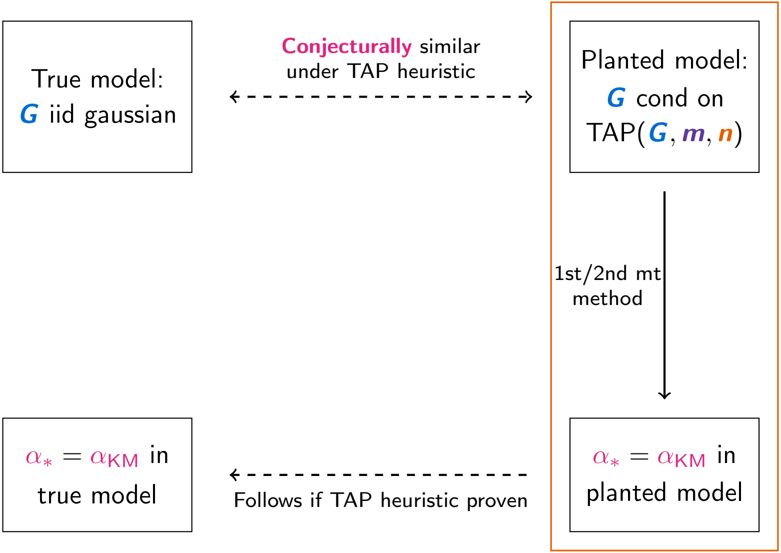
$$\mathbb{E}[|S(\alpha N)||\boldsymbol{m}, \boldsymbol{n}] \quad \approx \quad \mathbb{E}[|S(\alpha N)|^2|\boldsymbol{m}, \boldsymbol{n}]^{1/2} \quad \approx \quad \exp(N \,\mathrm{Vol}(\alpha))$$

under our + DS18's numerical conditions ⇒ typically $|S(\alpha N)| \approx \exp(N\,\mathrm{Vol}(\alpha))$

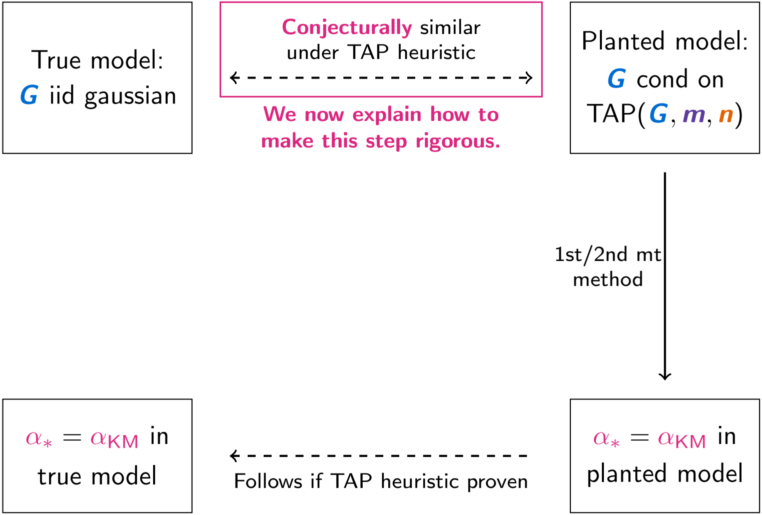


Since Vol has root α_{KM}, planted model has capacity α_{KM}

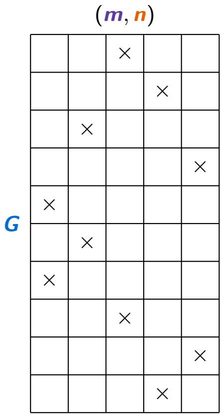
Proof roadmap



Proof roadmap



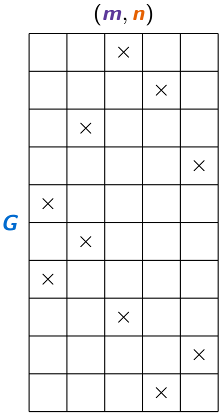
Key issue: linking true and planted models



True model $\leftrightarrow$ random row

Planted model $\leftrightarrow$ random col, then random $\times$ in col

Key issue: linking true and planted models



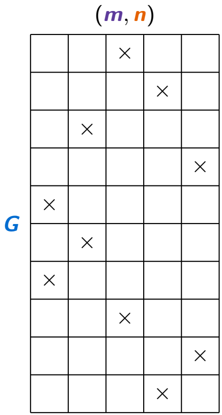
True model $\leftrightarrow$ random row

Planted model $\leftrightarrow$ random col, then random $\times$ in col

Under TAP prediction, most rows have exactly one $\times$

so true $\approx$ planted

Key issue: linking true and planted models

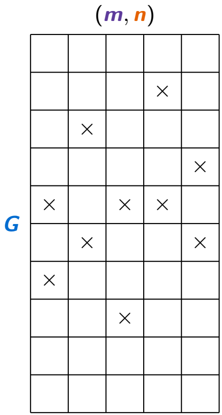


True model $\leftrightarrow$ random row
Planted model $\leftrightarrow$ random col, then random $\times$ in col

Under TAP prediction, most rows have exactly one $\times$
so true $\approx$ planted

but... **we don't actually know this** ☹

Key issue: linking true and planted models



True model ↔ random row

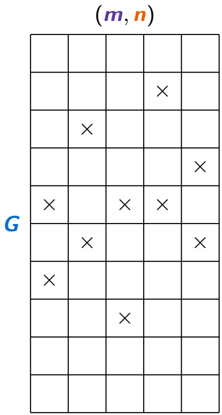
Planted model ↔ random col, then random × in col

Under TAP prediction, most rows have exactly one ×

so true ≈ planted

but. . . **we don't actually know this** ☹

Key issue: linking true and planted models



True model $\leftrightarrow$ random row

Planted model $\leftrightarrow$ random col, then random $\times$ in col

Under TAP prediction, most rows have exactly one $\times$

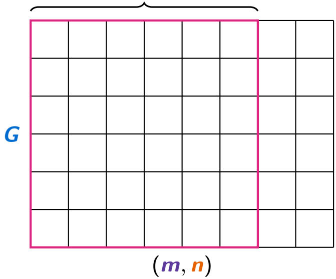
so true $\approx$ planted

but. . . **we don't actually know this** ☹

$\Rightarrow$ planted / true models can a priori be very different

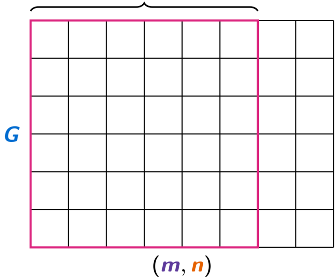
This work: contiguity of true / planted models

$T = \{(m, n)$ with $\approx \underbrace{\text{ideal coordinate profiles}}\}$ (high-probability set)



This work: contiguity of true / planted models

$\boldsymbol{T} = \{(\boldsymbol{m}, \boldsymbol{n})$ with $\approx$ ideal coordinate profiles$\}$ (high-probability set)

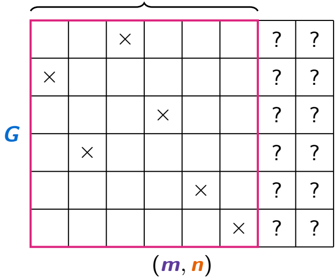


We show, for $\boldsymbol{G} \sim$ true model:

1. Existence: $\boldsymbol{G}$ has TAP solution $(\boldsymbol{m}, \boldsymbol{n}) \in \boldsymbol{T}$ whp (most rows have a $\times$)
2. Uniqueness: $\mathbb{E}[\#\text{TAP solutions in } \boldsymbol{T}] = 1 + o(1)$ (rows average $1 + o(1)$ $\times$'s)

This work: contiguity of true / planted models

$\boldsymbol{T} = \{(\boldsymbol{m}, \boldsymbol{n})$ with $\approx$ ideal coordinate profiles$\}$ (high-probability set)

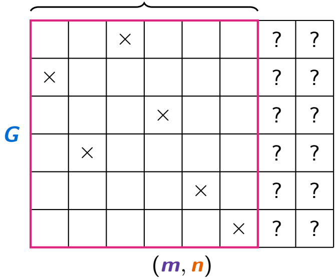


We show, for $\boldsymbol{G} \sim$ true model:

1. Existence: $\boldsymbol{G}$ has TAP solution $(\boldsymbol{m}, \boldsymbol{n}) \in \boldsymbol{T}$ whp (most rows have a $\times$)
2. Uniqueness: $\mathbb{E}[\#\text{TAP solutions in } \boldsymbol{T}] = 1 + o(1)$ (rows average $1 + o(1)$ $\times$'s)

This work: contiguity of true / planted models

$T = \{(m, n)$ with $\approx$ ideal coordinate profiles$\}$ (high-probability set)



We show, for $G \sim$ true model:

1. Existence: G has TAP solution $(m, n) \in T$ whp (most rows have a $\times$)
2. Uniqueness: $\mathbb{E}[\#\text{TAP solutions in } T] = 1 + o(1)$ (rows average $1 + o(1)$ $\times$'s)

This shows true $\approx$ planted. That is, $\forall$ event E,

$$\mathbb{P}_{\text{true}}(E) \quad \leq \quad C \sup_{(m,n) \in T} \mathbb{P}_{\text{planted}}(E | m, n) + o(1)$$

Existence: algorithmic proof

Want: $\boldsymbol{G} \sim$ true model, $\boldsymbol{G}$ has TAP fixed pt $(\boldsymbol{m}, \boldsymbol{n}) \in \boldsymbol{T} = \{\text{correct profiles}\}$ whp

Existence: algorithmic proof

Want: $\boldsymbol{G} \sim$ true model, $\boldsymbol{G}$ has TAP fixed pt $(\boldsymbol{m}, \boldsymbol{n}) \in \boldsymbol{T} = \{\text{correct profiles}\}$ whp

Will show **approximate message passing** (AMP) finds such a point:

$$\boldsymbol{m}^{k+1} = \dot{F}\left(\frac{\boldsymbol{G}^\top \boldsymbol{n}^k}{\sqrt{N}} - d\boldsymbol{m}^k\right) \qquad \boldsymbol{n}^k = \widehat{F}\left(\frac{\boldsymbol{G}\boldsymbol{m}^k}{\sqrt{N}} - b\boldsymbol{n}^{k-1}\right)$$

Existence: algorithmic proof

Want: $\boldsymbol{G} \sim$ true model, $\boldsymbol{G}$ has TAP fixed pt $(\boldsymbol{m}, \boldsymbol{n}) \in \boldsymbol{T} = \{\text{correct profiles}\}$ whp

Will show **approximate message passing** (AMP) finds such a point:

$$\boldsymbol{m}^{k+1} = \dot{F}\left(\frac{\boldsymbol{G}^\top \boldsymbol{n}^k}{\sqrt{N}} - d\boldsymbol{m}^k\right) \qquad \boldsymbol{n}^k = \widehat{F}\left(\frac{\boldsymbol{G}\boldsymbol{m}^k}{\sqrt{N}} - b\boldsymbol{n}^{k-1}\right)$$

---

1. TAP fixed points are critical points of **TAP free energy** $\mathcal{F}_{\mathrm{TAP}}(\boldsymbol{m}, \boldsymbol{n}; \boldsymbol{G})$

Existence: algorithmic proof

Want: $\boldsymbol{G} \sim$ true model, $\boldsymbol{G}$ has TAP fixed pt $(\boldsymbol{m}, \boldsymbol{n}) \in \boldsymbol{T} = \{\text{correct profiles}\}$ whp

Will show **approximate message passing** (AMP) finds such a point:

$$\boldsymbol{m}^{k+1} = \dot{F}\left(\frac{\boldsymbol{G}^\top \boldsymbol{n}^k}{\sqrt{N}} - d\boldsymbol{m}^k\right) \qquad \boldsymbol{n}^k = \widehat{F}\left(\frac{\boldsymbol{G}\boldsymbol{m}^k}{\sqrt{N}} - b\boldsymbol{n}^{k-1}\right)$$

---

1. TAP fixed points are critical points of **TAP free energy** $\mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m}, \boldsymbol{n}; \boldsymbol{G})$
2. State evolution (Bolthausen 14, Bayati Montanari 11) $\Rightarrow$ for large $k = O(1)$,

$$(\boldsymbol{m}^k, \boldsymbol{n}^k) \in \boldsymbol{T} \qquad \|\nabla \mathcal{F}_{\mathsf{TAP}}(\boldsymbol{m}^k, \boldsymbol{n}^k)\| = o_k(1)$$

Existence: algorithmic proof

Want: $\boldsymbol{G} \sim$ true model, $\boldsymbol{G}$ has TAP fixed pt $(\boldsymbol{m}, \boldsymbol{n}) \in \boldsymbol{T} = \{\text{correct profiles}\}$ whp

Will show **approximate message passing** (AMP) finds such a point:

$$\boldsymbol{m}^{k+1} = \dot{F}\left(\frac{\boldsymbol{G}^\top \boldsymbol{n}^k}{\sqrt{N}} - d\boldsymbol{m}^k\right) \qquad \boldsymbol{n}^k = \widehat{F}\left(\frac{\boldsymbol{G}\boldsymbol{m}^k}{\sqrt{N}} - b\boldsymbol{n}^{k-1}\right)$$

---

1. TAP fixed points are critical points of **TAP free energy** $\mathcal{F}_{\mathrm{TAP}}(\boldsymbol{m}, \boldsymbol{n}; \boldsymbol{G})$
2. State evolution (Bolthausen 14, Bayati Montanari 11) $\Rightarrow$ for large $k = O(1)$,

$$(\boldsymbol{m}^k, \boldsymbol{n}^k) \in \boldsymbol{T} \qquad \|\nabla \mathcal{F}_{\mathrm{TAP}}(\boldsymbol{m}^k, \boldsymbol{n}^k)\| = o_k(1)$$

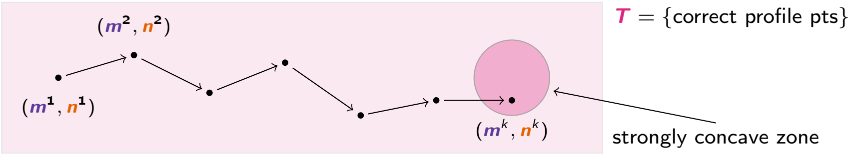
That is, AMP finds an **approximate** critical point in $\boldsymbol{T}$

Existence: from approximate to exact critical point

Celentano Fan Mei 21: if $\mathcal{F}_{\text{TAP}}$ **strongly concave** near the approximate critical point $(\boldsymbol{m}^k, \boldsymbol{n}^k)$, exists exact critical point nearby

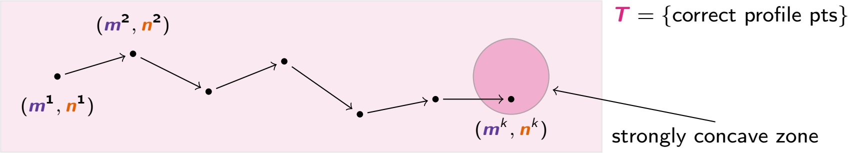
Existence: from approximate to exact critical point

Celentano Fan Mei 21: if $\mathcal{F}_{\text{TAP}}$ **strongly concave** near the approximate critical point $(\boldsymbol{m}^k, \boldsymbol{n}^k)$, exists exact critical point nearby



Existence: from approximate to exact critical point

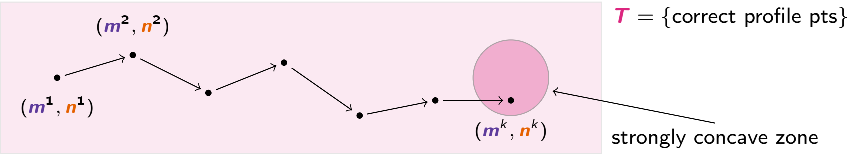
Celentano Fan Mei 21: if $\mathcal{F}_{\text{TAP}}$ **strongly concave** near the approximate critical point $(\boldsymbol{m}^k, \boldsymbol{n}^k)$, exists exact critical point nearby



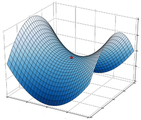
In our setting, $\mathcal{F}_{\text{TAP}}(\boldsymbol{m}, \boldsymbol{n})$ **not strongly concave** near $(\boldsymbol{m}^k, \boldsymbol{n}^k)$...

Existence: from approximate to exact critical point

Celentano Fan Mei 21: if $\mathcal{F}_{\text{TAP}}$ **strongly concave** near the approximate critical point $(\boldsymbol{m}^k, \boldsymbol{n}^k)$, exists exact critical point nearby



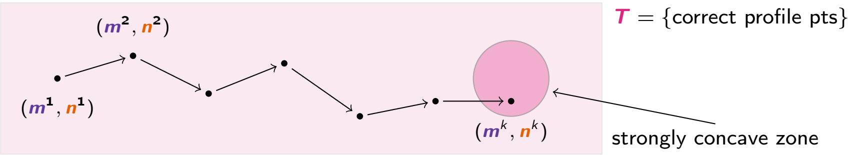
In our setting, $\mathcal{F}_{\text{TAP}}(\boldsymbol{m}, \boldsymbol{n})$ **not strongly concave** near $(\boldsymbol{m}^k, \boldsymbol{n}^k)$...



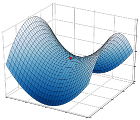
...but is **strongly convex-concave**, which also works

Existence: from approximate to exact critical point

Celentano Fan Mei 21: if $\mathcal{F}_{\mathrm{TAP}}$ **strongly concave** near the approximate critical point $(\boldsymbol{m}^k, \boldsymbol{n}^k)$, exists exact critical point nearby



In our setting, $\mathcal{F}_{\mathrm{TAP}}(\boldsymbol{m}, \boldsymbol{n})$ **not strongly concave** near $(\boldsymbol{m}^k, \boldsymbol{n}^k)$...



...but is **strongly convex-concave**, which also works
Proof adapts AMP-conditioned gaussian comparison approach of Celentano 22

Uniqueness: double-counting argument

Want: for $\boldsymbol{G} \sim$ true model, $\mathbb{E}[\#\text{TAP fixed pts of } \boldsymbol{G} \text{ in } \boldsymbol{T}] = 1 + o(1)$

Uniqueness: double-counting argument

Want: for $\boldsymbol{G} \sim$ true model, $\mathbb{E}[\#\text{TAP fixed pts of } \boldsymbol{G} \text{ in } \boldsymbol{T}] = 1 + o(1)$

This also has an algorithmic proof! Following claim implies uniqueness:

Uniqueness: double-counting argument

Want: for $\boldsymbol{G} \sim$ true model, $\mathbb{E}[\#\text{TAP fixed pts of } \boldsymbol{G} \text{ in } \boldsymbol{T}] = 1 + o(1)$

This also has an algorithmic proof! Following claim implies uniqueness:

Fix $(\boldsymbol{m}, \boldsymbol{n}) \in \boldsymbol{T} = \{\text{correct profiles}\}$. Sample $\boldsymbol{G}$ conditioned on $\text{TAP}(\boldsymbol{G}, \boldsymbol{m}, \boldsymbol{n})$.

AMP run on $\boldsymbol{G}$ finds **the planted point** $(\boldsymbol{m}, \boldsymbol{n})$ whp

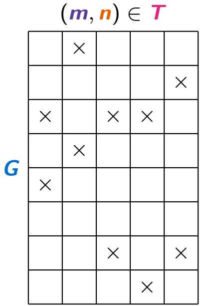
Uniqueness: double-counting argument

Want: for $\boldsymbol{G} \sim$ true model, $\mathbb{E}[\#\text{TAP fixed pts of } \boldsymbol{G} \text{ in } \boldsymbol{T}] = 1 + o(1)$

💡 This also has an algorithmic proof! Following claim implies uniqueness:

Fix $(\boldsymbol{m}, \boldsymbol{n}) \in \boldsymbol{T} = \{\text{correct profiles}\}$. Sample $\boldsymbol{G}$ conditioned on $\text{TAP}(\boldsymbol{G}, \boldsymbol{m}, \boldsymbol{n})$.

AMP run on $\boldsymbol{G}$ finds **the planted point** $(\boldsymbol{m}, \boldsymbol{n})$ whp



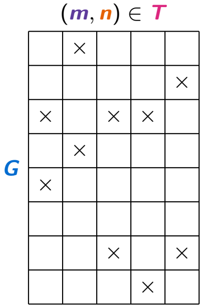
Uniqueness: double-counting argument

Want: for $G \sim$ true model, $\mathbb{E}[\#\text{TAP fixed pts of } G \text{ in } T] = 1 + o(1)$

This also has an algorithmic proof! Following claim implies uniqueness:

Fix $(m, n) \in T = \{\text{correct profiles}\}$. Sample G conditioned on $\text{TAP}(G, m, n)$.

AMP run on G finds **the planted point** (m, n) whp



Experiment: choose $(m, n) \in T$

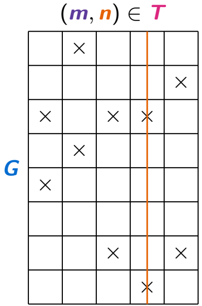
Uniqueness: double-counting argument

Want: for $G \sim$ true model, $\mathbb{E}[\#\text{TAP fixed pts of } G \text{ in } T] = 1 + o(1)$

This also has an algorithmic proof! Following claim implies uniqueness:

Fix $(m, n) \in T = \{\text{correct profiles}\}$. Sample G conditioned on $\text{TAP}(G, m, n)$.

AMP run on G finds **the planted point** (m, n) whp



Experiment: choose $(m, n) \in T$

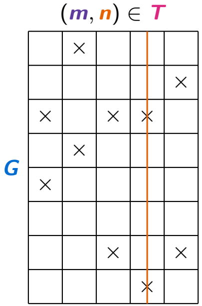
Uniqueness: double-counting argument

Want: for $G \sim$ true model, $\mathbb{E}[\#\text{TAP fixed pts of } G \text{ in } T] = 1 + o(1)$

This also has an algorithmic proof! Following claim implies uniqueness:

Fix $(m, n) \in T = \{\text{correct profiles}\}$. Sample G conditioned on $\mathrm{TAP}(G, m, n)$.

AMP run on G finds **the planted point** (m, n) whp



Experiment: choose $(m, n) \in T$

Sample G conditional on $\mathrm{TAP}(G, m, n)$

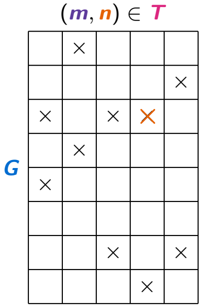
Uniqueness: double-counting argument

Want: for $G \sim$ true model, $\mathbb{E}[\#\text{TAP fixed pts of } G \text{ in } T] = 1 + o(1)$

💡 This also has an algorithmic proof! Following claim implies uniqueness:

Fix $(m, n) \in T = \{\text{correct profiles}\}$. Sample G conditioned on $\mathrm{TAP}(G, m, n)$.

AMP run on G finds **the planted point** (m, n) whp



Experiment: choose $(m, n) \in T$

Sample G conditional on $\mathrm{TAP}(G, m, n)$

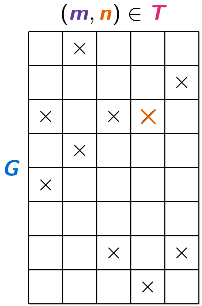
Uniqueness: double-counting argument

Want: for $\boldsymbol{G} \sim$ true model, $\mathbb{E}[\#\text{TAP fixed pts of } \boldsymbol{G} \text{ in } \boldsymbol{T}] = 1 + o(1)$

💡 This also has an algorithmic proof! Following claim implies uniqueness:

Fix $(\boldsymbol{m}, \boldsymbol{n}) \in \boldsymbol{T} = \{\text{correct profiles}\}$. Sample $\boldsymbol{G}$ conditioned on $\text{TAP}(\boldsymbol{G}, \boldsymbol{m}, \boldsymbol{n})$.

AMP run on $\boldsymbol{G}$ finds **the planted point** $(\boldsymbol{m}, \boldsymbol{n})$ whp



Experiment: choose $(\boldsymbol{m}, \boldsymbol{n}) \in \boldsymbol{T}$

Sample $\boldsymbol{G}$ conditional on $\text{TAP}(\boldsymbol{G}, \boldsymbol{m}, \boldsymbol{n})$

Run AMP on disorder $\boldsymbol{G}$

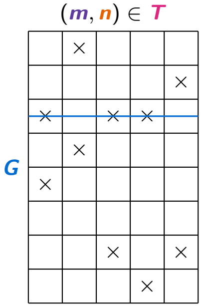
Uniqueness: double-counting argument

Want: for $\boldsymbol{G} \sim$ true model, $\mathbb{E}[\#\text{TAP fixed pts of } \boldsymbol{G} \text{ in } \boldsymbol{T}] = 1 + o(1)$

💡 This also has an algorithmic proof! Following claim implies uniqueness:

Fix $(\boldsymbol{m}, \boldsymbol{n}) \in \boldsymbol{T} = \{\text{correct profiles}\}$. Sample $\boldsymbol{G}$ conditioned on $\mathrm{TAP}(\boldsymbol{G}, \boldsymbol{m}, \boldsymbol{n})$.

AMP run on $\boldsymbol{G}$ finds **the planted point** $(\boldsymbol{m}, \boldsymbol{n})$ whp



Experiment: choose $(\boldsymbol{m}, \boldsymbol{n}) \in \boldsymbol{T}$

Sample $\boldsymbol{G}$ conditional on $\mathrm{TAP}(\boldsymbol{G}, \boldsymbol{m}, \boldsymbol{n})$

Run AMP on disorder $\boldsymbol{G}$

Uniqueness: double-counting argument

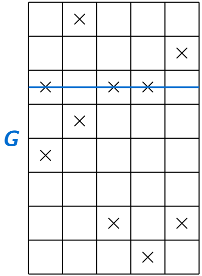
Want: for $G \sim$ true model, $\mathbb{E}[\#\text{TAP fixed pts of } G \text{ in } T] = 1 + o(1)$

💡 This also has an algorithmic proof! Following claim implies uniqueness:

Fix $(m, n) \in T = \{\text{correct profiles}\}$. Sample G conditioned on $\text{TAP}(G, m, n)$.

AMP run on G finds **the planted point** (m, n) whp

$(m, n) \in T$



Experiment: choose $(m, n) \in T$

Sample G conditional on $\text{TAP}(G, m, n)$

Run AMP on disorder G

Did AMP return to (m, n)?

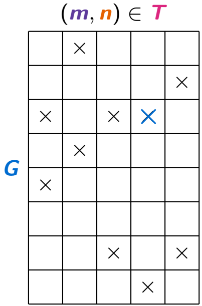
Uniqueness: double-counting argument

Want: for $G \sim$ true model, $\mathbb{E}[\#\text{TAP fixed pts of } G \text{ in } T] = 1 + o(1)$

This also has an algorithmic proof! Following claim implies uniqueness:

Fix $(m, n) \in T = \{\text{correct profiles}\}$. Sample G conditioned on $\mathrm{TAP}(G, m, n)$.

AMP run on G finds **the planted point** (m, n) whp



Experiment: choose $(m, n) \in T$

Sample G conditional on $\mathrm{TAP}(G, m, n)$

Run AMP on disorder G

Did AMP return to (m, n)?

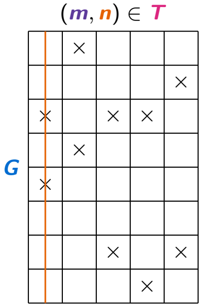
Uniqueness: double-counting argument

Want: for $G \sim$ true model, $\mathbb{E}[\#\text{TAP fixed pts of } G \text{ in } T] = 1 + o(1)$

This also has an algorithmic proof! Following claim implies uniqueness:

Fix $(m, n) \in T = \{\text{correct profiles}\}$. Sample G conditioned on $\text{TAP}(G, m, n)$.

AMP run on G finds **the planted point** (m, n) whp



Experiment: choose $(m, n) \in T$

Sample G conditional on $\text{TAP}(G, m, n)$

Run AMP on disorder G

Did AMP return to (m, n)?

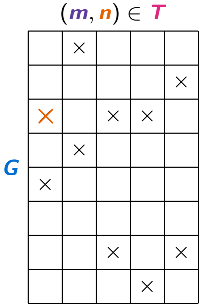
Uniqueness: double-counting argument

Want: for $G \sim$ true model, $\mathbb{E}[\#\text{TAP fixed pts of } G \text{ in } T] = 1 + o(1)$

💡 This also has an algorithmic proof! Following claim implies uniqueness:

Fix $(m, n) \in T = \{\text{correct profiles}\}$. Sample G conditioned on $\mathrm{TAP}(G, m, n)$.

AMP run on G finds **the planted point** (m, n) whp



Experiment: choose $(m, n) \in T$

Sample G conditional on $\mathrm{TAP}(G, m, n)$

Run AMP on disorder G

Did AMP return to (m, n)?

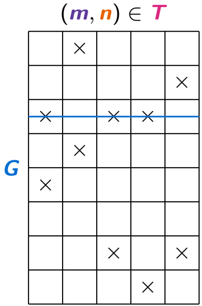
Uniqueness: double-counting argument

Want: for $\boldsymbol{G} \sim$ true model, $\mathbb{E}[\#\text{TAP fixed pts of } \boldsymbol{G} \text{ in } \boldsymbol{T}] = 1 + o(1)$

This also has an algorithmic proof! Following claim implies uniqueness:

Fix $(\boldsymbol{m}, \boldsymbol{n}) \in \boldsymbol{T} = \{\text{correct profiles}\}$. Sample $\boldsymbol{G}$ conditioned on $\text{TAP}(\boldsymbol{G}, \boldsymbol{m}, \boldsymbol{n})$.

AMP run on $\boldsymbol{G}$ finds **the planted point** $(\boldsymbol{m}, \boldsymbol{n})$ whp



Experiment: choose $(\boldsymbol{m}, \boldsymbol{n}) \in \boldsymbol{T}$

Sample $\boldsymbol{G}$ conditional on $\text{TAP}(\boldsymbol{G}, \boldsymbol{m}, \boldsymbol{n})$

Run AMP on disorder $\boldsymbol{G}$

Did AMP return to $(\boldsymbol{m}, \boldsymbol{n})$?

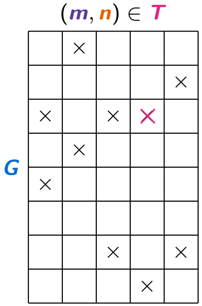
Uniqueness: double-counting argument

Want: for $G \sim$ true model, $\mathbb{E}[\#\text{TAP fixed pts of } G \text{ in } T] = 1 + o(1)$

This also has an algorithmic proof! Following claim implies uniqueness:

Fix $(m, n) \in T = \{\text{correct profiles}\}$. Sample G conditioned on $\text{TAP}(G, m, n)$.

AMP run on G finds **the planted point** (m, n) whp



Experiment: choose $(m, n) \in T$

Sample G conditional on $\text{TAP}(G, m, n)$

Run AMP on disorder G

Did AMP return to (m, n)?

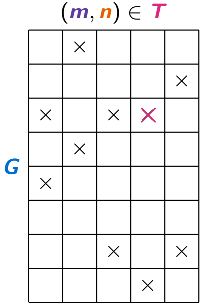
Uniqueness: double-counting argument

Want: for $G \sim$ true model, $\mathbb{E}[\#\text{TAP fixed pts of } G \text{ in } T] = 1 + o(1)$

This also has an algorithmic proof! Following claim implies uniqueness:

Fix $(m, n) \in T = \{\text{correct profiles}\}$. Sample G conditioned on $\text{TAP}(G, m, n)$.

AMP run on G finds **the planted point** (m, n) whp



Experiment: choose $(m, n) \in T$

Sample G conditional on $\text{TAP}(G, m, n)$

Run AMP on disorder G

Did AMP return to (m, n)?

Experiment succeeds for at most one $\times$ per row

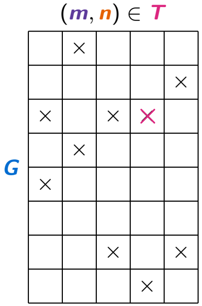
Uniqueness: double-counting argument

Want: for $\boldsymbol{G} \sim$ true model, $\mathbb{E}[\#\text{TAP fixed pts of } \boldsymbol{G} \text{ in } \boldsymbol{T}] = 1 + o(1)$

This also has an algorithmic proof! Following claim implies uniqueness:

Fix $(\boldsymbol{m}, \boldsymbol{n}) \in \boldsymbol{T} = \{\text{correct profiles}\}$. Sample $\boldsymbol{G}$ conditioned on $\text{TAP}(\boldsymbol{G}, \boldsymbol{m}, \boldsymbol{n})$.

AMP run on $\boldsymbol{G}$ finds **the planted point** $(\boldsymbol{m}, \boldsymbol{n})$ whp



Experiment: choose $(\boldsymbol{m}, \boldsymbol{n}) \in \boldsymbol{T}$

Sample $\boldsymbol{G}$ conditional on $\text{TAP}(\boldsymbol{G}, \boldsymbol{m}, \boldsymbol{n})$

Run AMP on disorder $\boldsymbol{G}$

Did AMP return to $(\boldsymbol{m}, \boldsymbol{n})$?

Experiment succeeds for at most one $\times$ per row

If too many rows have > 1 $\times$s, claim cannot be true!

Uniqueness: AMP returns home in planted model

Want: for $(\boldsymbol{m}, \boldsymbol{n}) \in \boldsymbol{T}$, $\boldsymbol{G}$ conditioned on $\text{TAP}(\boldsymbol{G}, \boldsymbol{m}, \boldsymbol{n})$,

AMP run on $\boldsymbol{G}$ finds **the planted point** $(\boldsymbol{m}, \boldsymbol{n})$ whp

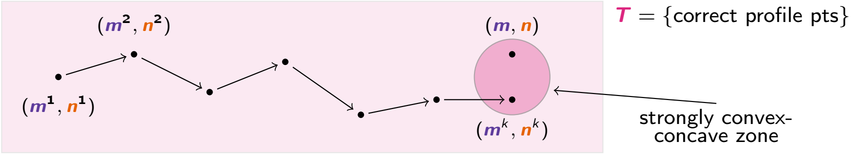
---

Uniqueness: AMP returns home in planted model

Want: for $(m, n) \in T$, G conditioned on TAP(G, m, n),

AMP run on G finds **the planted point** (m, n) whp

This can be proved by the same strategy!



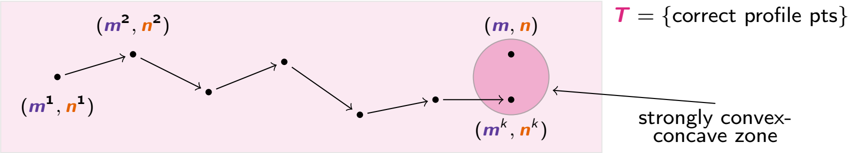
Uniqueness: AMP returns home in planted model

Want: for $(\boldsymbol{m}, \boldsymbol{n}) \in \boldsymbol{T}$, $\boldsymbol{G}$ conditioned on $\mathrm{TAP}(\boldsymbol{G}, \boldsymbol{m}, \boldsymbol{n})$,

AMP run on $\boldsymbol{G}$ finds **the planted point** $(\boldsymbol{m}, \boldsymbol{n})$ whp

---

This can be proved by the same strategy!



By adapting state evolution & gaussian comparison analyses to planted model:

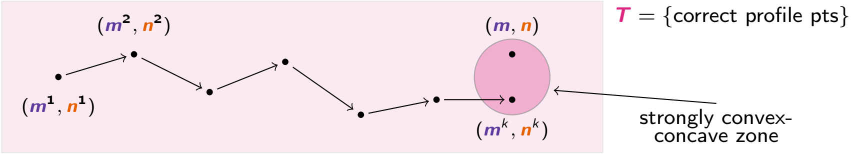
Uniqueness: AMP returns home in planted model

Want: for $(\boldsymbol{m}, \boldsymbol{n}) \in \boldsymbol{T}$, $\boldsymbol{G}$ conditioned on $\mathrm{TAP}(\boldsymbol{G}, \boldsymbol{m}, \boldsymbol{n})$,

AMP run on $\boldsymbol{G}$ finds **the planted point** $(\boldsymbol{m}, \boldsymbol{n})$ whp

---

This can be proved by the same strategy!



By adapting state evolution & gaussian comparison analyses to planted model:

- $(\boldsymbol{m}^k, \boldsymbol{n}^k) \in \boldsymbol{T}$, $\|\nabla \mathcal{F}_{\mathrm{TAP}}(\boldsymbol{m}^k, \boldsymbol{n}^k)\| = o_k(1)$, and $\|(\boldsymbol{m}^k, \boldsymbol{n}^k) - (\boldsymbol{m}, \boldsymbol{n})\| = o_k(1)$

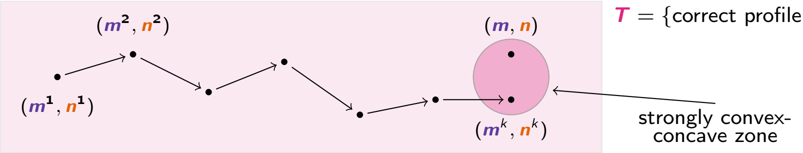
Uniqueness: AMP returns home in planted model

Want: for $(\boldsymbol{m}, \boldsymbol{n}) \in \boldsymbol{T}$, $\boldsymbol{G}$ conditioned on $\mathrm{TAP}(\boldsymbol{G}, \boldsymbol{m}, \boldsymbol{n})$,

AMP run on $\boldsymbol{G}$ finds **the planted point** $(\boldsymbol{m}, \boldsymbol{n})$ whp

---

This can be proved by the same strategy!

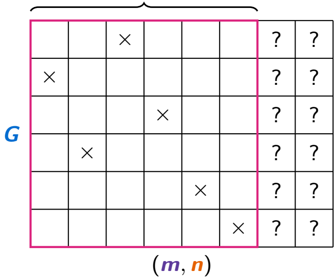


By adapting state evolution & gaussian comparison analyses to planted model:

- $(\boldsymbol{m}^k, \boldsymbol{n}^k) \in \boldsymbol{T}$, $\|\nabla \mathcal{F}_{\mathrm{TAP}}(\boldsymbol{m}^k, \boldsymbol{n}^k)\| = o_k(1)$, and $\|(\boldsymbol{m}^k, \boldsymbol{n}^k) - (\boldsymbol{m}, \boldsymbol{n})\| = o_k(1)$
- $\mathcal{F}_{\mathrm{TAP}}$ strongly convex-concave near $(\boldsymbol{m}^k, \boldsymbol{n}^k)$

Recap: contiguity of true / planted models

$T = \{(m, n) \text{ with} \approx \text{ideal coordinate profiles}\}$



We show, for $G \sim$ true model:

1. Existence: G has TAP solution $(m, n) \in T$ whp (most rows have a $\times$)
2. Uniqueness: $\mathbb{E}[\#\text{TAP solutions in } T] = 1 + o(1)$ (rows average $1 + o(1)$ $\times$'s)

This shows true $\approx$ planted.

Recap: proof roadmap

True model:
G iid gaussian

Conjecturally similar
under TAP heuristic

Our work proves this. . .

Planted model:
G cond on
TAP(**G**, **m**, **n**)

1st/2nd mt
method

$\alpha_* = \alpha_{\mathrm{KM}}$ in
true model

. . .justifying this implication

Follows if TAP heuristic proven

$\alpha_* = \alpha_{\mathrm{KM}}$ in
planted model

Conclusion

- We develop method to link true model & model with planted TAP fixed point
- Then 1st/2nd moment in planted model identifies capacity $\alpha_* = \alpha_{\mathsf{KM}}$

Conclusion

- We develop method to link true model & model with planted TAP fixed point
- Then 1st/2nd moment in planted model identifies capacity $\alpha_* = \alpha_{\mathsf{KM}}$

- Contiguity follows from existence / uniqueness of TAP fixed point
- Algorithmic proof of uniqueness via "AMP returns home in planted model"

Conclusion

- We develop method to link true model & model with planted TAP fixed point
- Then 1st/2nd moment in planted model identifies capacity $\alpha_* = \alpha_{\mathsf{KM}}$

- Contiguity follows from existence / uniqueness of TAP fixed point
- Algorithmic proof of uniqueness via "AMP returns home in planted model"

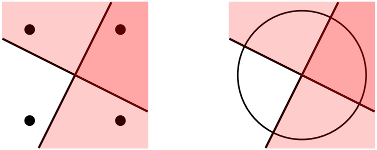
- **Open Q**: capacity of non replica symmetric models? E.g. spherical $\kappa < 0$

Conclusion

- We develop method to link true model & model with planted TAP fixed point
- Then 1st/2nd moment in planted model identifies capacity $\alpha_* = \alpha_{\mathsf{KM}}$

- Contiguity follows from existence / uniqueness of TAP fixed point
- Algorithmic proof of uniqueness via "AMP returns home in planted model"

- **Open Q**: capacity of non replica symmetric models? E.g. spherical $\kappa < 0$



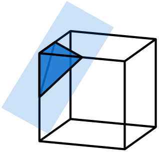
Thanks for your attention!

Earlier work: AMP-conditioned moment method

$$\boldsymbol{m}^{k+1} = \dot{F}\left(\frac{\boldsymbol{G}^\top \boldsymbol{n}^k}{\sqrt{N}} - d\boldsymbol{m}^k\right) \qquad \boldsymbol{n}^k = \widehat{F}\left(\frac{\boldsymbol{G}\boldsymbol{m}^k}{\sqrt{N}} - b\boldsymbol{n}^{k-1}\right)$$

Plan: for large $k = O(1)$, condition on $\boldsymbol{m}^1, \boldsymbol{n}^1, \ldots, \boldsymbol{m}^k, \boldsymbol{n}^k$, then 1st/2nd moment

No existence / uniqueness issue, but now $\mathbb{E}[|S(M)|\,|\text{AMP}]$ is k-dim optimization
(Over codimension-k slices of $\{\pm 1\}^N$ orthogonal to $\boldsymbol{m}^1, \ldots, \boldsymbol{m}^k$)



DS18: **for lower bound**, tractable 1st/2nd moment on **truncated** count

$$|S(M) \cap \{\text{correct affine slice}\}|$$

Upper bound: can't do truncation, optimization intractable