

Strong Freezing in Binary Perceptron Models

Shuyang Gong

LU-UMN Joint Probability Seminar

Joint work with



Shuangping Li

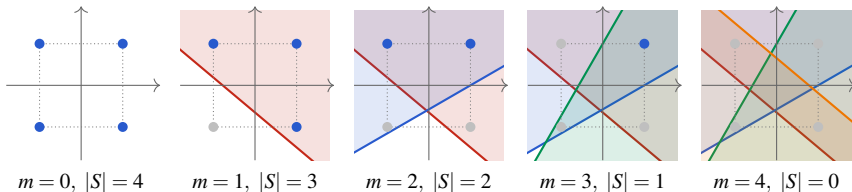
Yale University

Plan of the Talk

1. Motivation: the binary perceptron
2. Background: the physics prediction, progress so far, and algorithms
3. Main result: strong freezing at every density
4. The proof: Poissonization, FKG, and separation
5. Open problems

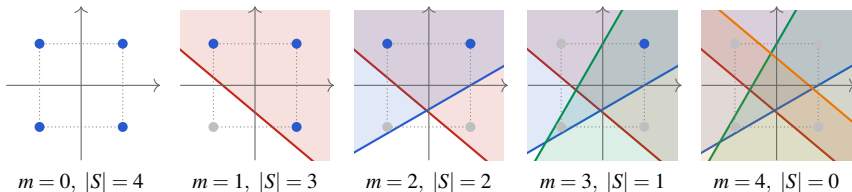
The Binary Perceptron Model

- Start with the discrete hypercube $\{\pm 1\}^n$



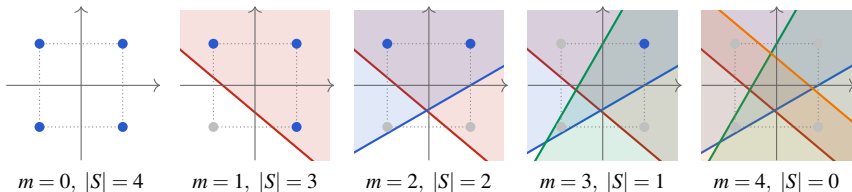
The Binary Perceptron Model

- Start with the discrete hypercube $\{\pm 1\}^n$
- m independent random half-spaces $H_1, \dots, H_m$
 - ▶ $H_a = \{\sigma \in \{\pm 1\}^n : \langle g_a, \sigma \rangle \geq \kappa \sqrt{n}\}$, with $g_1, \dots, g_m$ i.i.d. $N(0, I_n)$

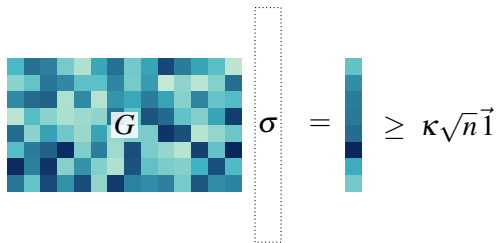


The Binary Perceptron Model

- Start with the discrete hypercube $\{\pm 1\}^n$
- m independent random half-spaces $H_1, \dots, H_m$
 - ▶ $H_a = \{\sigma \in \{\pm 1\}^n : \langle g_a, \sigma \rangle \geq \kappa \sqrt{n}\}$, with $g_1, \dots, g_m$ i.i.d. $N(0, I_n)$
- Define the solution set $S = \{\pm 1\}^n \cap H_1 \cap \dots \cap H_m$



The Binary Perceptron Model



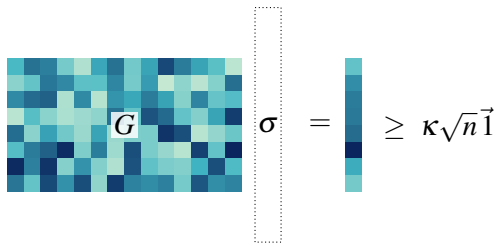
The diagram shows a matrix G (represented as a heatmap) multiplied by a vector σ (represented as a column of blue squares). The result is a vector of values greater than or equal to $\kappa\sqrt{n}\vec{1}$.

$$G \sigma = \geq \kappa\sqrt{n}\vec{1}$$

- $G \in \mathbb{R}^{m \times n}$ with i.i.d. standard normal entries; fix $\kappa \in \mathbb{R}$

$$S = \{ \sigma \in \{\pm 1\}^n : G\sigma \geq \kappa\sqrt{n} \text{ for each entry} \}$$

The Binary Perceptron Model

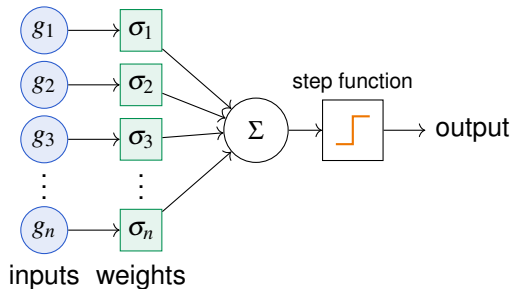

$$G \sigma = \geq \kappa \sqrt{n} \vec{1}$$

- $G \in \mathbb{R}^{m \times n}$ with i.i.d. standard normal entries; fix $\kappa \in \mathbb{R}$

$$S = \{ \sigma \in \{\pm 1\}^n : G \sigma \geq \kappa \sqrt{n} \text{ for each entry} \}$$

- $m/n \rightarrow \alpha$: the **constraint density**. This is the **asymmetric binary perceptron (ABP)**

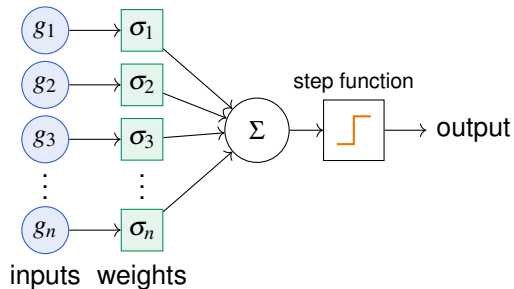
Connections to Neural Networks



$$\text{output} = \text{sign} \langle g, \sigma \rangle$$

- A perceptron with **binary weights**
 $\sigma \in \{\pm 1\}^n$

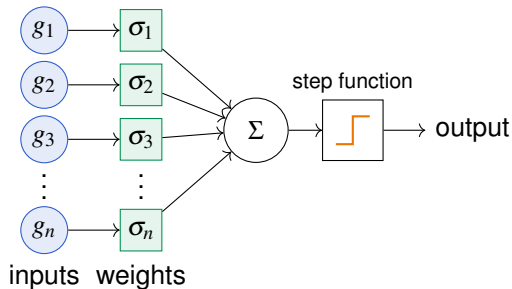
Connections to Neural Networks



$$\text{output} = \text{sign} \langle g, \sigma \rangle$$

- A perceptron with **binary weights**
 $\sigma \in \{\pm 1\}^n$
- Data $(g_a, \theta_a)_{a \leq m}$: Gaussian inputs, random labels $\theta_a = \pm 1$

Connections to Neural Networks



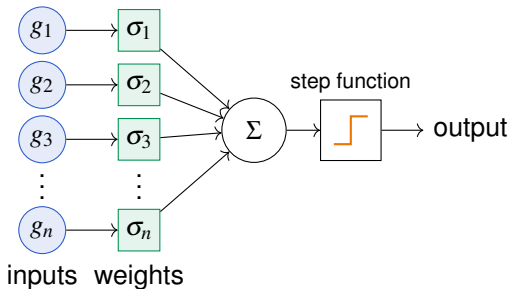
$$\text{output} = \text{sign} \langle g, \sigma \rangle$$

- A perceptron with **binary weights**
 $\sigma \in \{\pm 1\}^n$
- Data $(g_a, \theta_a)_{a \leq m}$: Gaussian inputs, random labels $\theta_a = \pm 1$
- Fit all the data:

$$\text{sign} \langle g_a, \sigma \rangle = \theta_a \iff \langle \theta_a g_a, \sigma \rangle \geq 0$$

- This is the ABP with $\kappa = 0$

Connections to Neural Networks



$$\text{output} = \text{sign} \langle g, \sigma \rangle$$

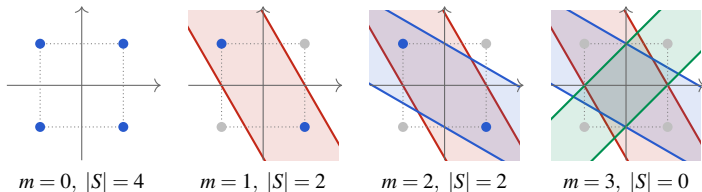
- A perceptron with **binary weights** $\sigma \in \{\pm 1\}^n$
- Data $(g_a, \theta_a)_{a \leq m}$: Gaussian inputs, random labels $\theta_a = \pm 1$
- Fit all the data:

$$\text{sign} \langle g_a, \sigma \rangle = \theta_a \iff \langle \theta_a g_a, \sigma \rangle \geq 0$$

- This is the ABP with $\kappa = 0$
- **Storage capacity**: how many random patterns can n binary weights store? Hopfield 82, Gardner 87

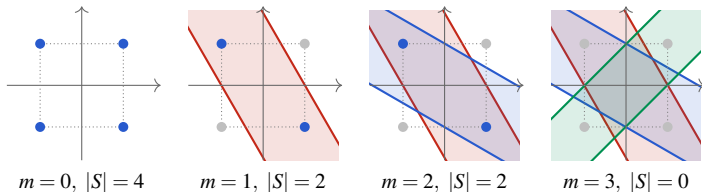
The Symmetric Binary Perceptron Model

- Start with the discrete hypercube $\{\pm 1\}^n$



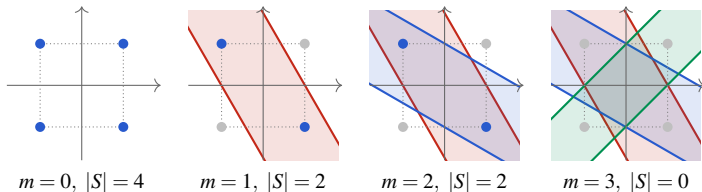
The Symmetric Binary Perceptron Model

- Start with the discrete hypercube $\{\pm 1\}^n$
- m independent random **slabs** $H_1, \dots, H_m$
 - ▶ $H_a = \{\sigma \in \{\pm 1\}^n : |\langle g_a, \sigma \rangle| \leq \kappa \sqrt{n}\}$, with $g_1, \dots, g_m$ i.i.d. $N(0, I_n)$

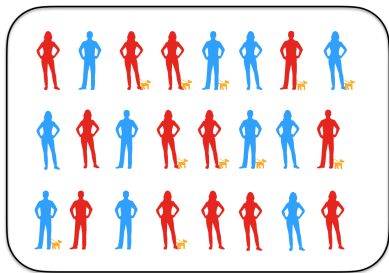


The Symmetric Binary Perceptron Model

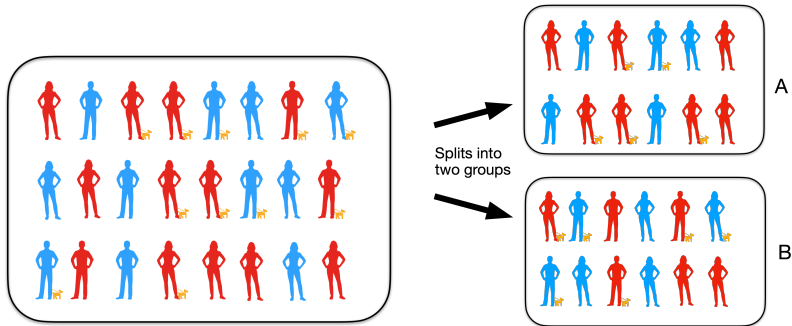
- Start with the discrete hypercube $\{\pm 1\}^n$
- m independent random **slabs** $H_1, \dots, H_m$
 - ▶ $H_a = \{\sigma \in \{\pm 1\}^n : |\langle g_a, \sigma \rangle| \leq \kappa \sqrt{n}\}$, with $g_1, \dots, g_m$ i.i.d. $N(0, I_n)$
- Define the solution set $S = \{\pm 1\}^n \cap H_1 \cap \dots \cap H_m$



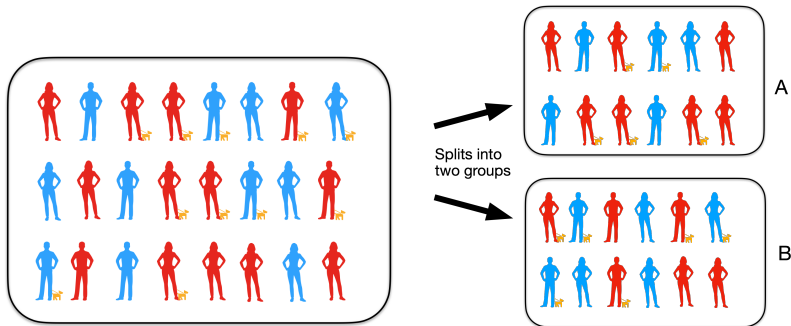
Discrepancy Minimization



Discrepancy Minimization



Discrepancy Minimization



Such that in both groups, the characteristics don't differ too much:

$$\# \text{ [Male Icon] in A} \approx \# \text{ [Male Icon] in B}$$

Gender

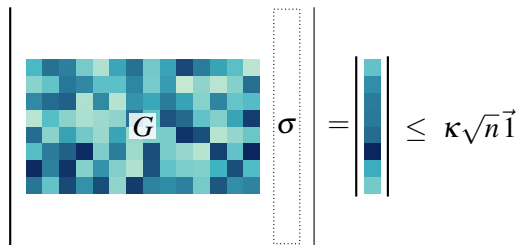
$$\# \text{ [Red Icon] in A} \approx \# \text{ [Red Icon] in B}$$

Color

$$\# \text{ [Dog Icon] in A} \approx \# \text{ [Dog Icon] in B}$$

Pet

Symmetric Perceptron—Discrepancy Minimization


$$\left| \begin{array}{c} \text{Matrix } G \\ \text{Vector } \sigma \end{array} \right| = \left| \begin{array}{c} \text{Vector of absolute values} \end{array} \right| \leq \kappa \sqrt{n} \vec{1}$$

- Fix $\kappa > 0$ Aubin Perkins Zdeborová 19

$$S = \{ \sigma \in \{\pm 1\}^n : |G\sigma| \leq \kappa \sqrt{n} \text{ for each entry} \}$$

Symmetric Perceptron—Discrepancy Minimization

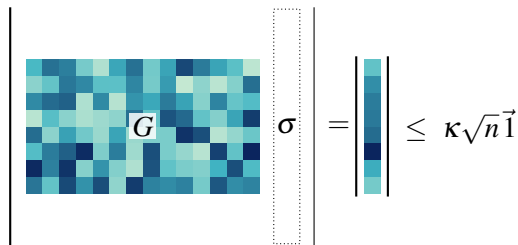
$$\left\| G \sigma \right\| = \left\| \begin{matrix} \text{Matrix } G \\ \text{Vector } \sigma \end{matrix} \right\| \leq \kappa \sqrt{n} \vec{1}$$

- Fix $\kappa > 0$ Aubin Perkins Zdeborová 19

$$S = \{ \sigma \in \{\pm 1\}^n : |G\sigma| \leq \kappa \sqrt{n} \text{ for each entry} \}$$

- The sign σ_i is the assignment: $+1 \mapsto$ group A, $-1 \mapsto$ group B

Symmetric Perceptron—Discrepancy Minimization


$$\left| G \sigma \right| = \left| \begin{matrix} \text{[Heatmap of } G \text{]} \\ \text{[Vector } \sigma \text{]} \end{matrix} \right| \leq \kappa \sqrt{n} \vec{1}$$

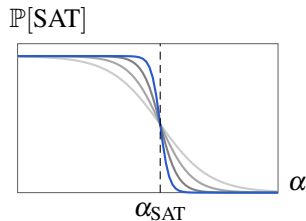
- Fix $\kappa > 0$ Aubin Perkins Zdeborová 19

$$S = \{ \sigma \in \{\pm 1\}^n : |G\sigma| \leq \kappa \sqrt{n} \text{ for each entry} \}$$

- The sign σ_i is the assignment: $+1 \mapsto$ group A, $-1 \mapsto$ group B
- Each row of G is a **feature**; we control its **imbalance** Spencer 85, Bansal 10, Lovett Meka 15

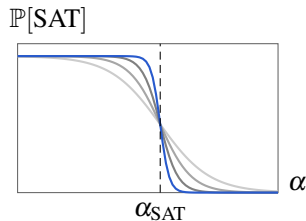
Questions for the Perceptron Model

- **Satisfiability threshold:** for which $\alpha = m/n$ are there solutions?
 - ▶ Is there a sharp transition? Where is it?



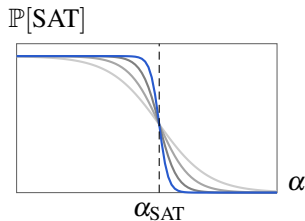
Questions for the Perceptron Model

- **Satisfiability threshold:** for which $\alpha = m/n$ are there solutions?
 - ▶ Is there a sharp transition? Where is it?
- **Structure of the solution set:**
 - ▶ Distances between solutions? Clusters?
 - ▶ What does a *typical* solution look like?



Questions for the Perceptron Model

- **Satisfiability threshold:** for which $\alpha = m/n$ are there solutions?
 - ▶ Is there a sharp transition? Where is it?
- **Structure of the solution set:**
 - ▶ Distances between solutions? Clusters?
 - ▶ What does a *typical* solution look like?
- **Algorithms:**
 - ▶ When are there polynomial-time algorithms?
 - ▶ *Which* solutions do they find?



Predictions from Physics and Rigorous Results

- Physics: **sharp transition** at $\alpha_{\text{SAT}} = \alpha_*(\kappa)$; $\alpha_*(0) \approx 0.833$ Krauth Mézard 89
- SBP: $\alpha_*(\kappa)$ given by the first moment Aubin Perkins Zdeborová 19

Predictions from Physics and Rigorous Results

- Physics: **sharp transition** at $\alpha_{\text{SAT}} = \alpha_*(\kappa)$; $\alpha_*(0) \approx 0.833$ Krauth Mézard 89
- SBP: $\alpha_*(\kappa)$ given by the first moment Aubin Perkins Zdeborová 19

Rigorous results: ABP

- $\varepsilon < \alpha_{\text{SAT}} < 1 - \varepsilon$ Kim Roche 98, Talagrand 99

Predictions from Physics and Rigorous Results

- Physics: **sharp transition** at $\alpha_{\text{SAT}} = \alpha_*(\kappa)$; $\alpha_*(0) \approx 0.833$ Krauth Mézard 89
- SBP: $\alpha_*(\kappa)$ given by the first moment Aubin Perkins Zdeborová 19

Rigorous results: ABP

- $\varepsilon < \alpha_{\text{SAT}} < 1 - \varepsilon$ Kim Roche 98, Talagrand 99
- **Sharp threshold sequence** Xu 21
- Extended to general disorder Nakajima Sun 23

Predictions from Physics and Rigorous Results

- Physics: **sharp transition** at $\alpha_{\text{SAT}} = \alpha_*(\kappa)$; $\alpha_*(0) \approx 0.833$ Krauth Mézard 89
- SBP: $\alpha_*(\kappa)$ given by the first moment Aubin Perkins Zdeborová 19

Rigorous results: ABP

- $\varepsilon < \alpha_{\text{SAT}} < 1 - \varepsilon$ Kim Roche 98, Talagrand 99
- **Sharp threshold sequence** Xu 21
- Extended to general disorder Nakajima Sun 23
- $\kappa = 0$: **lower bound** $\alpha_{\text{SAT}} \geq \alpha_*$, **second moment method*** Ding Sun 19

Predictions from Physics and Rigorous Results

- Physics: **sharp transition** at $\alpha_{\text{SAT}} = \alpha_*(\kappa)$; $\alpha_*(0) \approx 0.833$ Krauth Mézard 89
- SBP: $\alpha_*(\kappa)$ given by the first moment Aubin Perkins Zdeborová 19

Rigorous results: ABP

- $\varepsilon < \alpha_{\text{SAT}} < 1 - \varepsilon$ Kim Roche 98, Talagrand 99
- **Sharp threshold sequence** Xu 21
- Extended to general disorder Nakajima Sun 23
- $\kappa = 0$: **lower bound** $\alpha_{\text{SAT}} \geq \alpha_*$, **second moment method*** Ding Sun 19
- $\kappa = 0$: **upper bound** $\alpha_{\text{SAT}} \leq \alpha_*$, **conditional first moment method*** Huang 24

Predictions from Physics and Rigorous Results

- Physics: **sharp transition** at $\alpha_{\text{SAT}} = \alpha_*(\kappa)$; $\alpha_*(0) \approx 0.833$ Krauth Mézard 89
- SBP: $\alpha_*(\kappa)$ given by the first moment Aubin Perkins Zdeborová 19

Rigorous results: ABP

- $\varepsilon < \alpha_{\text{SAT}} < 1 - \varepsilon$ Kim Roche 98, Talagrand 99
- **Sharp threshold sequence** Xu 21
- Extended to general disorder Nakajima Sun 23
- $\kappa = 0$: **lower bound** $\alpha_{\text{SAT}} \geq \alpha_*$, **second moment method*** Ding Sun 19
- $\kappa = 0$: **upper bound** $\alpha_{\text{SAT}} \leq \alpha_*$, **conditional first moment method*** Huang 24
- *Similar **numerical conditions**, now verified Shmalo 26

Predictions from Physics and Rigorous Results

- Physics: **sharp transition** at $\alpha_{\text{SAT}} = \alpha_*(\kappa)$; $\alpha_*(0) \approx 0.833$ Krauth Mézard 89
- SBP: $\alpha_*(\kappa)$ given by the first moment Aubin Perkins Zdeborová 19

Rigorous results: ABP

- $\varepsilon < \alpha_{\text{SAT}} < 1 - \varepsilon$ Kim Roche 98, Talagrand 99
- **Sharp threshold sequence** Xu 21
- Extended to general disorder Nakajima Sun 23
- $\kappa = 0$: **lower bound** $\alpha_{\text{SAT}} \geq \alpha_*$, **second moment method*** Ding Sun 19
- $\kappa = 0$: **upper bound** $\alpha_{\text{SAT}} \leq \alpha_*$, **conditional first moment method*** Huang 24
- *Similar **numerical conditions**, now verified Shmalo 26

Rigorous results: SBP

- Model introduced; α_* from the first moment Aubin Perkins Zdeborová 19

Predictions from Physics and Rigorous Results

- Physics: **sharp transition** at $\alpha_{\text{SAT}} = \alpha_*(\kappa)$; $\alpha_*(0) \approx 0.833$ Krauth Mézard 89
- SBP: $\alpha_*(\kappa)$ given by the first moment Aubin Perkins Zdeborová 19

Rigorous results: ABP

- $\varepsilon < \alpha_{\text{SAT}} < 1 - \varepsilon$ Kim Roche 98, Talagrand 99
- **Sharp threshold sequence** Xu 21
- Extended to general disorder Nakajima Sun 23
- $\kappa = 0$: **lower bound** $\alpha_{\text{SAT}} \geq \alpha_*$, **second moment method*** Ding Sun 19
- $\kappa = 0$: **upper bound** $\alpha_{\text{SAT}} \leq \alpha_*$, **conditional first moment method*** Huang 24
- *Similar **numerical conditions**, now verified Shmalo 26

Rigorous results: SBP

- Model introduced; α_* from the first moment Aubin Perkins Zdeborová 19
- **Sharp threshold** at α_* : first and second moment methods Perkins Xu 21, Abbe Li Sly 21
- Lognormal limit of $Z/\mathbb{E}Z$ Abbe Li Sly 21

Predictions from Physics and Rigorous Results

- Physics: **sharp transition** at $\alpha_{\text{SAT}} = \alpha_*(\kappa)$; $\alpha_*(0) \approx 0.833$ Krauth Mézard 89
- SBP: $\alpha_*(\kappa)$ given by the first moment Aubin Perkins Zdeborová 19

Rigorous results: ABP

- $\varepsilon < \alpha_{\text{SAT}} < 1 - \varepsilon$ Kim Roche 98, Talagrand 99
- **Sharp threshold sequence** Xu 21
- Extended to general disorder Nakajima Sun 23
- $\kappa = 0$: **lower bound** $\alpha_{\text{SAT}} \geq \alpha_*$, **second moment method*** Ding Sun 19
- $\kappa = 0$: **upper bound** $\alpha_{\text{SAT}} \leq \alpha_*$, **conditional first moment method*** Huang 24
- *Similar **numerical conditions**, now verified Shmalo 26

Rigorous results: SBP

- Model introduced; α_* from the first moment Aubin Perkins Zdeborová 19
- **Sharp threshold** at α_* : first and second moment methods Perkins Xu 21, Abbe Li Sly 21
- Lognormal limit of $Z/\mathbb{E}Z$ Abbe Li Sly 21
- **Critical window** Altschuler 23, Sah Sawhney 23

Predictions from Physics and Rigorous Results

- Physics: **sharp transition** at $\alpha_{\text{SAT}} = \alpha_*(\kappa)$; $\alpha_*(0) \approx 0.833$ Krauth Mézard 89
- SBP: $\alpha_*(\kappa)$ given by the first moment Aubin Perkins Zdeborová 19

Rigorous results: ABP

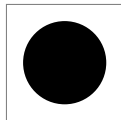
- $\varepsilon < \alpha_{\text{SAT}} < 1 - \varepsilon$ Kim Roche 98, Talagrand 99
- **Sharp threshold sequence** Xu 21
- Extended to general disorder Nakajima Sun 23
- $\kappa = 0$: **lower bound** $\alpha_{\text{SAT}} \geq \alpha_*$, **second moment method*** Ding Sun 19
- $\kappa = 0$: **upper bound** $\alpha_{\text{SAT}} \leq \alpha_*$, **conditional first moment method*** Huang 24
- *Similar **numerical conditions**, now verified Shmalo 26

Rigorous results: SBP

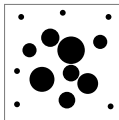
- Model introduced; α_* from the first moment Aubin Perkins Zdeborová 19
- **Sharp threshold** at α_* : first and second moment methods Perkins Xu 21, Abbe Li Sly 21
- Lognormal limit of $Z/\mathbb{E}Z$ Abbe Li Sly 21
- **Critical window** Altschuler 23, Sah Sawhney 23
- Symmetry $\Rightarrow Z \approx \mathbb{E}Z$, so moment methods work

Predictions from Physics: Strong Freezing

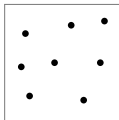
- Next: the **geometry of the solution space**
- Does the solution space have clusters?



connected



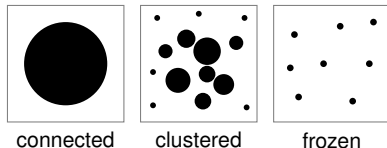
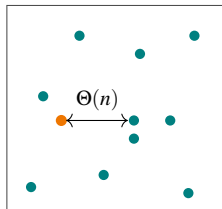
clustered



frozen

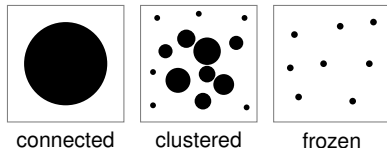
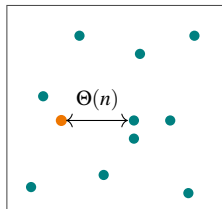
Predictions from Physics: Strong Freezing

- Next: the geometry of the solution space
- Does the solution space have clusters?
- Physicists conjecture: for **any** $0 < \alpha < \alpha_*$, a **strong freezing** phenomenon
 - ▶ Typical solutions are **isolated**
 - ▶ No other solution within Hamming distance $\Theta(n)$ Krauth Mézard 89, Huang Wong Kabashima 13, Huang Kabashima 14



Predictions from Physics: Strong Freezing

- Next: the **geometry of the solution space**
- Does the solution space have clusters?
- Physicists conjecture: for **any** $0 < \alpha < \alpha_*$, a **strong freezing** phenomenon
 - ▶ Typical solutions are **isolated**
 - ▶ No other solution within Hamming distance $\Theta(n)$ Krauth Mézard 89, Huang Wong Kabashima 13, Huang Kabashima 14
- Same prediction for the SBP Aubin Perkins Zdeborová 19



Strong Freezing

Theorem (Perkins and Xu 21; Abbe, Li, Sly 21):

For the symmetric binary perceptron, for any $\alpha < \alpha_{\text{SAT}}$, w.h.p.:

- almost all solutions are isolated;
- they are at Hamming distance $\Theta(n)$ from every other solution.

Strong Freezing

Theorem (Perkins and Xu 21; Abbe, Li, Sly 21):

For the symmetric binary perceptron, for any $\alpha < \alpha_{\text{SAT}}$, w.h.p.:

- almost all solutions are isolated;
 - they are at Hamming distance $\Theta(n)$ from every other solution.
-
- For the ABP: **open**. Both proofs use $Z \approx \mathbb{E}Z$, which fails for the ABP.

Strong Freezing

Theorem (Perkins and Xu 21; Abbe, Li, Sly 21):

For the symmetric binary perceptron, for any $\alpha < \alpha_{\text{SAT}}$, w.h.p.:

- almost all solutions are isolated;
 - they are at Hamming distance $\Theta(n)$ from every other solution.
-
- For the ABP: **open**. Both proofs use $Z \approx \mathbb{E}Z$, which fails for the ABP.
 - Why is strong freezing important?
 - ▶ Long-standing belief: **geometry** of solutions $\leftrightarrow$ **tractability** of algorithms
 - ▶ **Physics heuristic**: strong freezing $\implies$ algorithmic hardness
 - ▶ Intuition: isolated solutions are hard to find by local moves

Strong Freezing

Theorem (Perkins and Xu 21; Abbe, Li, Sly 21):

For the symmetric binary perceptron, for any $\alpha < \alpha_{\text{SAT}}$, w.h.p.:

- almost all solutions are isolated;
 - they are at Hamming distance $\Theta(n)$ from every other solution.
-
- For the ABP: **open**. Both proofs use $Z \approx \mathbb{E}Z$, which fails for the ABP.
 - Why is strong freezing important?
 - ▶ Long-standing belief: **geometry** of solutions $\leftrightarrow$ **tractability** of algorithms
 - ▶ **Physics heuristic**: strong freezing $\implies$ algorithmic hardness
 - ▶ Intuition: isolated solutions are hard to find by local moves
 - In sparse CSPs (k -SAT, coloring), freezing is partial, and only at positive density Achlioptas Coja-Oghlan 08, Zdeborová Mézard 08

But Algorithms Are Known for the Model

For the binary perceptron model:

- **Rigorous**: algorithms work at small density
 - ▶ Multiscale majority, $\alpha < 0.005$, $\kappa = 0$ Kim Roche 98
 - ▶ Discrepancy-type algorithm, $\alpha < 0.1$, $\kappa = 0$ Li Schramm Zhou 24
 - ▶ Rare clusters and an algorithm Abbe Li Sly 22

But Algorithms Are Known for the Model

For the binary perceptron model:

- **Rigorous**: algorithms work at small density
 - ▶ Multiscale majority, $\alpha < 0.005$, $\kappa = 0$ Kim Roche 98
 - ▶ Discrepancy-type algorithm, $\alpha < 0.1$, $\kappa = 0$ Li Schramm Zhou 24
 - ▶ Rare clusters and an algorithm Abbe Li Sly 22
- **Heuristic**: up to $\alpha \approx 0.69$ – 0.75 at $\kappa = 0$
 - ▶ Braunstein Zecchina 06, Baldassi Braunstein Brunel Zecchina 07, Baldassi 09, Baldassi Braunstein 15

But Algorithms Are Known for the Model

For the binary perceptron model:

- **Rigorous**: algorithms work at small density
 - ▶ Multiscale majority, $\alpha < 0.005$, $\kappa = 0$ Kim Roche 98
 - ▶ Discrepancy-type algorithm, $\alpha < 0.1$, $\kappa = 0$ Li Schramm Zhou 24
 - ▶ Rare clusters and an algorithm Abbe Li Sly 22
- **Heuristic**: up to $\alpha \approx 0.69$ – 0.75 at $\kappa = 0$
 - ▶ Braunstein Zecchina 06, Baldassi Braunstein Brunel Zecchina 07, Baldassi 09, Baldassi Braunstein 15

For the symmetric binary perceptron model:

- **Rigorous**: discrepancy minimization algorithms Bansal 10, Lovett Meka 15, Rothvoss 17, Eldan Singh 18

The Seemingly Contradictory Phenomenon

Theorem (Perkins Xu 21, Abbe Li Sly 21)

The SBP has the strong freezing property.

Conjecture (Huang Wong Kabashima 13)

The ABP has the strong freezing property.

physics
heuristic

Algorithmic Hardness

The Seemingly Contradictory Phenomenon

Theorem (Perkins Xu 21, Abbe Li Sly 21)

The SBP has the strong freezing property.

Conjecture (Huang Wong Kabashima 13)

The ABP has the strong freezing property.

physics
heuristic

Algorithmic Hardness

×

Algorithms exist
ABP and SBP

The Seemingly Contradictory Phenomenon

Theorem (Perkins Xu 21, Abbe Li Sly 21)

The SBP has the strong freezing property.

Conjecture (Huang Wong Kabashima 13)

The ABP has the strong freezing property.

physics
heuristic



Algorithmic Hardness

×

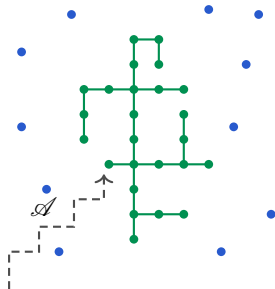


Algorithms exist
ABP and SBP

Actually, no contradiction!

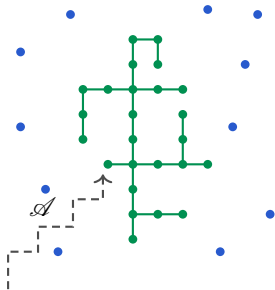
Resolving the Contradiction

- What kind of solutions do the algorithms find?



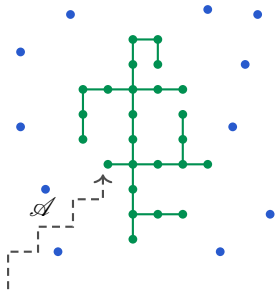
Resolving the Contradiction

- What kind of solutions do the algorithms find?
- Heuristic resolution Baldassi et al. 15, 20:
 - ▶ There are **well-connected clusters** of solutions that are **rare** (not typical)
 - ▶ Algorithms find solutions there



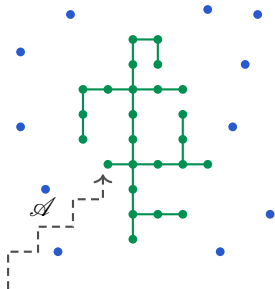
Resolving the Contradiction

- What kind of solutions do the algorithms find?
- Heuristic resolution Baldassi et al. 15, 20:
 - ▶ There are **well-connected clusters** of solutions that are **rare** (not typical)
 - ▶ Algorithms find solutions there
- Recall strong freezing: **almost all** solutions are isolated



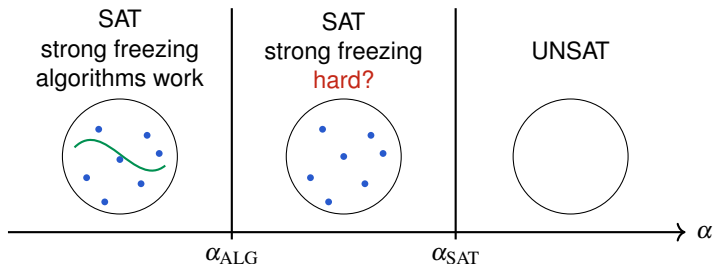
Resolving the Contradiction

- What kind of solutions do the algorithms find?
- Heuristic resolution Baldassi et al. 15, 20:
 - ▶ There are **well-connected clusters** of solutions that are **rare** (not typical)
 - ▶ Algorithms find solutions there
- Recall strong freezing: **almost all** solutions are isolated
- To justify this picture, two things must be shown:
 1. rare well-connected clusters exist:
rigorous at small α Abbe Li Sly 22;
 2. isolated solutions are hard to find.



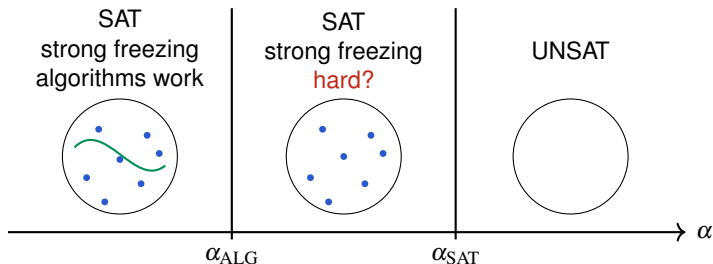
One Last Piece: Are Isolated Solutions Hard to Find?

- Sampling a uniform solution is hard for any $0 < \alpha < \alpha_{\text{SAT}}$ El Alaoui Gamarnik 25



One Last Piece: Are Isolated Solutions Hard to Find?

- Sampling a uniform solution is hard for any $0 < \alpha < \alpha_{\text{SAT}}$ El Alaoui Gamarnik 25
- They asked: can stable algorithms find **any** isolated solution? also Perkins Xu 21



Our Algorithmic Result

Theorem (G., Huang, Li, Sellke 26):

If a **stable** algorithm finds a solution with probability $1 - o(1)$, then its output is a cn -**isolated** solution only with probability $o(1)$.

Our Algorithmic Result

Theorem (G., Huang, Li, Sellke 26):

If a **stable** algorithm finds a solution with probability $1 - o(1)$, then its output is a cn -**isolated** solution only with probability $o(1)$.

- **Stable**: the output moves by $o(\sqrt{n})$ when G is slightly perturbed
- Examples: low-degree polynomials (degree $o(n/\log n)$); multiscale majority

Kim Roche 98

Our Algorithmic Result

Theorem (G., Huang, Li, Sellke 26):

If a **stable** algorithm finds a solution with probability $1 - o(1)$, then its output is a cn -**isolated** solution only with probability $o(1)$.

- **Stable**: the output moves by $o(\sqrt{n})$ when G is slightly perturbed
- Examples: low-degree polynomials (degree $o(n/\log n)$); multiscale majority
Kim Roche 98
- Successful stable algorithms **avoid** isolated solutions

The Missing Piece

G., Huang, Li, Sellke 26

successful stable algorithms output **non-isolated** solutions

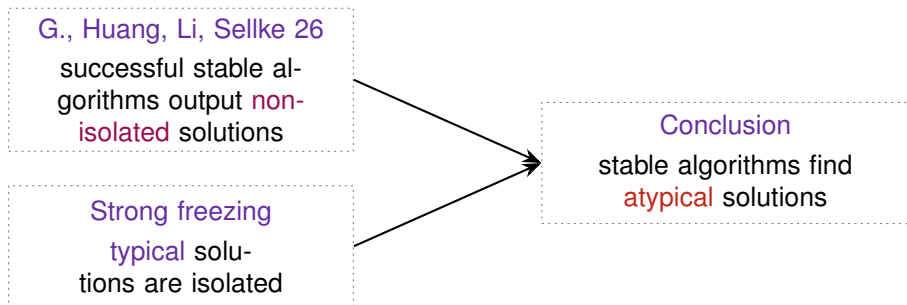
Strong freezing

typical solutions are isolated

Conclusion

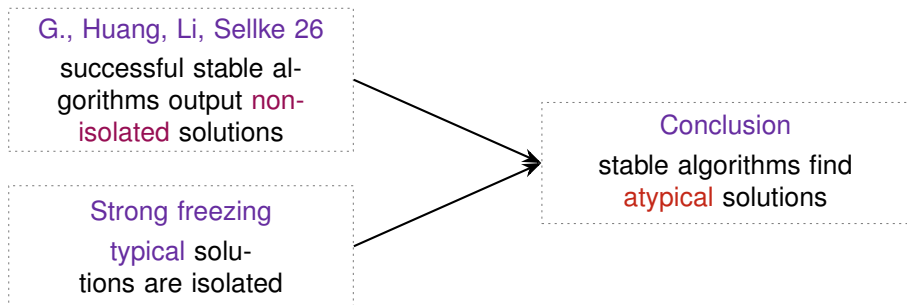
stable algorithms find **atypical** solutions

The Missing Piece



- SBP: strong freezing is known.
- ABP: strong freezing was **open**.

The Missing Piece



- SBP: strong freezing is known.
- ABP: strong freezing was **open**.
- **This talk**: strong freezing for the ABP, at every density and every margin.

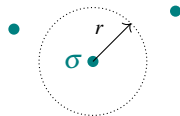
Isolation and Strong Freezing

Non-isolated solutions

Hamming distance $d_H(\sigma, \tau) = \#\{i : \sigma_i \neq \tau_i\}$. For $r \geq 1$,

$$A_r(G) = \{\sigma \in S : \exists \tau \in S, 1 \leq d_H(\sigma, \tau) \leq r\},$$

$$R_r(G) = |A_r(G)|/Z(G).$$



isolated: $\sigma \notin A_r$



not isolated: $\sigma \in A_r$

Isolation and Strong Freezing

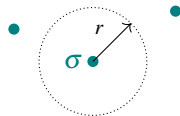
Non-isolated solutions

Hamming distance $d_H(\sigma, \tau) = \#\{i : \sigma_i \neq \tau_i\}$. For $r \geq 1$,

$$A_r(G) = \{\sigma \in S : \exists \tau \in S, 1 \leq d_H(\sigma, \tau) \leq r\},$$

$$R_r(G) = |A_r(G)|/Z(G).$$

- $R_r(G)$: the probability that a **uniformly random** solution is not isolated



isolated: $\sigma \notin A_r$



not isolated: $\sigma \in A_r$

Isolation and Strong Freezing

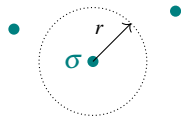
Non-isolated solutions

Hamming distance $d_H(\sigma, \tau) = \#\{i : \sigma_i \neq \tau_i\}$. For $r \geq 1$,

$$A_r(G) = \{\sigma \in S : \exists \tau \in S, 1 \leq d_H(\sigma, \tau) \leq r\},$$

$$R_r(G) = |A_r(G)|/Z(G).$$

- $R_r(G)$: the probability that a **uniformly random** solution is not isolated
- **Strong freezing**: $R_{\delta n}(G) \rightarrow 0$ for some $\delta > 0$



isolated: $\sigma \notin A_r$



not isolated: $\sigma \in A_r$

The Model in This Talk

- Entries g_{ai} i.i.d., mean 0, variance 1 (Gaussian, Rademacher, ...)

The Model in This Talk

- Entries g_{ai} i.i.d., mean 0, variance 1 (Gaussian, Rademacher, ...)
- For simplicity, take $K = [\kappa, \infty)$, the ABP:

$$S(G) = \{ \boldsymbol{\sigma} \in \{\pm 1\}^n : \langle g_a, \boldsymbol{\sigma} \rangle \geq \kappa \sqrt{n} \quad \forall a \leq m \}$$

The Model in This Talk

- Entries g_{ai} i.i.d., mean 0, variance 1 (Gaussian, Rademacher, ...)
- For simplicity, take $K = [\kappa, \infty)$, the ABP:

$$S(G) = \{ \sigma \in \{\pm 1\}^n : \langle g_a, \sigma \rangle \geq \kappa \sqrt{n} \quad \forall a \leq m \}$$

More generally

The same proof works when the constraint is $\langle g_a, \sigma \rangle / \sqrt{n} \in K$ for a finite union of intervals $K \neq \mathbb{R}$.

For example, the SBP: $K = [-\kappa, \kappa]$.

Main Result—Strong Freezing

Theorem (G., Li 26):

Let $K = [\kappa, \infty)$, or any finite union of intervals $K \neq \mathbb{R}$. For every $\alpha > 0$ there are $\delta, c > 0$ such that for large n ,

$$\mathbb{E}R_{\delta n}(G) \leq e^{-2c\sqrt{n}}, \quad \mathbb{P}(R_{\delta n}(G) > e^{-c\sqrt{n}}) \leq e^{-c\sqrt{n}}.$$

Main Result—Strong Freezing

Theorem (G., Li 26):

Let $K = [\kappa, \infty)$, or any finite union of intervals $K \neq \mathbb{R}$. For every $\alpha > 0$ there are $\delta, c > 0$ such that for large n ,

$$\mathbb{E}R_{\delta n}(G) \leq e^{-2c\sqrt{n}}, \quad \mathbb{P}(R_{\delta n}(G) > e^{-c\sqrt{n}}) \leq e^{-c\sqrt{n}}.$$

- W.h.p., almost every solution has **no other solution** within distance δn

Main Result—Strong Freezing

Theorem (G., Li 26):

Let $K = [\kappa, \infty)$, or any finite union of intervals $K \neq \mathbb{R}$. For every $\alpha > 0$ there are $\delta, c > 0$ such that for large n ,

$$\mathbb{E} R_{\delta n}(G) \leq e^{-2c\sqrt{n}}, \quad \mathbb{P}(R_{\delta n}(G) > e^{-c\sqrt{n}}) \leq e^{-c\sqrt{n}}.$$

- W.h.p., almost every solution has **no other solution** within distance δn
- For **every** $\alpha > 0$ and every margin κ ; no input on α_{SAT}

Main Result—Strong Freezing

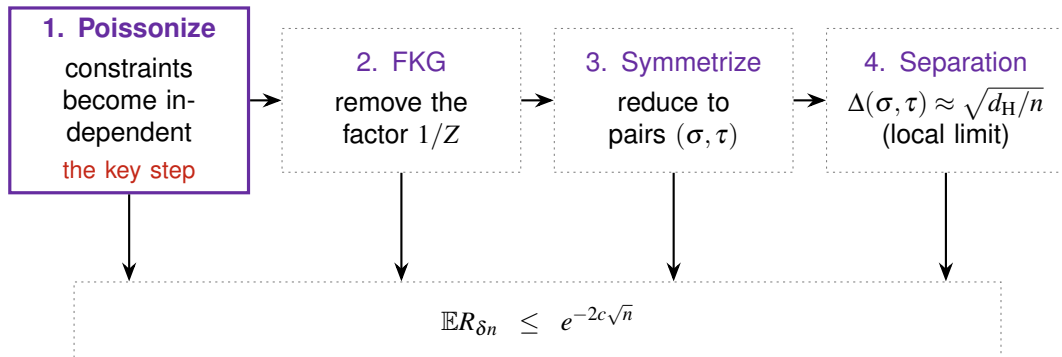
Theorem (G., Li 26):

Let $K = [\kappa, \infty)$, or any finite union of intervals $K \neq \mathbb{R}$. For every $\alpha > 0$ there are $\delta, c > 0$ such that for large n ,

$$\mathbb{E} R_{\delta n}(G) \leq e^{-2c\sqrt{n}}, \quad \mathbb{P}(R_{\delta n}(G) > e^{-c\sqrt{n}}) \leq e^{-c\sqrt{n}}.$$

- W.h.p., almost every solution has **no other solution** within distance δn
- For **every** $\alpha > 0$ and every margin κ ; no input on α_{SAT}
- ABP: confirms Krauth Mézard 89, Huang Wong Kabashima 13, Huang Kabashima 14
- SBP: recovers Perkins Xu 21, Abbe Li Sly 21

Proof Roadmap



The Obstacle: the Factor $1/Z$

A non-isolated solution has a neighbour. So

$$\mathbb{E}R_r(G) \leq \sum_{\sigma \in \{\pm 1\}^n} \sum_{\tau: 1 \leq d_H(\sigma, \tau) \leq r} \mathbb{E} \left[\frac{\mathbf{1}\{\sigma, \tau \in S\}}{Z} \right].$$

The Obstacle: the Factor $1/Z$

A non-isolated solution has a neighbour. So

$$\mathbb{E}R_r(G) \leq \sum_{\sigma \in \{\pm 1\}^n} \sum_{\tau: 1 \leq d_H(\sigma, \tau) \leq r} \mathbb{E} \left[\frac{\mathbf{1}\{\sigma, \tau \in S\}}{Z} \right].$$

- $1/Z$ is correlated with $\{\sigma \in S\}$ and $\{\tau \in S\}$

The Obstacle: the Factor $1/Z$

A non-isolated solution has a neighbour. So

$$\mathbb{E}R_r(G) \leq \sum_{\sigma \in \{\pm 1\}^n} \sum_{\tau: 1 \leq d_H(\sigma, \tau) \leq r} \mathbb{E} \left[\frac{\mathbf{1}\{\sigma, \tau \in S\}}{Z} \right].$$

- $1/Z$ is correlated with $\{\sigma \in S\}$ and $\{\tau \in S\}$
- Controlling it directly needs fluctuations of Z , not available for the ABP

The Obstacle: the Factor $1/Z$

A non-isolated solution has a neighbour. So

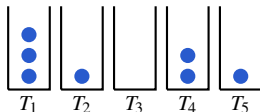
$$\mathbb{E}R_r(G) \leq \sum_{\sigma \in \{\pm 1\}^n} \sum_{\tau: 1 \leq d_H(\sigma, \tau) \leq r} \mathbb{E} \left[\frac{\mathbf{1}\{\sigma, \tau \in S\}}{Z} \right].$$

- $1/Z$ is correlated with $\{\sigma \in S\}$ and $\{\tau \in S\}$
- Controlling it directly needs fluctuations of Z , not available for the ABP
- **Idea**: make the constraints independent, then use FKG

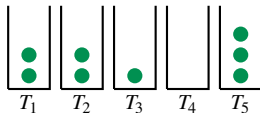
The Key Idea: Poissonization

- Constraint a cuts out a set $C_a \subseteq \{\pm 1\}^n$: a ball in the bin $T = C_a$

fixed m : counts sum to m



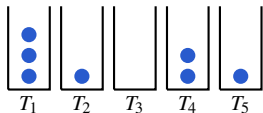
Poisson: independent counts



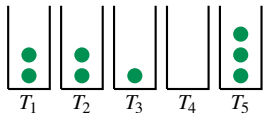
The Key Idea: Poissonization

- Constraint a cuts out a set $C_a \subseteq \{\pm 1\}^n$: a **ball** in the **bin** $T = C_a$
- Exactly m balls: bin counts are **dependent**

fixed m : counts sum to m



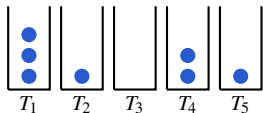
Poisson: independent counts



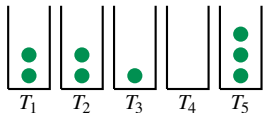
The Key Idea: Poissonization

- Constraint a cuts out a set $C_a \subseteq \{\pm 1\}^n$: a **ball** in the **bin** $T = C_a$
- Exactly m balls: bin counts are **dependent**
- $\text{Pois}(m)$ balls: bin counts are **independent**

fixed m : counts sum to m



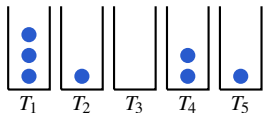
Poisson: independent counts



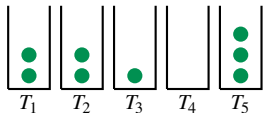
The Key Idea: Poissonization

- Constraint a cuts out a set $C_a \subseteq \{\pm 1\}^n$: a **ball** in the **bin** $T = C_a$
- Exactly m balls: bin counts are **dependent**
- $\text{Pois}(m)$ balls: bin counts are **independent**
- A new ball only **removes** solutions
- So everything is **monotone** in independent inputs $\Rightarrow$ **FKG**

fixed m : counts sum to m



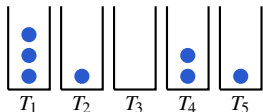
Poisson: independent counts



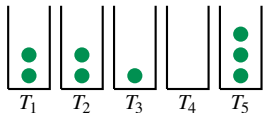
The Key Idea: Poissonization

- Constraint a cuts out a set $C_a \subseteq \{\pm 1\}^n$: a **ball** in the **bin** $T = C_a$
- Exactly m balls: bin counts are **dependent**
- $\text{Pois}(m)$ balls: bin counts are **independent**
- A new ball only **removes** solutions
- So everything is **monotone** in independent inputs $\Rightarrow$ **FKG**
- Back to exactly m balls: costs a factor $\leq 3\sqrt{m}$, which is negligible

fixed m : counts sum to m



Poisson: independent counts



Step 1: Poissonization

- $C_a = \{\sigma : \langle g_a, \sigma \rangle / \sqrt{n} \in K\}$, so $S = \bigcap_{a \leq m} C_a$

Step 1: Poissonization

- $C_a = \{\sigma : \langle g_a, \sigma \rangle / \sqrt{n} \in K\}$, so $S = \bigcap_{a \leq m} C_a$
- Replace m by $N(\lambda) \sim \text{Pois}(\lambda)$

Step 1: Poissonization

- $C_a = \{\sigma : \langle g_a, \sigma \rangle / \sqrt{n} \in K\}$, so $S = \bigcap_{a \leq m} C_a$
- Replace m by $N(\lambda) \sim \text{Pois}(\lambda)$
- **Poisson colouring**: for $T \subseteq \{\pm 1\}^n$, let $\text{Count}(T) = \#\{a : C_a = T\}$
- These counts are **independent** Poisson variables

Step 1: Poissonization

- $C_a = \{\sigma : \langle g_a, \sigma \rangle / \sqrt{n} \in K\}$, so $S = \bigcap_{a \leq m} C_a$
- Replace m by $N(\lambda) \sim \text{Pois}(\lambda)$
- **Poisson colouring**: for $T \subseteq \{\pm 1\}^n$, let $\text{Count}(T) = \#\{a : C_a = T\}$
- These counts are **independent** Poisson variables
- Conditioning on $N(\lambda) = m$ gives back the original model

Step 1: What Poissonization Gives

- $\sigma \in S$ iff $\text{Count}(T) = 0$ for all $T \not\supset \sigma$. So

$$\mathbb{P}(\sigma \in S) = \exp(-\lambda \mathbb{P}(\sigma \notin C_1))$$

Step 1: What Poissonization Gives

- $\sigma \in S$ iff $\text{Count}(T) = 0$ for all $T \not\ni \sigma$. So

$$\mathbb{P}(\sigma \in S) = \exp(-\lambda \mathbb{P}(\sigma \notin C_1))$$

- Given $\sigma \in S$, the event $\{\tau \in S\}$ depends on independent indicators $\{\text{Count}(T) \geq 1\}$, $T \ni \sigma$

Step 1: What Poissonization Gives

- $\sigma \in S$ iff $\text{Count}(T) = 0$ for all $T \not\ni \sigma$. So

$$\mathbb{P}(\sigma \in S) = \exp(-\lambda \mathbb{P}(\sigma \notin C_1))$$

- Given $\sigma \in S$, the event $\{\tau \in S\}$ depends on independent indicators $\{\text{Count}(T) \geq 1\}$, $T \ni \sigma$
- And

$$\mathbb{P}(\tau \in S \mid \sigma \in S) = \exp(-\lambda \mathbb{P}(\sigma \in C_1, \tau \notin C_1))$$

Step 2: FKG Decoupling

Claim:

$$\mathbb{E}\left[\frac{\mathbf{1}\{\sigma, \tau \in S\}}{Z}\right] \leq \underbrace{\exp(-\lambda \mathbb{P}(\sigma \in C_1, \tau \notin C_1))}_{= \mathbb{P}(\tau \in S \mid \sigma \in S)} \cdot \mathbb{E}\left[\frac{\mathbf{1}\{\sigma \in S\}}{Z}\right]$$

Step 2: FKG Decoupling

Claim:

$$\mathbb{E}\left[\frac{\mathbf{1}\{\sigma, \tau \in S\}}{Z}\right] \leq \underbrace{\exp(-\lambda \mathbb{P}(\sigma \in C_1, \tau \notin C_1))}_{= \mathbb{P}(\tau \in S \mid \sigma \in S)} \cdot \mathbb{E}\left[\frac{\mathbf{1}\{\sigma \in S\}}{Z}\right]$$

- Turning on an indicator only **removes** solutions

Step 2: FKG Decoupling

Claim:

$$\mathbb{E}\left[\frac{\mathbf{1}\{\sigma, \tau \in S\}}{Z}\right] \leq \underbrace{\exp(-\lambda \mathbb{P}(\sigma \in C_1, \tau \notin C_1))}_{= \mathbb{P}(\tau \in S \mid \sigma \in S)} \cdot \mathbb{E}\left[\frac{\mathbf{1}\{\sigma \in S\}}{Z}\right]$$

- Turning on an indicator only **removes** solutions
- So $\mathbf{1}\{\tau \in S\}$ is decreasing, and $1/Z$ is increasing

Step 2: FKG Decoupling

Claim:

$$\mathbb{E}\left[\frac{\mathbf{1}\{\sigma, \tau \in S\}}{Z}\right] \leq \underbrace{\exp(-\lambda \mathbb{P}(\sigma \in C_1, \tau \notin C_1))}_{= \mathbb{P}(\tau \in S \mid \sigma \in S)} \cdot \mathbb{E}\left[\frac{\mathbf{1}\{\sigma \in S\}}{Z}\right]$$

- Turning on an indicator only **removes** solutions
- So $\mathbf{1}\{\tau \in S\}$ is decreasing, and $1/Z$ is increasing
- **Harris–FKG**: oppositely monotone functions are negatively correlated,
 $\mathbb{E}[fh] \leq \mathbb{E}f \cdot \mathbb{E}h$

Step 3: Symmetrization

- Let $\Delta(\sigma, \tau) = \mathbb{P}(\mathbf{1}\{\sigma \in C_1\} \neq \mathbf{1}\{\tau \in C_1\})$

Step 3: Symmetrization

- Let $\Delta(\sigma, \tau) = \mathbb{P}(\mathbf{1}\{\sigma \in C_1\} \neq \mathbf{1}\{\tau \in C_1\})$
- Use Step 2 in both orders. The exponents add up to $\lambda\Delta$:

$$\mathbb{E}\left[\frac{\mathbf{1}\{\sigma, \tau \in S\}}{Z}\right] \leq \frac{1}{2} e^{-\lambda\Delta(\sigma, \tau)/2} \left(\mathbb{E}\left[\frac{\mathbf{1}\{\sigma \in S\}}{Z}\right] + \mathbb{E}\left[\frac{\mathbf{1}\{\tau \in S\}}{Z}\right] \right)$$

Step 3: Symmetrization

- Let $\Delta(\sigma, \tau) = \mathbb{P}(\mathbf{1}\{\sigma \in C_1\} \neq \mathbf{1}\{\tau \in C_1\})$
- Use Step 2 in both orders. The exponents add up to $\lambda\Delta$:

$$\mathbb{E}\left[\frac{\mathbf{1}\{\sigma, \tau \in S\}}{Z}\right] \leq \frac{1}{2} e^{-\lambda\Delta(\sigma, \tau)/2} \left(\mathbb{E}\left[\frac{\mathbf{1}\{\sigma \in S\}}{Z}\right] + \mathbb{E}\left[\frac{\mathbf{1}\{\tau \in S\}}{Z}\right] \right)$$

- Sum over pairs, using $\sum_{\sigma} \mathbb{E}[\mathbf{1}\{\sigma \in S\}/Z] = \mathbb{P}(Z > 0) \leq 1$

Step 3: Symmetrization

- Let $\Delta(\sigma, \tau) = \mathbb{P}(\mathbf{1}\{\sigma \in C_1\} \neq \mathbf{1}\{\tau \in C_1\})$
- Use Step 2 in both orders. The exponents add up to $\lambda\Delta$:

$$\mathbb{E}\left[\frac{\mathbf{1}\{\sigma, \tau \in S\}}{Z}\right] \leq \frac{1}{2} e^{-\lambda\Delta(\sigma, \tau)/2} \left(\mathbb{E}\left[\frac{\mathbf{1}\{\sigma \in S\}}{Z}\right] + \mathbb{E}\left[\frac{\mathbf{1}\{\tau \in S\}}{Z}\right] \right)$$

- Sum over pairs, using $\sum_{\sigma} \mathbb{E}[\mathbf{1}\{\sigma \in S\}/Z] = \mathbb{P}(Z > 0) \leq 1$

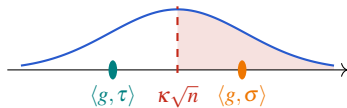
Proposition:

With $\lambda = m$, $\mathbb{E}R_r(G) \leq \underbrace{3\sqrt{m}}_{\text{Poissonization}} \max_{\sigma} \sum_{\tau: 1 \leq d_H(\sigma, \tau) \leq r} e^{-m\Delta(\sigma, \tau)/2}.$

The factor $3\sqrt{m} \geq 1/\mathbb{P}(N(m) = m)$ is the price of going back from $\text{Pois}(m)$ to exactly m constraints.

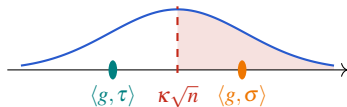
Step 4: Separation

- $\Delta(\sigma, \tau)$: chance that **one** constraint separates σ and τ



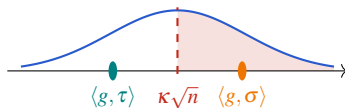
Step 4: Separation

- $\Delta(\sigma, \tau)$: chance that **one** constraint separates σ and τ
- $d_H(\sigma, \tau) = k$: $\langle g, \sigma \rangle$ and $\langle g, \tau \rangle$ differ by $\approx \sqrt{k}$



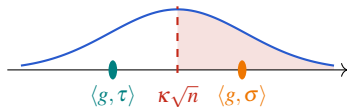
Step 4: Separation

- $\Delta(\sigma, \tau)$: chance that **one** constraint separates σ and τ
- $d_H(\sigma, \tau) = k$: $\langle g, \sigma \rangle$ and $\langle g, \tau \rangle$ differ by $\approx \sqrt{k}$
- Separated iff $\kappa\sqrt{n}$ lies **between** them



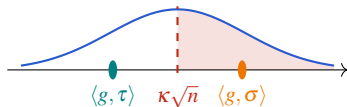
Step 4: Separation

- $\Delta(\sigma, \tau)$: chance that **one** constraint separates σ and τ
- $d_H(\sigma, \tau) = k$: $\langle g, \sigma \rangle$ and $\langle g, \tau \rangle$ differ by $\approx \sqrt{k}$
- Separated iff $\kappa\sqrt{n}$ lies **between** them
- Density $\approx 1/\sqrt{n}$ there, by the **local limit theorem**, so $\Delta(\sigma, \tau) \approx \sqrt{k/n}$



Step 4: Separation

- $\Delta(\sigma, \tau)$: chance that **one** constraint separates σ and τ
- $d_H(\sigma, \tau) = k$: $\langle g, \sigma \rangle$ and $\langle g, \tau \rangle$ differ by $\approx \sqrt{k}$
- Separated iff $\kappa\sqrt{n}$ lies **between** them
- Density $\approx 1/\sqrt{n}$ there, by the **local limit theorem**, so $\Delta(\sigma, \tau) \approx \sqrt{k/n}$



Lemma:

$\Delta(\sigma, \tau) \geq c\sqrt{d_H(\sigma, \tau)/n}$ whenever $1 \leq d_H(\sigma, \tau) \leq \delta_0 n$.

Completing the Proof

Summing over the distance $k = d_H(\sigma, \tau)$:

$$\mathbb{E}R_{\delta n}(G) \leq 3\sqrt{m} \sum_{k=1}^{\delta n} \underbrace{\binom{n}{k}}_{\text{cost}} \underbrace{e^{-c\sqrt{nk}}}_{\text{gain}}$$

$c, c', C > 0$ denote constants depending only on α and κ .

Completing the Proof

Summing over the distance $k = d_H(\sigma, \tau)$:

$$\mathbb{E}R_{\delta n}(G) \leq 3\sqrt{m} \sum_{k=1}^{\delta n} \underbrace{\binom{n}{k}}_{\text{cost}} \underbrace{e^{-c\sqrt{nk}}}_{\text{gain}}$$

- **Gain:** each τ at distance k gives $e^{-m\Delta(\sigma, \tau)/2} \leq e^{-c\sqrt{nk}}$, since $m\Delta(\sigma, \tau) \gtrsim \sqrt{nk}$

$c, c', C > 0$ denote constants depending only on α and κ .

Completing the Proof

Summing over the distance $k = d_H(\sigma, \tau)$:

$$\mathbb{E}R_{\delta n}(G) \leq 3\sqrt{m} \sum_{k=1}^{\delta n} \underbrace{\binom{n}{k}}_{\text{cost}} \underbrace{e^{-c\sqrt{nk}}}_{\text{gain}}$$

- **Gain**: each τ at distance k gives $e^{-m\Delta(\sigma, \tau)/2} \leq e^{-c\sqrt{nk}}$, since $m\Delta(\sigma, \tau) \gtrsim \sqrt{nk}$
- **Cost**: the number of τ at distance k is $\binom{n}{k} \leq e^{k \log(en/k)}$

$c, c', C > 0$ denote constants depending only on α and κ .

Completing the Proof

Summing over the distance $k = d_H(\sigma, \tau)$:

$$\mathbb{E}R_{\delta n}(G) \leq 3\sqrt{m} \sum_{k=1}^{\delta n} \underbrace{\binom{n}{k}}_{\text{cost}} \underbrace{e^{-c\sqrt{nk}}}_{\text{gain}}$$

- **Gain**: each τ at distance k gives $e^{-m\Delta(\sigma, \tau)/2} \leq e^{-c\sqrt{nk}}$, since $m\Delta(\sigma, \tau) \gtrsim \sqrt{nk}$
- **Cost**: the number of τ at distance k is $\binom{n}{k} \leq e^{k \log(en/k)}$
- Gain beats cost for $k \leq \delta n$: $\mathbb{E}R_{\delta n}(G) \leq e^{-2c'\sqrt{n}}$, and by Markov

$$\mathbb{P}(R_{\delta n}(G) > e^{-c'\sqrt{n}}) \leq e^{-c'\sqrt{n}} \quad (\text{the theorem}) \blacksquare$$

$c, c', C > 0$ denote constants depending only on α and κ .

Completing the Proof

Summing over the distance $k = d_H(\sigma, \tau)$:

$$\mathbb{E}R_{\delta n}(G) \leq 3\sqrt{m} \sum_{k=1}^{\delta n} \underbrace{\binom{n}{k}}_{\text{cost}} \underbrace{e^{-c\sqrt{nk}}}_{\text{gain}}$$

- **Gain**: each τ at distance k gives $e^{-m\Delta(\sigma, \tau)/2} \leq e^{-c\sqrt{nk}}$, since $m\Delta(\sigma, \tau) \gtrsim \sqrt{nk}$
- **Cost**: the number of τ at distance k is $\binom{n}{k} \leq e^{k \log(en/k)}$
- Gain beats cost for $k \leq \delta n$: $\mathbb{E}R_{\delta n}(G) \leq e^{-2c'\sqrt{n}}$, and by Markov

$$\mathbb{P}(R_{\delta n}(G) > e^{-c'\sqrt{n}}) \leq e^{-c'\sqrt{n}} \quad (\text{the theorem}) \blacksquare$$

- **Why $e^{-\sqrt{n}}$ is sharp**: given $\sigma \in S$, a neighbour of σ at distance 1 is also a solution with probability $\geq e^{-C\sqrt{n}}$

$c, c', C > 0$ denote constants depending only on α and κ .

Summary

- The **binary perceptron**: a random CSP connected to **neural networks**, **statistical physics**, and **discrepancy minimization**

Summary

- The **binary perceptron**: a random CSP connected to **neural networks**, **statistical physics**, and **discrepancy minimization**
- It breaks the physics intuition: **strong freezing** holds, yet algorithms find solutions

Summary

- The **binary perceptron**: a random CSP connected to **neural networks**, **statistical physics**, and **discrepancy minimization**
- It breaks the physics intuition: **strong freezing** holds, yet algorithms find solutions
- We prove **strong freezing** at **every** density, for the ABP at every margin, and for the SBP

Summary

- The **binary perceptron**: a random CSP connected to **neural networks**, **statistical physics**, and **discrepancy minimization**
- It breaks the physics intuition: **strong freezing** holds, yet algorithms find solutions
- We prove **strong freezing** at **every** density, for the ABP at every margin, and for the SBP
- With G., Huang, Li, Sellke 26: successful stable algorithms find **atypical** solutions

Summary

- The **binary perceptron**: a random CSP connected to **neural networks**, **statistical physics**, and **discrepancy minimization**
- It breaks the physics intuition: **strong freezing** holds, yet algorithms find solutions
- We prove **strong freezing** at **every** density, for the ABP at every margin, and for the SBP
- With G., Huang, Li, Sellke 26: successful stable algorithms find **atypical** solutions
- Proof: Poissonize, FKG, symmetrize, CLT

Future Directions

- **Other models:** smooth activations, the spherical perceptron. What should “isolated” mean?

Future Directions

- **Other models:** smooth activations, the spherical perceptron. What should “isolated” mean?
- **Other random CSPs:** can Poissonization and FKG help study the geometry of solutions there?

Thank you!