Recent advances on mean-field spin glasses

Wei-Kuo Chen
University of Minnesota

Joint work with
A. Auffinger, M. Handschy, D. Gamarnik, G. Lerman, D. Panchenko, M. Rahman, A. Sen, Q. Zeng

July, 2018

What are spin glasses?

What are spin glasses?

- Spin Glasses are alloys with strange magnetic properties. Ex: CuMn
  - In physics: spin + glass
  - In mathematics: quenched disorder + frustration

What are spin glasses?

- Spin Glasses are alloys with strange magnetic properties. Ex: CuMn
  - In physics: spin + glass
  - In mathematics: quenched disorder + frustration
- Spin glass features appear in many real world problems:
  - Traveling salesman problem.
  - Hopfield neural network.
  - Spike detection and recovery problems.

Edwards-Anderson model

- Consider a finite graph (V, E) on $\mathbb{Z}^d$.
- Hamiltonian: For $\sigma \in \{-1, 1\}^V$,

$$H(\sigma) = \sum_{(i,j) \in E} g_{ij} \sigma_i \sigma_j,$$

  where g_{ij} are i.i.d. $N(0, 1)$.

Edwards-Anderson model

- Consider a finite graph (V, E) on $\mathbb{Z}^d$.
- Hamiltonian: For $\sigma \in \{-1, 1\}^V$,

$$H(\sigma) = \sum_{(i,j)\in E} g_{ij}\sigma_i\sigma_j,$$

  where g_{ij} are i.i.d. $N(0, 1)$.
- Frustration appears when computing $\max H_N(\sigma)$.

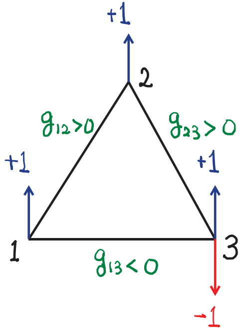


Figure: Frustration

Mean field approach: The Sherrington-Kirkpatrick model

- Hamiltonian:

$$H_N(\sigma) = \frac{1}{\sqrt{N}} \sum_{i,j=1}^{N} g_{ij}\sigma_i\sigma_j + h \sum_{i=1}^{N} \sigma_i$$

for $\sigma \in \{-1, +1\}^N$, where $g_{ij} \overset{i.i.d.}{\sim} N(0, 1)$.

Mean field approach: The Sherrington-Kirkpatrick model

- Hamiltonian:

$$H_N(\sigma) = \frac{1}{\sqrt{N}} \sum_{i,j=1}^{N} g_{ij}\sigma_i\sigma_j + h \sum_{i=1}^{N} \sigma_i$$

for $\sigma \in \{-1,+1\}^N$, where $g_{ij} \overset{i.i.d.}{\sim} N(0,1)$.

- Covariance Structure:

$$\mathbb{E}\Big(\frac{1}{\sqrt{N}} \sum_{i,j=1}^{N} g_{ij}\sigma_i^1\sigma_j^1\Big)\Big(\frac{1}{\sqrt{N}} \sum_{i,j=1}^{N} g_{ij}\sigma_i^2\sigma_j^2\Big) = N\big(R(\sigma^1,\sigma^2)\big)^2,$$

where

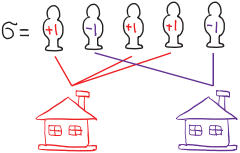
$$R(\sigma^1,\sigma^2) = \frac{1}{N}\sum_{i=1}^{N} \sigma_i^1\sigma_i^2.$$

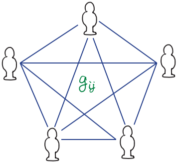
Dean's problem

Assign N students into two dorms and avoid conflicts.

Dean’s problem

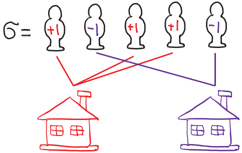
Assign N students into two dorms and avoid conflicts.

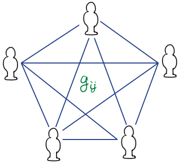




Dean's problem

Assign N students into two dorms and avoid conflicts.



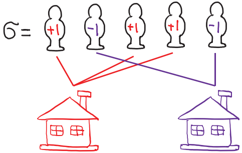


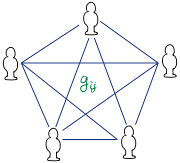
σ_i	σ_j	g_{ij}	$g_{ij}\sigma_i\sigma_j$
+1	+1	>0	>0
		<0	<0
+1	-1	>0	<0
		<0	>0
-1	-1	>0	>0
		<0	<0

Peace

Dean's problem

Assign N students into two dorms and avoid conflicts.





σ_i	σ_j	g_{ij}	$g_{ij}\sigma_i\sigma_j$	
+1	+1	>0	>0	Peace
		<0	<0	
+1	-1	>0	<0	
		<0	>0	Peace
-1	-1	>0	>0	Peace
		<0	<0	

Dean's problem: Find the optimizer of

$$\max_{\sigma \in \{-1,+1\}^N} \sum_{i,j=1}^{N} g_{ij}\sigma_i\sigma_j.$$

A soft approximation: Free energy

- For any $\beta = \frac{1}{T} > 0$ (inverse temperature), define the free energy

$$F_N(\beta) = \frac{1}{\beta N} \log \sum_{\sigma \in \{-1,+1\}^N} e^{\beta H_N(\sigma)}$$

A soft approximation: Free energy

- For any $\beta = \frac{1}{T} > 0$ (inverse temperature), define the free energy

$$F_N(\beta) = \frac{1}{\beta N} \log \sum_{\sigma \in \{-1,+1\}^N} e^{\beta H_N(\sigma)}$$

- Simple observation:

$$\max_{\sigma \in \{-1,+1\}^N} \frac{H_N(\sigma)}{N} \leq F_N(\beta) \leq \max_{\sigma \in \{-1,+1\}^N} \frac{H_N(\sigma)}{N} + \frac{\log 2}{\beta}$$

A soft approximation: Free energy

- For any $\beta = \frac{1}{T} > 0$ (inverse temperature), define the free energy

$$F_N(\beta) = \frac{1}{\beta N} \log \sum_{\sigma \in \{-1,+1\}^N} e^{\beta H_N(\sigma)}$$

- Simple observation:

$$\max_{\sigma \in \{-1,+1\}^N} \frac{H_N(\sigma)}{N} \leq F_N(\beta) \leq \max_{\sigma \in \{-1,+1\}^N} \frac{H_N(\sigma)}{N} + \frac{\log 2}{\beta}$$

- Physicists' replica method:

$$\lim_{N\to\infty} \frac{1}{N} \mathbb{E} \log Z_N = \lim_{N\to\infty} \lim_{n\downarrow 0} \frac{\mathbb{E} \log Z_N^n}{nN} \overset{?}{=} \lim_{n\downarrow 0} \lim_{N\to\infty} \frac{\log \mathbb{E} Z_N^n}{nN}$$

Theorem (Parisi formula)

- (Talagrand '06)

$$\lim_{N\to\infty} F_N(\beta) = \inf_{\alpha}\Big(\Phi_{\alpha,\beta}(0,h) - \frac{1}{2}\int_0^1 \beta\alpha(s)sds\Big),\ a.s.,$$

where for any CDF α *on* $[0,1]$,

$$\partial_s \Phi_{\alpha,\beta} = -\frac{1}{2}\Big(\partial_{xx}\Phi_{\alpha,\beta} + \beta\alpha(s)(\partial_x\Phi_{\alpha,\beta})^2\Big), \forall (s,x) \in [0,1) \times \mathbb{R}$$

with

$$\Phi_{\alpha,\beta}(1,x) = \frac{1}{\beta}\log\cosh(\beta x).$$

- (Guerra' 03) *Minimizer exists.*
- (Auffinger-C. '14) *Minimizer is unique.*

Denote this minimizer by α_β and call it the Parisi measure.

Significance of the Parisi measure

Three major predictions:

(1) α_β is the limiting distribution of the overlap:

$$R(\sigma^1,\sigma^2) \overset{d}{\Rightarrow} \alpha_\beta,$$

where σ^1,σ^2 are i.i.d. samplings from the Gibbs measure

$$G_N(\sigma) = \frac{e^{\beta H_N(\sigma)}}{\sum_{\sigma'} e^{\beta H_N(\sigma')}}.$$

(2) Phase Transition:

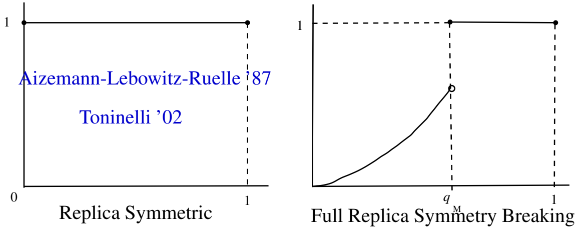
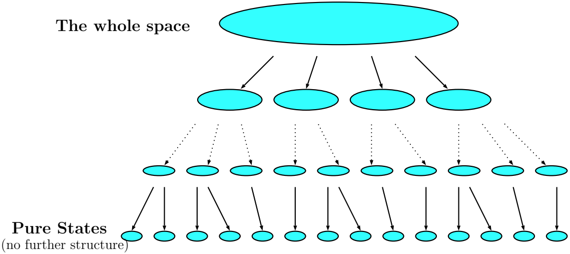


Figure: SK model with $h = 0$

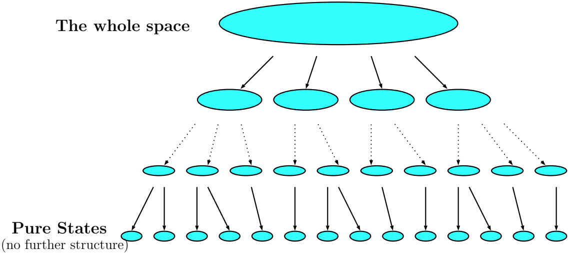
(3) Ultrametricity: with probab.≈ 1, for i.i.d. $\sigma^1, \sigma^2, \sigma^3 \sim G_N$,

$$\|\sigma^1 - \sigma^2\| \leq \max\big(\|\sigma^1 - \sigma^3\|, \|\sigma^2 - \sigma^3\|\big) + o(1).$$



(3) Ultrametricity: with probab.≈ 1, for i.i.d. $\sigma^1, \sigma^2, \sigma^3 \sim G_N$,

$$\|\sigma^1 - \sigma^2\| \leq \max\big(\|\sigma^1 - \sigma^3\|, \|\sigma^2 - \sigma^3\|\big) + o(1).$$



Panchenko '11: Ultrametricity holds for the SK model with a vanishing perturbation, but we do not know if it is still true without perturbation.

Theorem (Auffinger-C.-Zeng '17)

The cardinality of supp α_β *diverges as* $\beta \to \infty$.

As a consequence: If we add perturbation so that ultrametricity holds, then the total levels of the trees diverge as $\beta \uparrow \infty$.

Parisi formula for the maximal energy

Parisi formula for the maximal energy

For any γ with $\gamma(s) = \mu([0,s])$ and $\int_0^1 \gamma(s)ds < \infty$, consider the PDE solution Ψ_γ,

$$\Psi_\gamma(1,x) = |x|,$$
$$\partial_s \Psi_\gamma = -\frac{1}{2}\Big(\partial_{xx}\Psi_\gamma + \gamma(s)(\partial_x \Psi_\gamma)^2\Big), \forall (s,x) \in [0,1) \times \mathbb{R}.$$

Parisi formula for the maximal energy

For any γ with $\gamma(s) = \mu([0,s])$ and $\int_0^1 \gamma(s)ds < \infty$, consider the PDE solution Ψ_γ,

$$\Psi_\gamma(1,x) = |x|,$$
$$\partial_s \Psi_\gamma = -\frac{1}{2}\Big(\partial_{xx}\Psi_\gamma + \gamma(s)(\partial_x \Psi_\gamma)^2\Big), \forall (s,x) \in [0,1) \times \mathbb{R}.$$

Theorem

- *(Auffinger-C. '16) Parisi formula at zero temperature:*

$$\lim_{N\to\infty} \mathbb{E} \max_{\sigma\in\{-1,+1\}^N} \frac{H_N(\sigma)}{N} = \inf_\gamma \Big(\Psi_\gamma(0,h) - \frac{1}{2}\int_0^1 s\gamma(s)ds\Big)$$

- *(C.-Handschy-Lerman '16) Minimizer γ_P exists and is unique.*

Energy landscape: multiple peaks

Overlap $R(\sigma, \sigma') = \frac{1}{N} \sum_{i=1}^{N} \sigma_i \sigma'_i$.

Energy landscape: multiple peaks

Overlap $R(\sigma, \sigma') = \frac{1}{N} \sum_{i=1}^{N} \sigma_i \sigma'_i$.

Theorem (Multiple peaks, C.-Handschy-Lerman '16)

Assume $h = 0$. For any $\varepsilon > 0$, there exists a constant $K > 0$ s.t. for any $N \geq 1$, with probability at least $1 - Ke^{-N/K}$, $\exists S_N \subset \{-1, +1\}^N$ such that

(i) $|S_N| \geq e^{N/K}$.

(ii) $\forall \sigma \in S_N$, $\left| \frac{H_N(\sigma)}{N} - \max_{\sigma' \in \Sigma_N} \frac{H_N(\sigma')}{N} \right| < \varepsilon$.

(iii) $\forall \sigma, \sigma' \in S_N$ *with* $\sigma \neq \sigma'$, $|R(\sigma, \sigma')| < \varepsilon$.

Energy landscape: multiple peaks

Overlap $R(\sigma, \sigma') = \frac{1}{N} \sum_{i=1}^{N} \sigma_i \sigma_i'$.

Theorem (Multiple peaks, C.-Handschy-Lerman '16)

Assume $h = 0$. For any $\varepsilon > 0$, there exists a constant $K > 0$ s.t. for any $N \geq 1$, with probability at least $1 - Ke^{-N/K}$, $\exists S_N \subset \{-1, +1\}^N$ such that

(i) $|S_N| \geq e^{N/K}$.

(ii) $\forall \sigma \in S_N$, $\left| \frac{H_N(\sigma)}{N} - \max_{\sigma' \in \Sigma_N} \frac{H_N(\sigma')}{N} \right| < \varepsilon$.

(iii) $\forall \sigma, \sigma' \in S_N$ *with* $\sigma \neq \sigma'$, $|R(\sigma, \sigma')| < \varepsilon$.

- Chatterjee '09: $|S_N| \geq (\log N)^c$.
- Ding-Eldan-Zhai '14: $|S_N| \geq N^c$.

Pure p-spin model for $p \geq 3$: Overlap gap property

- Hamiltonian:

$$H_N(\sigma) = \frac{1}{N^{(p-1)/2}} \sum_{1 \leq i_1, \ldots, i_p \leq N} g_{i_1, \ldots, i_p} \sigma_{i_1} \cdots \sigma_{i_p}.$$

- (Overlap gap property) There exist $c, C > 0$ such that with overwhelming probability, any two near ground states σ^1 and σ^2 satisfy

$$|R(\sigma^1, \sigma^2)| \notin [c, C].$$

- Computational hardness:

- Computational hardness:
  - Suppose σ^1 is a near ground state and σ^2 is the ground state so that
$$|R(\sigma^1, \sigma^2)| \leq c.$$

- Computational hardness:
  - Suppose σ^1 is a near ground state and σ^2 is the ground state so that
$$|R(\sigma^1, \sigma^2)| \leq c.$$
  - Locally update algorithms take exponential time to find the ground state since
$$|R(\sigma^1, \sigma(n))| \notin [c, C].$$

- Computational hardness:
  - Suppose σ^1 is a near ground state and σ^2 is the ground state so that
$$|R(\sigma^1, \sigma^2)| \leq c.$$
  - Locally update algorithms take exponential time to find the ground state since
$$|R(\sigma^1, \sigma(n))| \notin [c, C].$$
- Results:
  - C.-Gamarnik-Rahman-Panchenko '17
  - Jagannath-Ben Arous '17

New challenges

Bipartite SK model: Let $N_1 = cN$ and $N_2 = (1-c)N$.

$$H_N(\sigma) = \frac{1}{\sqrt{N}} \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} g_{ij} \tau_i \rho_j$$

for $\sigma = (\tau, \rho) \in \{-1,+1\}^{N_1} \times \{-1,+1\}^{N_2}$. Note

$$\mathbb{E} H_N(\sigma) H_N(\sigma') = c(1-c) N R(\tau, \tau') R(\rho, \rho').$$

Questions:

- Free energy?
- Ground state energy?
- Energy landscape?

Thank you for your attention.