Complexity of spin-glass Hamiltonians

Jiří Černý

University of Basel

October 4, 2023

Mean-field spin glasses

Hamiltonian. N interacting 'spins'

$$H_{N,p}(\boldsymbol{\sigma}) = \frac{1}{N^{(p-1)/2}} \sum_{i_1,\dots,i_p=1}^{N} J_{i_1,\dots,i_p} \sigma_{i_1} \dots \sigma_{i_p},$$

$J_\cdot$ i.i.d. standard Gaussian r.v.'s

p-spin SK: configuration space $\sigma \in \{-1,1\}^N$.

spherical p-spin SK: $\sigma \in S^{N-1}(\sqrt{N})$.

Hamiltonian is a centred Gaussian process with covariance

$$\mathbb{E}\big[H_{N,p}(\boldsymbol{\sigma}) H_{N,p}(\boldsymbol{\sigma}')\big] = N^{1-p}\big(\boldsymbol{\sigma} \cdot \boldsymbol{\sigma}'\big)^p.$$

Classical questions: statics of spin-glasses

Gibbs measure:

$$\mu_{N,p,\beta}(\mathrm{d}\sigma) = \frac{1}{Z_{N,p}(\beta)} e^{-\beta H_{N,p}(\sigma)} \Lambda_N(\mathrm{d}\sigma),$$

where Λ_N is the uniform probability measure on $S^{N-1}(\sqrt{N})$

Partition function.

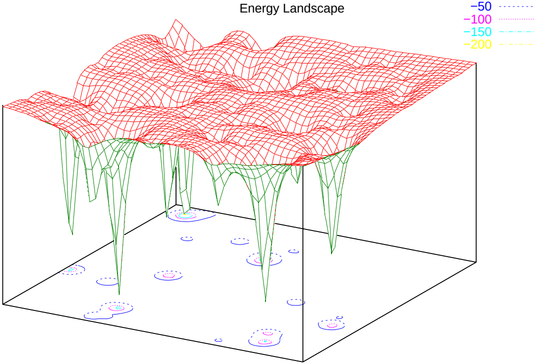
$$Z_{N,p}(\beta) = \int e^{-\beta H_{N,p}(\sigma)} \Lambda_N(\mathrm{d}\sigma).$$

Free energy.

$$\lim_{N\to\infty} \frac{1}{N} \log \mathbb{E}[Z_{N,p}(\beta)] = F_p(\beta).$$

Contributions by: Parisi, Talagrand, Guera, Toninelli, Panchenko, ...

Caricature of the process



"Typical realization of the Gibbs measure $e^{-\beta H_{N,p}(\sigma)}$"

Original motivation: Study a dynamics in this random landscape.

Motivation 2: Typical Morse functions on the N-sphere

Morse functions. $f : M \to \mathbb{R}$, such that all its critical points are non-degenerate. I.e. $\det(\nabla^2 f(x) \neq 0)$.

Typical functions = Gaussian processes.
By Schoenberg's theorem (1937), the only covariance functions that work on spheres of arbitrary dimension are mixtures of p-spins, i.e.

$$\mathbb{E}[f(x)f(x')] = \sum_{p=0}^{\infty} a_p (x \cdot x')^p, \qquad x, x' \in S^{N-1}, a_p \geq 0$$

(taking $f(x) = N^{-\frac{1}{2}} H_{N,p}(xN^{\frac{1}{2}})$ corresponds to taking $a_p = \delta_p$.)

Such processes are a.s. Morse functions.

Questions. Can we say more about them?

- number of critical points of various index
- they mutual position
- links by saddles
- Euler characteristics of level sets
- gradient flows, stable manifolds, ...

Complexity

Complexity.

The number of critical points of a given index k with value in $B \subset \mathbb{R}$

$$\mathrm{Crt}_{N,k}(B) = \# \left\{ \boldsymbol{\sigma} : \begin{array}{c} H_{N,p}(\boldsymbol{\sigma}) \in NB, \\ \nabla H_{N,p}(\boldsymbol{\sigma}) = 0, \\ i(\nabla^2 H_{N,p}(\boldsymbol{\sigma})) = k \end{array} \right\}.$$

∇, ∇^2 - gradient and Hessian restricted to $S^{N-1}(\sqrt{N})$
$i(\nabla^2 H_{N,p}(\boldsymbol{\sigma}))$ - index of $\nabla^2 H_{N,p} = \#$ negative eigenvalues

---

Total complexity.

$$\mathrm{Crt}_N(B) = \sum_k \mathrm{Crt}_{N,k}(B).$$

Complexity and random matrices

Result 1:

The (expected) complexity of the spherical p-spin spin glass
is related to the spectrum of random GOE matrices.

GOE: Probability distribution of the set of $N \times N$ real symmetric random matrices.

- The entries $(M_{ij}^N, i \leq j)$ are independent centred Gaussian r.v.s

$$\mathbb{E}[(M_{ij}^N)^2] = \frac{1+\delta_{ij}}{2N}.$$

- *Spectrum.* $\lambda_0^N \leq \lambda_1^N \leq \cdots \leq \lambda_{N-1}^N$
- *Spectral measure.* $L_N = \frac{1}{N}\sum_{i=0}^{N-1} \delta_{\lambda_i^N}$

Complexity and random matrices

Theorem (ABČ'12)

For all N, $p \geq 2$, $k \in \{0, \dots, N-1\}$, and $B \subset \mathbb{R}$,

$$\mathbb{E}[\mathrm{Crt}_{N,k}(B)] = C_{p,N} \times \mathbb{E}^N_{GOE}\left[e^{-NF_p(\lambda_k^N)}\mathbf{1}\left\{\lambda_k^N \in c_p B\right\}\right].$$

where

$$C_{p,N} = 2\sqrt{\frac{2}{p}}(p-1)^{\frac{N}{2}}, \quad F_p(x) = \frac{p-2}{2p}x^2, \quad c_p = \sqrt{\frac{p}{2(p-1)}}$$

Large deviation analysis

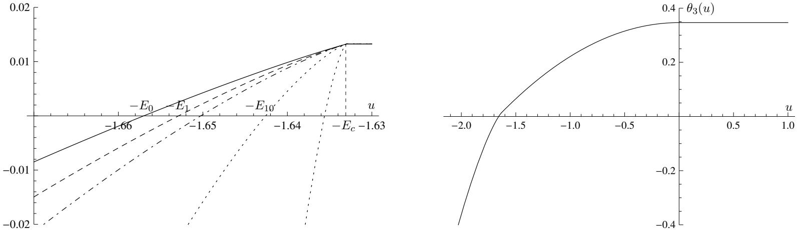
Define $\text{Crt}_{\cdot}(u) = \text{Crt}_{\cdot}((-\infty, u])$.

Theorem

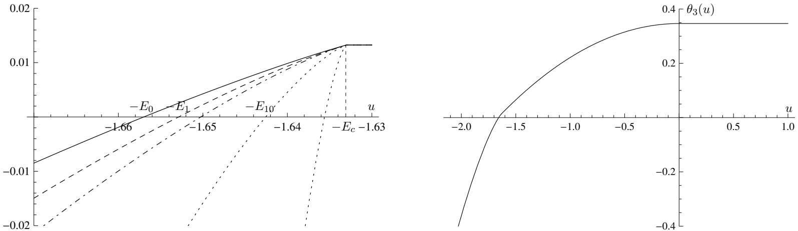
For every k, p fixed

$$\lim_{N\to\infty} \frac{1}{N} \log \mathbb{E}\, \text{Crt}_{N,k}(u) = \theta_{k,p}(u)$$

where the functions $\theta_{k,p}$, θ_p look like (for $p = 3$)



Implications for the energy landscape



Layered structure of the landscape.

- All critical points of index k are between $-E_k$ and $-E_c$.
- Below $-E_1$ there are only local minima
- Below $-E_2$ there are local minima and saddles of index 1, ...
- There are no critical points of finite index above $-E_c + \varepsilon$
- All critical points of index αN have value in a small interval around $E(\alpha)$

$-E_0$ is the same as the ground state energy computed from Parisi formula.

Proof: Kac-Rice formula for the complexity

Main tool in the proofs:

Theorem (Kac–Rice Formula, e.g. Adler-Taylor '07)

Under some mild conditions, with $H = H_{N,p}$

$$\mathbb{E}\,\mathrm{Crt}_{N,k}(B) = \int_{S^{N-1}} \mathrm{d}\sigma\, \phi_{\nabla H(\sigma)}(0)\mathbb{E}\Big[\big|\det \nabla^2 H(\sigma)\big| \mathbf{1}_{H(\sigma)\in B, i(\nabla^2 H(\sigma))=k} \,\Big|\, \nabla H(\sigma) = 0\Big].$$

Further results

Kac–Rice formula allows to compute the expectation of generalised complexity

$$\mathrm{Crt}_N(A) = \#\{\sigma \in S^{N-1} : \nabla_{\mathrm{sp}} H_N(\sigma) = 0, \mathbf{H}_N(\sigma) \in A_N\}$$

where $A = (A_N)_{N \geq 1}$ and $\mathbf{H}_N(\sigma) \in A_N$ means

$$\Big(\sigma, \tfrac{1}{N} H_N(\sigma), \tfrac{1}{N} \nabla H_N(\sigma), \tfrac{1}{N} \nabla^2 H_N(\sigma)\Big) \in A_N$$

and of

$$\theta_N(A) = \frac{1}{N} \log \mathbb{E}\, \mathrm{Crt}_N(A)$$

in many related models.

Examples:

- mixture of p-spins (given energy and index) [Auffinger-Ben Arous '13]
- TAP complexity [Fan-Mei-Montanari'18],
- bipartite spin glasses [McKenna'21]
- elastic manifold [Ben Arous-Bourgade-McKenna'22]
- ...
- in physics: Fyodorov et al., Ros, Biroli, Cammarota, Pacco, ...

Main problem

Is this computation useful?

- $\mathrm{Crt}_N(A)$ does not need to concentrate around its expectation
- In general only

$$\lim_{N\to\infty} \mathbb{E}\theta_N(A) \leq \lim_{N\to\infty} \frac{1}{N} \log \mathbb{E}\,\mathrm{Crt}_N(A)$$

- It is useful in certain cases: Trivialisation

Main problem

Is this computation useful?

- $\mathrm{Crt}_N(A)$ does not need to concentrate around its expectation
- In general only

$$\lim_{N\to\infty} \mathbb{E}\theta_N(A) \le \lim_{N\to\infty} \frac{1}{N} \log \mathbb{E}\,\mathrm{Crt}_N(A)$$

- It is useful in certain cases: Trivialisation

How to prove the concentration for the complexity?

Concentration: second moment method

A version of Kac–Rice formula allows to compute $\mathbb{E}(\mathrm{Crt}_N(A)^2)$
The calculation involves

$$\mathbb{E}\Big(|\det \nabla^2_{\mathrm{sp}} H_N(\sigma) \det \nabla^2_{\mathrm{sp}} H_N(\sigma')| \mathbf{1}_A(\mathbf{H}_N(\sigma,\sigma')) \,\Big|\, \nabla_{\mathrm{sp}} H_N(\sigma) = \nabla_{\mathrm{sp}} H_N(\sigma') = 0\Big)$$

- Subag (2017) for pure p-spin, all critical points with $H_N(\sigma) \le EN$, $E \le -E_c$

$$\frac{\mathbb{E}(\mathrm{Crt}_N(E)^2)}{(\mathbb{E}\,\mathrm{Crt}_N(E))^2} \xrightarrow{N\to\infty} 1$$

- Auffinger–Gold (2020), critical points of a given finite index
- Kivimae (2022), bipartite spherical p, q-spin
- . . . Belius–Schmidt (2023+) . . .

Concentration: second moment method

A version of Kac–Rice formula allows to compute $\mathbb{E}(\mathrm{Crt}_N(A)^2)$
The calculation involves

$$\mathbb{E}\Big(|\det \nabla^2_{\mathrm{sp}} H_N(\sigma) \det \nabla^2_{\mathrm{sp}} H_N(\sigma')| \mathbf{1}_A(\mathbf{H}_N(\sigma,\sigma')) \,\Big|\, \nabla_{\mathrm{sp}} H_N(\sigma) = \nabla_{\mathrm{sp}} H_N(\sigma') = 0\Big)$$

- Subag (2017) for pure p-spin, all critical points with $H_N(\sigma) \leq EN$, $E \leq -E_c$

$$\frac{\mathbb{E}(\mathrm{Crt}_N(E)^2)}{(\mathbb{E}\,\mathrm{Crt}_N(E))^2} \xrightarrow{N\to\infty} 1$$

- Auffinger–Gold (2020), critical points of a given finite index
- Kivimae (2022), bipartite spherical p, q-spin
- . . . Belius–Schmidt (2023+) . . .

These computations are difficult . . .
Can we do something else?

General setting: mixed p-spin models

Hamiltonian:

$$H_N(\sigma, J) = \sum_{p=1}^{P} a_p H_N^p(\sigma, J) + f_N(\sigma)$$

where

$$H_N^p(\sigma, J) = \sqrt{N} \sum_{i_1,\dots,i_p=1}^{N} J_{i_1,\dots,i_p} \sigma_{i_1} \dots \sigma_{i_p}, \quad \sigma \in S^{N-1} \text{ (or } B_N)$$

and $J = (J_{i_1,\dots,i_p}) \in \mathbb{R}^{N+N^2+\dots+N^P}$ is a standard Gaussian vector.

External field f_N:

- might be random, independent of J
- we assume that

$$c_f = \frac{1}{N} \sup_{\sigma \in B_{N-1}} \max\left(|\nabla f_N(\sigma)|, \|\nabla^2 f_N(\sigma)\|, \|\nabla^3 f_N(\sigma)\|\right) < \infty$$

Result: non-flat critical points

Definition. A critical point σ of H_N is called η-**non-flat** if

$$\operatorname{spec} N^{-1}\nabla^2_{\mathrm{sp}} H_N(\sigma, J) \cap [-\eta, \eta] = \emptyset.$$

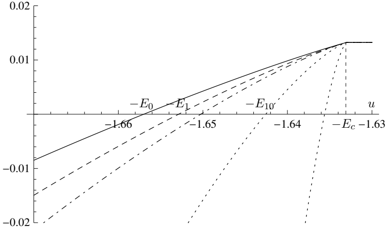
Number of non-flat critical points with properties A:

$$\mathrm{Crt}_N^{\eta}(A, J) = \#\{\sigma : \sigma \text{ is } \eta\text{-non-flat critical point}, \mathbf{H}_N(\sigma, J) \in A_N\}.$$

and

$$\theta_N^{\eta}(A, J) = \frac{1}{N} \log \mathrm{Crt}_N^{\eta}(A, J)$$

Remark. Minima under $-E_\infty$ in pure p-spin are typically non-flat



The concentration result

Theorem (Belius–Č'23+)

For all $0 < \eta_1 < \eta_2 \le 1$ *there is* $c = c(\eta_1) > 0$ *and* $\kappa = \kappa(c_f, mixture, \eta_1) > 0$ *such that:*
For $0 < \delta \le \min(c, \eta_2 - \eta_1)$, $N \ge \kappa\delta^{-2}$ *and sets* $A_2 \subset A_1$ *with* $d(A_2, A_1^c) \ge \delta$*:*

$$\mathbb{P}\big(\mathrm{Crt}_N^{\eta_1}(A_1) \le \mathrm{Med}\,\mathrm{Crt}_N^{\eta_2}(A_2)\big) \le \mathrm{e}^{-\kappa\delta^2 N},$$

and

$$\mathbb{P}\big(\mathrm{Crt}_N^{\eta_2}(A_2) \ge \mathrm{Med}\,\mathrm{Crt}_N^{\eta_1}(A_1)\big) \le \mathrm{e}^{-\kappa\delta^2 N}.$$

Corollary

The same hold if Crt_N *is replaced by* θ_N*:*

$$\mathbb{P}\big(\theta_N^{\eta_1}(A_1) \le \mathrm{Med}\,\theta_N^{\eta_2}(A_2)\big) \le \mathrm{e}^{-\kappa\delta^2 N},$$

$$\mathbb{P}\big(\theta_N^{\eta_2}(A_2) \ge \mathrm{Med}\,\theta_N^{\eta_1}(A_1)\big) \le \mathrm{e}^{-\kappa\delta^2 N}.$$

Applications

Corollary

If

$$(A,\eta) \mapsto \theta^\eta(A) := \lim_{N\to\infty} \operatorname{Med} \theta^\eta_N(A)$$

is continuous in (A,η), *then*

$$\lim_{N\to\infty} \theta^\eta_N(A) = \theta^\eta(A), \qquad \textit{in probability}$$

Pros: Pretty general, in typical points the median should be continuous

Cons:

- How to compute $\operatorname{Med} \theta^\eta_N(A)$? Does it converge?
- How to compare with the Kac–Rice computation?

Applications: Comparison with Kac–Rice

Corollary

If

$$\frac{1}{N}\log\mathbb{E}\big(\mathrm{Crt}^{\eta}(A)^2\big) = \frac{2}{N}\log\mathbb{E}\big(\mathrm{Crt}^{\eta}(A)\big) + o(1),$$

and $(A,\eta)\mapsto \lim_{N\to\infty}\log\mathbb{E}\,\mathrm{Crt}_N^{\eta}(A)$ *is continuous in* (A,η)*, then*

$$\lim_{N\to\infty}\theta_N^{\eta}(A) = \lim_{N\to\infty}\log\mathbb{E}\,\mathrm{Crt}_N^{\eta}(A).$$

Pros:

- Compares to Kac–Rice
- Requires weaker second moment computation, (cf. [BBM])
- Proof is robust

Cons: Still requires second moment computation.

Concrete application

Concentration for number of critical points with given radial derivative.

Theorem (Belius-Schmidt '23+)

Consider mixed p-spin Hamiltonian without external field. Let

$$A = A(x, \varepsilon) = \{N^{-1} \partial_r H_N(\sigma) \in (x - \varepsilon, x + \varepsilon)\}$$

If $x \in [x_-, x_+]$, *then, in probability,*

$$\lim_{\varepsilon \downarrow 0} \lim_{N \to \infty} \theta_N(A(x, \varepsilon)) = \theta(x).$$

Tools: Gaussian isoperimetric inequality

- P_n standard Gaussian measure on $\mathbb{R}^n$
- For $B \subset \mathbb{R}^n$, define t-blowup as

$$B_t = \{x \in \mathbb{R}^n : d(x, B) \leq t\}$$

Theorem

For any $n \geq 1$, $t \geq 0$ and $B \subset \mathbb{R}^n$ measurable

$$P^n(B_t) \geq 1 - \exp\left\{ -\tfrac{1}{2}\left(t + \Psi^{-1}(P^n(B))^2\right)^2\right\}$$

We will apply this for J's, that is:

- $n = N + N^2 + \cdots + N^P$, (typical norm of J is $\sqrt{n}$),
- $B = \{\mathrm{Crt}_N^\eta(A, J) \geq \mathrm{Med}\,\mathrm{Crt}_N^\eta(A, J)\}$,
- $t \sim \sqrt{N}$

Quantitative implicit function theorem

Theorem

Let $x_0 \in \mathbb{R}^n, y_0 \in \mathbb{R}^m$, $\delta_x, \delta_y > 0$ *and let* $F : B_n(x_0, \delta_x) \times B_m(y_0, \delta_y) \to \mathbb{R}^n$ *be a* C^1*-function such that* $F(x_0, y_0) = 0$, $D_x F(x_0, y_0)$ *is an invertible matrix and*

$$\sup_{\substack{x,y: |x-x_0| \leq \delta_x \\ |y-y_0| \leq \delta_y}} \left\| I - (D_x F(x_0, y_0))^{-1} D_x F(x, y) \right\|_{op} \leq \frac{1}{2}.$$

Let

$$H = \sup_{x,y: |x-x_0| \leq \delta_x, |y-y_0| \leq \delta_y} \|D_y F(x, y)\|_{op},$$

$$M = \|(D_x F(x_0, y_0))^{-1}\|_{op},$$

$$\overline{\delta}_y = \min\left(\delta_x/(2MH), \delta_y\right).$$

Then there exists a continuous function $g : B_m(y_0, \overline{\delta}_y) \to \mathbb{R}^m$ *such that* $(x, y) \in B_m(y_0, \delta_x) \times B_m(y_0, \overline{\delta}_y)$ *is a solution to* $F(x, y) = 0$ *iff* $x = g(y)$. *Furthermore* g *is Lipschitz continuous with constant at most* $2MH$.

Will be applied for $F = \nabla_{\mathrm{sp}} H_N(\sigma, J)$, $x \leftrightarrow \sigma$ and $y \leftrightarrow J$.

Regularity estimates

Let $\mathbf{H}_N$ be any of $H_N, \nabla H_n, \nabla^2 H_n$

Lemma

(a) *For every* J, J' *and* $\sigma \in S^{N-1}$

$$\|\mathbf{H}_N(\sigma, J + J') - \mathbf{H}_N(\sigma, J)\| \leq c\|J'\|\sqrt{N}$$

(b) *On the set* $\mathcal{G}_N$ *of* J*'s of probability at least* $1 - \mathrm{e}^{-2N}$*, for all* $\sigma, \sigma' \in S^{N-1}$

$$\|\mathbf{H}_N(\sigma, J) - \mathbf{H}_N(\sigma', J)\| \leq cN\|\sigma - \sigma'\|.$$

Key lemma

As consequence of the implicit function theorem, the non-flat critical points cannot appear/disappear after perturbation of J's of order $\sqrt{N}$

Lemma

Let $\eta \in (0, 1]$ and $\delta \in (0, c(\eta))$, and $J \in \mathcal{G}_N$.

(a) *If σ is η-flat critical point of $H_N(\cdot, J)$ and $\|J'\| \leq \delta\sqrt{N}$, then there is exactly one critical point σ' of $H_N(\cdot, J + J')$ which is $(\eta - c\delta)$-flat such that*

$$\|\sigma - \sigma'\| \leq c\delta$$
$$N^{-1}\|\mathbf{H}_N(\sigma', J + J') - \mathbf{H}_N(\sigma, J)\| \leq c\delta$$

(b) *If σ is a η-flat critical point of $H_N(\cdot, J + J')$, then ...*

Key lemma

As consequence of the implicit function theorem, the non-flat critical points cannot appear/disappear after perturbation of J's of order $\sqrt{N}$

Lemma

Let $\eta \in (0,1]$ *and* $\delta \in (0, c(\eta))$, *and* $J \in \mathcal{G}_N$.

(a) *If* σ *is* η*-flat critical point of* $H_N(\cdot, J)$ *and* $\|J'\| \leq \delta\sqrt{N}$, *then there is exactly one critical point* σ' *of* $H_N(\cdot, J+J')$ *which is* $(\eta - c\delta)$*-flat such that*

$$\|\sigma - \sigma'\| \leq c\delta$$
$$N^{-1}\|\mathbf{H}_N(\sigma', J+J') - \mathbf{H}_N(\sigma, J)\| \leq c\delta$$

(b) *If* σ *is a* η*-flat critical point of* $H_N(\cdot, J+J')$, *then* ...

Corollary

For $J \in \mathcal{G}_N$ *and* $\|J'\| \leq \delta\sqrt{N}$

$$\mathrm{Crt}_N^{\eta - c\delta}(A_{c\delta}, J+J') \geq \mathrm{Crt}_N^{\eta}(A, J) \geq \mathrm{Crt}_N^{\eta + c\delta}(A_{-c\delta}, J+J').$$

Summary

- We obtained "concentration" for the number of *non-flat* critical points
- The estimates are very robust :
  - use only regularity of the landscape and basic techniques
  - can be extended to other domains than S^{N-1} (TAP equations)
- Can be generalised to infinite mixtures
- We hope that they will be useful

Thank you!