Journal of Mathematical Physics

Sharp complexity asymptotics and topological trivialization for the (*p*, *k*) spiked tensor model

Cite as: J. Math. Phys. **63**, 043303 (2022); https://doi.org/10.1063/5.0070300
Submitted: 05 September 2021 • Accepted: 21 March 2022 • Published Online: 08 April 2022

Antonio Auffinger, Gerard Ben Arous and Zhehua Li

Journal of Mathematical Physics ARTICLE scitation.org/journal/jmp

Sharp complexity asymptotics and topological trivialization for the (p,k) spiked tensor model

Cite as: J. Math. Phys. **63**, 043303 (2022); doi: 10.1063/5.0070300
Submitted: 5 September 2021 • Accepted: 21 March 2022 •
Published Online: 8 April 2022

View Online Export Citation CrossMark

Antonio Auffinger,[1,a] Gerard Ben Arous,[2,b] and Zhehua Li[1,c]

AFFILIATIONS

[1] Northwestern University, Evanston, Illinois 60208, USA
[2] New York University, New York, New York 10012, USA

Note: This paper is part of the Special Collection in Honor of Freeman Dyson.
[a] **E-mail:** tuca@northwestern.edu
[b] **Author to whom correspondence should be addressed:** benarous@cims.nyu.edu
[c] **E-mail:** zhehua.li@northwestern.edu

ABSTRACT

Using precise random matrix theory tools and the Kac–Rice formula, we provide sharp $O(1)$ asymptotics for the average number of deep minima of the (p,k) spiked tensor model. These sharp estimates allow us to prove that, when the signal-to-noise ratio is large enough, the expected number of deep minima is asymptotically finite as N tends to infinity and to establish the occurrence of topological trivialization by showing that this number vanishes when the strength of the signal-to-noise ratio diverges. We also derive an explicit formula for the value of the absolute minimum (the limiting ground state energy) on the N-dimensional sphere, similar to the recent work of Jagannath, Lopatto, and Miolane [Ann. Appl. Probab. **4**, 1910–1933 (2020)].

I. INTRODUCTION

It is now well established that random smooth functions of many variables, and the landscapes they define, may be topologically complex. This complexity may be described through topological quantities, such as the Betti numbers of their level sets, or by the number of their critical points, which both may happen to be exponentially large (in the number of variables). These landscapes could, for instance, be the energy or Hamiltonian landscape in statistical physics of disordered media. This question was introduced some time ago for the simplest example of such landscapes, i.e., the Hamiltonian of spherical spin-glasses.[3,9,10] The topological complexity of random landscapes is also relevant for certain high-dimensional statistical estimation tasks, for instance, the landscape defined by natural empirical risks.[20] The mathematical tools to study these complexity questions are naturally given by the Kac–Rice formula, which translate the question into a random matrix problem.

We study here the case of the (p,k) spiked tensor model, introduced in Ref. 24 and defined more precisely below. This function is simply the restriction to the sphere of the sum of a random Gaussian homogeneous polynomial of N variables and of degree p, which is probably the simplest example of a complex random landscape, and of a degree k monomial, which is obviously not a complex function.

More precisely, let $S^{N-1}(\sqrt{N}) = \{\sigma\in\mathbb{R}^N : \sum_{i=1}^N \sigma_i^2 = N\}$ be the N-sphere of radius $\sqrt{N}$ and fix $\mathbf{v}_0\in S^{N-1}(\sqrt{N})$. Given integers $p,k\geq 1$, $\lambda\in\mathbb{R}$, let

$$H_N(\sigma) = \frac{1}{N^{\frac{p-1}{2}}}\sum_{1\leq i_1,i_2,\ldots,i_p\leq N} J_{i_1,i_2,\ldots,i_p}\sigma_{i_1}\sigma_{i_2}\ldots\sigma_{i_p} - \frac{\lambda N}{k}\left(\frac{\sigma\cdot\mathbf{v}_0}{N}\right)^k, \tag{1.1}$$

where $\sigma=(\sigma_1,\ldots,\sigma_N)\in S^{N-1}(\sqrt{N})$ and $(J_{i_1,i_2,\ldots,i_p})_{1\leq i_1,\ldots,i_p\leq N}$ are independent standard Gaussian random variables. We call H_N the Hamiltonian of the (p,k) spiked tensor model.

Note that, when $p=k\geq 3$, this function is simply the log-likelihood function for the classical spiked tensor model whose complexity has been studied in Ref. 7, and then, the parameter λ has a natural interpretation related to the signal-to-noise ratio and the sample size. When

$p = 1$, this function is the Hamiltonian of a pure spherical p-spin in a magnetic field, and λ is then the strength of the magnetic field (up to a normalization).

Without loss of generality, we refer to the direction of $\mathbf{v}_0$ as the North Pole of the model, and we let

$$m(\sigma) = \sigma \cdot \mathbf{v}_0 / N \in [-1, 1]$$

be the overlap of σ with the signal $\mathbf{v}_0$. The aim of this paper is to investigate the landscape of the random function H_N around its ground state energy

$$L_N := \min_{\sigma \in S^{N-1}(\sqrt{N})} H_N(\sigma)$$

and the overlap between its ground state and the signal,

$$m_N := \left(\arg\min_{\sigma \in S^{N-1}(\sqrt{N})} H_N(\sigma)\right) \cdot \mathbf{v}_0.$$

For each $\lambda > 0$, define

$$m_\lambda := \min\left\{1, \left(\frac{(p-2)\sqrt{p}}{\lambda\sqrt{p-1}}\right)^{\frac{1}{k}}\right\}. \tag{1.2}$$

As illustrated in the transformative work of Ros *et al.*[24] (see also Secs. 2.3 and 2.4 in Ref. 7), in the "low-latitude" region, $|m| \le m_\lambda$, H_N has a rugged energy landscape, with exponentially many critical values in N, resembling the spherical p-spin spin glass models,[2] while in the "high-latitude" region, $|m| \ge m_\lambda$, it resembles a convex potential. The study of phase transitions in the topology of level sets of H_N and limit theorems for m_N and L_N have drawn much attention recently; see, for instance, Refs. 11, 18, 22, and 23.

Here, we focus on providing a better understanding of the model in the presence of a strong signal, that is, when λ in (1.1) is large. In this case, low energy level sets of Hamiltonian H_N will go through a phenomenon called "topology trivialization," a term pioneered by Fyodorov and Le Doussal[16] and discussed in Fyodorov's remarkable work.[14,15] In short, for λ large, one does not expect exponentially many critical values of H_N with energy near the ground state energy L_N. Instead, as λ increases, the location of ground state energy starts to align with the signal (the North Pole in Fig. 1). Furthermore, near the ground state energy, the landscape should only have a finite number of critical values. In this paper, we confirm the existence of such a phase (the trivialization phase) for the (p,k)-spiked model.

We will now describe our results. Let ∇ and ∇^2 denote the spherical gradient and Hessian with respect to the standard metric on $S^{N-1}(\sqrt{N})$. For open sets $M \subseteq [-1,1]$ and $E \subseteq \mathbb{R}$, we denote the total number of critical points of H_N that have overlap with $\mathbf{v}_0$ in M and whose critical values are in NE by

$$\mathrm{Crt}_N(M,E) := \sum_{\sigma \in S^{N-1}(\sqrt{N}), \nabla H_N(\sigma)=0} \mathbf{1}_{\{\sigma\cdot\mathbf{v}_0/N \in M\}} \cdot \mathbf{1}_{\{H_N(\sigma)/N \in E\}}$$

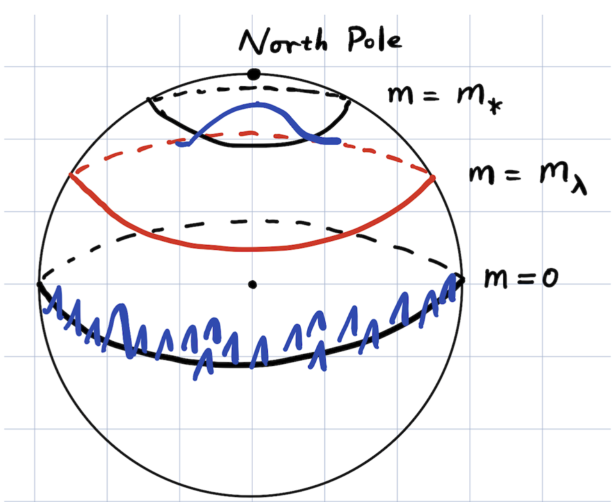

FIG. 1. The landscape of $H_N(\sigma)$ on S^{N-1}. $\mathbf{v}_0$ is the North Pole, $m = \langle \sigma, \mathbf{v}_0 \rangle / N$. The spikes around the equator represent numerous local maxima (minima) that are possibly exponential in N in the "low-latitude" region $|m| \le m_\lambda$, where m_λ is defined in (1.2). When $m \ge m_\lambda$, there are only a few critical points on a parallel $m = m_*$.

and the corresponding number of critical points of index $l = 0, \ldots, N-1$ by

$$\mathrm{Crt}_{N,l}(M,E) := \sum_{\sigma \in S^{N-1}(\sqrt{N}), \nabla H_N(\sigma)=0} \mathbf{1}_{\{\sigma \cdot \mathbf{v}_0/N \in M\}} \cdot \mathbf{1}_{\{H_N(\sigma)/N \in E\}} \mathbf{1}_{\{i(\nabla^2 H_N)=l\}}.$$

Here, the index $i(\cdot)$ is the number of negative eigenvalues of the corresponding matrix. When $l = 0$, $Ctr_{N,0}(M,E)$ counts the number of local minima that have overlap with $\mathbf{v}_0$ in M and whose critical values are in NE.

Let $m_* := m_*(\lambda)$ be the largest solution of

$$\frac{\lambda m^k}{\sqrt{p}} = \frac{m^2}{\sqrt{1-m^2}} \tag{1.3}$$

on $(0,1]$. Such m_* exists when $\lambda \geq \lambda^{(1)}(p,k) = \begin{cases} 0, & k = 1,2, \\ \sqrt{p\dfrac{(k-1)^{k-1}}{(k-2)^{k-2}}}, & k > 2. \end{cases}$ Set

$$x_*(\lambda) := -\frac{\lambda m_*^k}{k} - \sqrt{p(1-m_*^2)}.$$

Our three main theorems explain the existence of the trivialization phase. In the first result, we provide sharp asymptotics of the average number of critical values. For λ sufficiently large, they remain of constant order, not diverging with N. Furthermore, in this regime, all critical values are, with probability going to one, local minima.

Theorem 1.1. Let M be an open interval of $(-1,1)$ and E be a bounded open interval on $\mathbb{R}$ such that

$$\sup E < -2\sqrt{\frac{p-1}{p}} - \left|\frac{1}{p} - \frac{1}{k}\right|\lambda. \tag{1.4}$$

There exists $c > 0$ such that for any $\lambda \geq c$, there exists a constant $C = C(\lambda,p,k)$ that does not depend on M and E such that

$$\lim_{N\to\infty} \mathbb{E}[\mathrm{Crt}_{N,0}(M,E)] = \lim_{N\to\infty} \mathbb{E}[\mathrm{Crt}_N(M,E)] = \begin{cases} C > 0 & \text{if } x_*(\lambda) \in E \text{ and } m_*(\lambda) \in M, \\ 0 & \text{otherwise.} \end{cases}$$

Remark 1.2. The role of (1.4) is to make sure we are counting critical values at the bottom of the landscape, the region that want to study. If we remove condition (1.4), critical values near zero (and of diverging order) will provide the main contribution of $\mathrm{Crt}_N(M,\mathbb{R})$.

The constant C in Theorem 1.1 is explicit, and we can further consider its asymptotics when $\lambda \to \infty$.

Theorem 1.3. Let $C(\lambda,p,k) > 0$ be the constant given in Theorem 1.1. Then, for any $p \geq 3$ and $k \geq 1$,

$$\lim_{\lambda\to\infty} C(\lambda,p,k) = 1.$$

Theorem 1.3 confirms the existence of the trivialization phase for the (p,k) spiked tensor model. It is believed that as $\lambda \to \infty$, the deterministic potential becomes stronger and the landscape should approach a convex potential with a unique minimum located exactly at the signal vector $\mathbf{v}_0$; see Ref. 24. Closest to our setting is the recent work of Belius *et al.*,[5] which deals with the mean number of critical points for mixed spherical spin glass models with an external field.

Our third result verifies that indeed the ground state and the ground state energy align with the signal once λ is large enough.

Theorem 1.4. For any integers $p \geq 3$ and $k \geq 1$, there exists $c > 0$ such that for $\lambda > c$, the following holds:

$$\lim_{N\to\infty} \frac{1}{N} m_N = m_*(\lambda) \quad \text{almost surely} \tag{1.5}$$

and

$$\lim_{N\to\infty} \frac{1}{N} L_N = -\frac{\lambda m_*^k(\lambda)}{k} - \sqrt{p(1-m_*^2(\lambda))} \quad \text{almost surely.} \tag{1.6}$$

A word of comment is needed here. Theorem 1.4 was first conjectured and proposed in the article of Ros *et al.*,[24] where the authors studied the number of local minima of H_N via a replica theoretic approach. The above formulas are not expected to be true when λ is small (see Refs. 17 and 24 and Remark 1.8) for any choices of (p,k). In the case $p = k$, the (p,k) spiked tensor model has a log-likelihood interpretation as tensor principle component analysis (PCA). This interpretation was used by Jagannath, Lopatto, and Miolane to derive asymptotic formulas for the ground state energy for all values of λ. Theorem 1.4 is an extension of Theorem 1.2 in Ref. 18 for $p \neq q$ and λ sufficiently large, although the method of the proof is different.

Let us now explain where Theorems 1.1, 1.3, and 1.4 come from. They will follow from the main technical contribution of this paper, which is the derivation of $\mathcal{O}(1)$ asymptotics of $\mathbb{E}[\mathrm{Crt}_N(M,E)]$ in the large N limit. Exponential asymptotics of $\mathbb{E}[\mathrm{Crt}_N(M,E)]$ were determined in Ref. 7 in the case $k \neq p$. Define $\tilde{S}_{p,k} : (-1,1)\times(-\infty,-\sqrt{2}) \to \mathbb{R}$ as

$$\tilde{S}_{p,k}(m,y) := \frac{1}{2}\log\big((1-m^2)(p-1)\big) + \frac{2-p}{2p}y^2 - \frac{\lambda m^k}{p}\sqrt{\frac{2(p-1)}{p}}\,y - \frac{\lambda^2 m^{2k-2}}{2p^2}\big(p+(1-p)m^2\big) - I_1(-y), \tag{1.7}$$

where

$$y = y(x,m) := \frac{px-(1-p/k)\lambda m^k}{\sqrt{2p(p-1)}} \tag{1.8}$$

and

$$I_1(z) = \int_{\sqrt{2}}^{z}\sqrt{t^2-2}\,dt \quad \text{for} \quad z \geq \sqrt{2}, \quad I_1(z) = \infty \quad \text{for} \quad z < \sqrt{2}.$$

The next two results do not require any assumptions on λ.

Theorem 1.5. Let M be an open interval of $(-1,1)$ such that $\bar{M} \subset (-1,1)$ and E be a bounded open interval on $\mathbb{R}$ such that sup $E < -2\sqrt{\frac{p-1}{p}} - \left|\frac{1}{p}-\frac{1}{k}\right|\lambda$; then, as $N\to\infty$,

$$\mathbb{E}[Crt_{N,0}(M,E)] = \frac{\sqrt{2}h(y_o)\Big(\sqrt{p}\big(1-m_o^2\big)^{-\frac{3}{2}} - \lambda(k-1)m_o^{k-2}J(m_o,y_o)\Big)}{\Big(\sqrt{y_o^2-2}-y_o\Big)p\sqrt{\left|\partial_{yy}\tilde{S}_{p,k}(m_o,y_o)g''(m_o)\right|}}e^{N\tilde{S}_{p,k}(m_o,y_o)}(1+o(N)), \tag{1.9}$$

where

$$h(y) = \left|\frac{y-\sqrt{2}}{y+\sqrt{2}}\right|^{\frac{1}{4}} + \left|\frac{y+\sqrt{2}}{y-\sqrt{2}}\right|^{\frac{1}{4}},$$

$$J(m,y) = \exp\left(-\left(\frac{\lambda^2}{2p^2}m^{2k-2}\big(p(1-m^2)+m^2\big) + \frac{\lambda m^k y}{2p}\sqrt{\frac{2(p-1)}{p}}\right)\right),$$

$$\tilde{E}_m := \sqrt{\frac{p}{2(p-1)}}\left(E - \lambda m^k\left(\frac{1}{p}-\frac{1}{k}\right)\right), \quad \forall m \in M,$$

$$y_o := y_o(m_o), \quad y_o(m) = \operatorname{argmax}_{y\in\tilde{E}_m}\tilde{S}_{p,k}(m,y),$$

and

$$m_o := \operatorname{argmax}_{m\in\bar{M}} g(m), \quad g(m) = \tilde{S}_{p,k}(m,y_o(m)).$$

Theorem 1.5 naturally leads to the following corollary.

Corollary 1.6. Let M and E be the same as in Theorem 1.5; then,

$$\lim_{N\to\infty}\frac{1}{N}\log\mathbb{E}[Crt_{N,0}(M,E)] = \sup_{m\in\bar{M}}\sup_{y\in\tilde{E}_m}\tilde{S}_{p,k}(m,y). \tag{1.10}$$

Remark 1.7. The function $S_{p,k}(m,x) := \tilde{S}_{p,k}(m, y(x,m))$ describes the exponential behavior of $\mathbb{E}[\mathrm{Crt}_N(M,E)]$ with respect to the dimension N, and it is called the annealed complexity, a function of $m \in [-1,1]$ and $x \in \mathbb{R}$ such that for any Borel sets $M \subset [-1,1]$ and $E \subset \mathbb{R}$,

$$\sup_{m\in M^{\circ}, x\in E^{\circ}} S_{p,k,0}(m,x) \le \liminf_{N\to\infty} \frac{1}{N} \log \mathbb{E}[\mathrm{Crt}_{N,0}(M,E)]$$
$$\le \limsup_{N\to\infty} \frac{1}{N} \mathbb{E}[\mathrm{Crt}_{N,0}(M,E)] \le \sup_{m\in\bar{M}, x\in\bar{E}} S_{p,k,0}(m,x). \tag{1.11}$$

We now mention one important remark. Looking at Theorems 1.1, 1.3, and 1.4, it is natural to ask how large λ needs to be so the model is in the trivialization phase. This leads to the notion of a trivialization threshold, which we define as the real number λ_* such that (1.5) and (1.6) hold for all $\lambda > \lambda_*$ and fail for all $\lambda < \lambda_*$.

For the (p,k)-spiked tensor, the remark below provides partial information on λ_*.

Remark 1.8 (trivialization threshold). Recall m_λ from (1.2). Let

$$\lambda^{(2)}(p,k) = \inf\left\{\lambda \ge \lambda^{(1)}(p,k) : m_*(\lambda) \ge m_\lambda\right\},$$

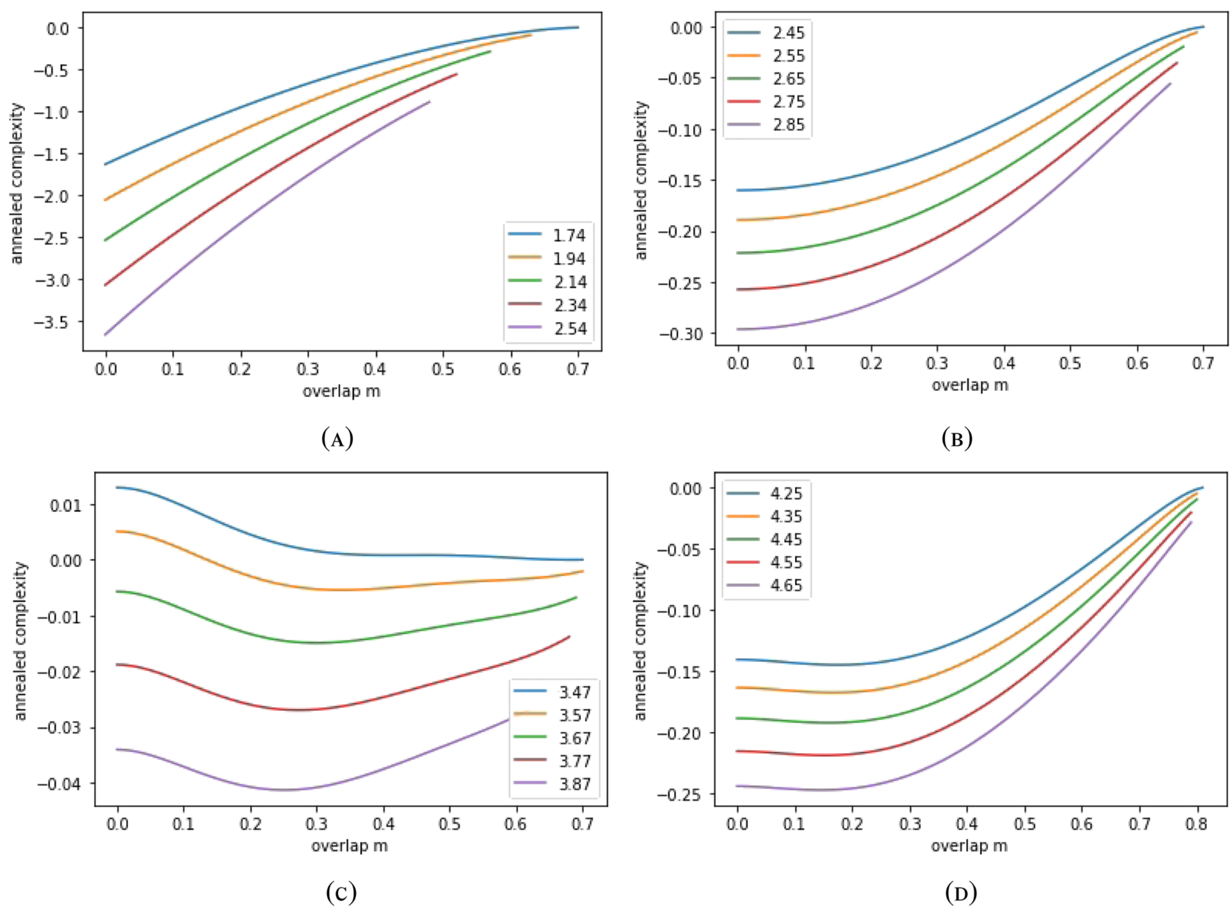

FIG. 2. $S_{p,k}(m, x_*(\lambda))1_{m\in[0,m_\lambda]}$ with different values of p, k, and λ. The numbers in the legends are values of λ. In each of the panels and for each $m \in [0,1]$, the values of $S_{p,k}(m, x_*(\lambda))1_{m\in[0,m_\lambda]}$ decrease as λ increases. For $p = 3, k = 1, \lambda^{(1)} = 0, \lambda^{(2)} = \lambda_{tr} = 1.732$. For $p = 3, k = 2, \lambda^{(1)} = 0, \lambda^{(2)} = \lambda_{tr} = 2.449$. For $p = 3, k = 3, \lambda^{(1)} = \lambda^{(2)} = 3.464, \lambda_{tr} = 3.619$. For $p = 4, k = 3, \lambda^{(1)} = 4, \lambda^{(2)} = \lambda_{tr} = 4.243$. (a) p = 3, k = 1. (b) p = 3, k = 2. (c) p = 3, k = 3. (d) p = 4, k = 3.

and define

$$\lambda_{tr} = \inf\Big\{\lambda \geq \lambda^{(2)}(p,k) : \sup_{0\leq m\leq m_\lambda} S_{p,k}(m, x_*(\lambda)) \leq 0 \text{ and} \tag{1.12}$$

$$S_{p,k}(m, x_*(\lambda))1_{m\in[0,m_\lambda]} \text{ is a decreasing function of } \lambda \text{ on } [\lambda^{(2)}(p,k), \infty)\Big\}. \tag{1.13}$$

Our Proof of Theorem 1.4 shows that (1.5) and (1.6) hold for all $\lambda > \lambda_{tr}$, that is,

$$\lambda_{tr} \geq \lambda_*.$$

We expect that the opposite inequality $\lambda_{tr} \leq \lambda_*$ also holds, and therefore, $\lambda_{tr} = \lambda_*$. Indeed, we expect that (1.12) implies (1.13) and that (1.5) and (1.6) fail for $\lambda < \lambda_{tr}$. Figure 2 shows a plot of the annealed complexity for various values of λ. For the spiked tensor model ($p = k > 2$), it has been shown in Refs. 7 and 18 that $\lambda^{(1)} = \lambda^{(2)} < \lambda_{tr}$.

We finish this Introduction mentioning a few related results and a brief description of the rest of this paper. First, the study of models such as the (p,k) spiked tensor along the direction of high-dimensional statistical inference was initiated by Montanari and Richard.[22] For the readers who are particularly interested in Tensor PCA and the spiked matrix-tensor model, we refer the reader to Refs. 18, 21, 22, and 24–26 and the references therein. A prototypical inference model called spiked matrix-tensor model, which is closely related to the case of $k = 2$ and $p \geq 3$, was extensively studied in Refs. 21, 25, and 26. In a recent paper by Maillard, Ben Arous, and Biroli, the complexity study (using the replicated Kac–Rice approach) is extended to current machine learning models, such as random generalized linear models and neural networks.

In Sec. II, we prove Theorem 1.5. We first show that the deep minima dominate the total number of critical points in Proposition 2.1. This result allows us to transform the problem of computing the mean number of deep minima into a problem of computing the mean Euler characteristic of level set for which we could use tools from random matrix theory to compute the characteristic polynomial of a deformed Gaussian Orthogonal Ensemble (GOE).

In Sec. III, we prove Theorems 1.1 and 1.3 where we study the mean number of deep minima (minima near the bottom of the energy landscape) and its asymptotic as $\lambda \to \infty$.

In Sec. IV, we analyze the ground state energy and prove Theorem 1.4. We first provide in Proposition 4.1 an upper bound of the ground state energy by restricting to energies with fixed latitude m, a method that was used[18] in the case of $k = p$. A matching lower bound is given in Proposition 4.2 by exploring the supremum of the annealed complexity near the bottom of the energy landscape.

II. PROOF OF THEOREM 1.5

The (normalized) GOE of size N (denoted by W_N) is a real symmetric random matrix $(W_{ij})_{1\leq i,j\leq N}$ such that $\{W_{ij}\}_{1\leq i\leq j\leq N}$ are independent zero mean normal random variables with $\mathbb{E}[W_{ij}^2] = \frac{1}{N}$ and $\mathbb{E}[W_{ii}^2] = \frac{2}{N}$.

We will work with the rescaled Hamiltonian f on the unit sphere S^{N-1},

$$f(\sigma) := \frac{H_N(\sqrt{N}\sigma)}{\sqrt{N}} = -\sum_{1\leq i_1,i_2,\ldots,i_p\leq N} J_{i_1,i_2,\ldots,i_p}\sigma_{i_1}\sigma_{i_2}\ldots\sigma_{i_p} - \frac{\lambda\sqrt{N}}{k}\langle\sigma, \hat{\mathbf{v}}_0\rangle^k, \tag{2.1}$$

where $\hat{\mathbf{v}}_0 := \mathbf{v}_0 \in S^{N-1}$. Then,

$$\mathrm{Crt}_N(M,E) = \sum_{\sigma\in S^{N-1},\nabla f(\sigma)=0} \mathbf{1}_{\{\sigma\cdot\hat{\mathbf{v}}_0\in M\}}\cdot\mathbf{1}_{\{f(\sigma)/\sqrt{N}\in E\}},$$

and the corresponding number of critical points of index $l = 0,\ldots,N-1$ is given by

$$\mathrm{Crt}_{N,l}(M,E) = \sum_{\sigma\in S^{N-1},\nabla f(\sigma)=0} \mathbf{1}_{\{\sigma\cdot\hat{\mathbf{v}}_0\in M\}}\cdot\mathbf{1}_{\{f(\sigma)/\sqrt{N}\in E\}}\mathbf{1}_{\{i(\nabla^2 f)=l\}}.$$

Proposition 2.1. Let M and E be the same as in Theorem 1.1. Then, for any $l \geq 1$,

$$\limsup_{N\to\infty}\frac{1}{N}\log\mathbb{E}[Crt_{N,l}(M,E)] < \sup_{m\in\bar{M},x\in\bar{E}} S_{p,k}(x,m). \tag{2.2}$$

We postpone the proof of this proposition to the end of this section. We now show how to prove Theorem 1.5.

Proof of Theorem 1.5. Since f is a Morse function almost surely, let

$$\mathcal{S}_N(M,E) := \Big\{\sigma\in S^{N-1} : f(\sigma)\in\sqrt{N}E, \sigma\cdot\hat{\mathbf{v}}_0\in M\Big\}.$$

Then, its Euler characteristic $\phi(\mathcal{S}_N(M,E))$ can be computed in terms of the numbers of critical points as follows:

$$\phi(\mathcal{S}_N(M,E)) = \sum_{l=0}^{N-1}(-1)^{l+2}Crt_{N,l}(M,E).$$

Using Proposition 2.1, as $N\to\infty$, we have

$$\mathbb{E}[Crt_{N,0}(M,E)] \sim \mathbb{E}[Crt_N(M,E)] \sim \mathbb{E}[\phi(\mathcal{S}_N(M,E))]. \tag{2.3}$$

Therefore, it suffices to compute the asymptotic of the mean Euler characteristic $\mathbb{E}[\phi(\mathcal{S}_N(M,E))]$. Applying Formula 12.4.4 in Ref. 1 (see also Eq. (6.22) in Ref. 4), we have

$$\mathbb{E}[\phi(\mathcal{S}_N(M,E))] = \int_{\sigma\cdot v_0\in M}\mathbb{E}\Big[\det\nabla^2 f(\sigma)\cdot 1_{\{f\in\sqrt{N}E\}}\mid \nabla f=0\Big]\phi_{\nabla f(\sigma)}(0)d\sigma, \tag{2.4}$$

where $\phi_{\nabla f(\sigma)}(0)$ is the density of $\nabla f(\sigma)$ at 0.

Let $\omega_{N-2} = \frac{(N-1)\pi^{\frac{N-1}{2}}}{\Gamma(\frac{N+1}{2})}$ be the surface area of the $N-2$ dimensional unit sphere; using the data in Lemma A.1, we get

$$\begin{aligned}&\int_{\sigma\cdot v_0\in M}\mathbb{E}\Big[\det\nabla^2 f(\sigma)\cdot 1_{\{f\in\sqrt{N}E\}}\mid \nabla f=0\Big]\phi_{\nabla f(\sigma)}(0)d\sigma\\ &=\frac{\omega_{N-2}\sqrt{N}}{(2\pi)^{\frac{N}{2}}p^{\frac{N-1}{2}}}\int_M\int_E\left(1-m^2\right)^{\frac{N-3}{2}}\exp\left(-\frac{N}{2}\left(\lambda^2m^{2k-2}\left(1-m^2\right)/p+\left(x+\lambda m^k/k\right)^2\right)\right)G_N(x,m)dxdm,\end{aligned}$$

where

$$G_N(x,m) = (2(N-1)p(p-1))^{\frac{N-1}{2}}\mathbb{E}\left[\det\left(W_{N-1}-\frac{\sqrt{N}}{\sqrt{N-1}}\theta e_{N-1}e_{N-1}^T-\frac{\sqrt{N}}{\sqrt{N-1}}yI_{N-1}\right)\right], \tag{2.5}$$

$$\theta=\theta(m) := \frac{\lambda(k-1)m^{k-2}\left(1-m^2\right)}{\sqrt{2p(p-1)}}, \tag{2.6}$$

and

$$y=y(x,m) := \sqrt{\frac{p}{2(p-1)}}\left(x-(1/p-1/k)\lambda m^k\right). \tag{2.7}$$

Using Lemmas A.3 and A.6, we can express G using Hermite polynomials (see Definition A.2),

$$G_N(x,m) = (-1)^{N-1}(p(p-1)/2)^{\frac{N-1}{2}}\left(h_{N-1}\left(\sqrt{N}y\right)+2\sqrt{N}\theta h_{N-2}\left(\sqrt{N}y\right)\right).$$

It follows that $\mathbb{E}[Crt_N(M,E)] = I + II$, where

$$\begin{aligned}I &= \frac{\omega_{N-2}\sqrt{N}(-1)^{N-1}(p-1)^{\frac{N-1}{2}}}{2^{N-\frac{1}{2}}\pi^{\frac{N}{2}}}\int_M dm\left(1-m^2\right)^{\frac{N-3}{2}}\\ &\times\int_E dx\,\exp\left(-\frac{N}{2}\left(\lambda^2m^{2k-2}\left(1-m^2\right)/p+\left(x+\lambda m^k/k\right)^2\right)\right)h_{N-1}\left(\sqrt{N}y\right)\end{aligned}$$

and

$$\begin{aligned}II &= \frac{\omega_{N-2}N(-1)^{N-1}(p-1)^{\frac{N-1}{2}}}{2^{N-\frac{3}{2}}\pi^{\frac{N}{2}}}\int_M dm\left(1-m^2\right)^{\frac{N-3}{2}}\\ &\times\int_E dx\,\exp\left(-\frac{N}{2}\left(\lambda^2m^{2k-2}\left(1-m^2\right)/p+\left(x+\lambda m^k/k\right)^2\right)\right)\theta h_{N-2}\left(\sqrt{N}y\right).\end{aligned}$$

We consider term I first. Using the Hermite function ϕ_{N-1} (see Definition A.2),

$$I = \frac{\omega_{N-2} N^{\frac{1}{2}} (-1)^{N-1} \left(2^{N-1}(N-1)!\sqrt{\pi}\right)^{\frac{1}{2}}}{2^{N-\frac{1}{2}} \pi^{\frac{N}{2}} \sqrt{p-1}} \int_M dm \left(1-m^2\right)^{-\frac{3}{2}}$$
$$\times \int_E dx \, \exp\left(\frac{N}{2}\left(y^2 + \log\left(1-m^2\right)(p-1) - \lambda^2 m^{2k-2}\left(1-m^2\right)/p - \left(x + \lambda m^k/k\right)^2\right)\right) \phi_{N-1}\left(\sqrt{N} y\right). \tag{2.8}$$

For any $m \in M$ and $x \in E$, $y = y(x,m) < -\sqrt{2}$. Using Lemma A.8 and letting $\tilde{h}(y) = \frac{\sqrt{2} h(y)}{\sqrt{y^2-2}-y}$, as $N \to \infty$, we have

$$I \sim \frac{\omega_{N-2} N^{\frac{1}{2}} \left(2^{N-1}(N-1)!\sqrt{\pi}\right)^{\frac{1}{2}}}{\sqrt{4\pi\sqrt{2N}} 2^{N-\frac{1}{2}} \pi^{\frac{N}{2}} \sqrt{p-1}} \int_M dm \left(1-m^2\right)^{-\frac{3}{2}}$$
$$\times \int_E dx \, \exp\left(\frac{N}{2}\left(y^2 + \log\left(1-m^2\right)(p-1) - \lambda^2 m^{2k-2}\left(1-m^2\right)/p - \left(x+\lambda m^k/k\right)^2 - 2I_1(-y)\right)\right) \tilde{h}(y)$$
$$\sim \frac{(N-1)\pi^{\frac{N-1}{2}} N^{\frac{1}{4}} \sqrt{(N-1)!}}{\Gamma\left(\frac{N+1}{2}\right) 2^{\frac{N}{2}+\frac{5}{4}} \pi^{\frac{N}{2}+\frac{1}{4}} \sqrt{p-1}} \int_M \int_E \left(1-m^2\right)^{-\frac{3}{2}} \tilde{h}(y(x,m)) \exp\left(N S_{p,k}(x,m)\right) dx dm$$
$$\sim \frac{N}{2\sqrt{2\pi}\sqrt{p-1}} \int_M \int_E \left(1-m^2\right)^{-\frac{3}{2}} \tilde{h}(y(x,m)) \exp\left(N S_{p,k}(x,m)\right) dx dm.$$

Substituting y for x, as $N \to \infty$, we have

$$I \sim \frac{N}{2\pi\sqrt{p}} \int_M \int_{\tilde{E}_m} \left(1-m^2\right)^{-\frac{3}{2}} \tilde{h}(y) \exp\left(N \tilde{S}_{p,k}(m,y)\right) dy dm, \tag{2.9}$$

where $\tilde{E}_m := \sqrt{\frac{p}{2(p-1)}}\left(E - \lambda m^k\left(\frac{1}{p} - \frac{1}{k}\right)\right)$.

Similarly,

$$II = \frac{2\lambda\omega_{N-2} N\sqrt{N-1}(-1)^{N-1}(p-1)^{\frac{N-2}{2}}(k-1)}{2^N \pi^{\frac{N}{2}} \sqrt{p}\sqrt{N-1}}$$
$$\times \int_M \int_E \left(1-m^2\right)^{\frac{N-1}{2}} m^{k-2} \exp\left(-\frac{N}{2}\left(\frac{\lambda^2}{p} m^{2k-2}\left(1-m^2\right) + \left(x + \frac{\lambda m^k}{k}\right)^2\right)\right) h_{N-2}\left(\sqrt{N} y\right) dx dm$$
$$= \frac{\lambda\omega_{N-2} N\sqrt{(N-2)!}\pi^{\frac{1}{4}}(k-1)(-1)^{N-1}}{\sqrt{p(2\pi)^N}} \int_M dm m^{k-2} (p-1)^{\frac{N-2}{2}} \left(1-m^2\right)^{\frac{N-1}{2}}$$
$$\times \int_E dx \, \exp\left(-\frac{N}{2}\left(\frac{\lambda^2}{p} m^{2k-2}\left(1-m^2\right) + \left(x + \frac{\lambda m^k}{k}\right)^2 - y^2\right)\right) \phi_{N-2}\left(\sqrt{N} y\right).$$

When N is large enough, $\sqrt{\frac{N}{N-1}} y < -\sqrt{2}$, so by Lemma A.8,

$$\phi_{N-2}\left(\sqrt{N} y\right) \sim (-1)^{N-2} \frac{e^{-(N-1) I_1\left(-\sqrt{\frac{N}{N-1}} y\right)}}{\sqrt{4\pi\sqrt{2(N-1)}}} \tilde{h}\left(\sqrt{\frac{N}{N-1}} y\right).$$

Therefore, as $N \to \infty$,

$$II \sim -\frac{\lambda N(k-1)}{2\sqrt{2p}\pi} \int_M dm (p-1)^{\frac{N-2}{2}} \left(1-m^2\right)^{\frac{N-1}{2}} m^{k-2} \int_E dx \tilde{h}\left(\sqrt{\frac{N}{N-1}} y(x)\right) L_N(m,x),$$

where

$$L_N(m,x) = \exp\left(-\frac{N}{2}\left(\frac{\lambda^2}{p} m^{2k-2}\left(1-m^2\right) + \left(x + \lambda m^k/k\right)^2 - y^2 + \frac{2(N-1)}{N} I_1\left(-\sqrt{\frac{N}{N-1}} y\right)\right)\right).$$

Let $z = \sqrt{\frac{N}{N-1}} y(x)$; then,

$$\tilde{L}_N(m,z) := \left((p-1)\left(1-m^2\right)\right)^{\frac{N-1}{2}} L_N(m,x) = \exp\left(-(N-1)\tilde{S}_{p,k}(m,z)\right) J_N(m,z),$$

where

$$J_N(m,z) = \exp\left(-\frac{\lambda^2}{2p^2} m^{2k-2}\left(p\left(1-m^2\right)+m^2\right) + \frac{\lambda m^k z}{p}\sqrt{\frac{2(p-1)}{p}}(N-1)\left(1-\sqrt{\frac{N}{N-1}}\right)\right).$$

Substituting z for x, as $N \to \infty$, we have

$$II \sim -\frac{\lambda(N-1)(k-1)}{2p\pi}\int_M dm m^{k-2}\int_{\tilde{E}_{m,N}} dz \tilde{h}(z)\exp\left((N-1)\tilde{S}_{p,k}(m,z)\right)J_N(m,z),$$

where $\tilde{E}_{m,N} = \sqrt{\frac{N}{N-1}}\tilde{E}_m$.

Since $\tilde{E}_m$ is precompact, $1_{\tilde{E}_{m,N}}(z)J_N(m,z)$ converges to $1_{\tilde{E}_m}(z)J(m,z)$ uniformly on $m \in M$ and $z \in \tilde{E}_m$. Therefore, as $N \to \infty$,

$$II \sim -\frac{\lambda(N-1)(k-1)}{2p\pi}\int_M dm m^{k-2}\int_{\tilde{E}_m} dy \tilde{h}(y)\exp\left((N-1)\tilde{S}_{p,k}(m,y)\right)J(m,y). \tag{2.10}$$

Combining Eqs. (2.9) and (2.10), we get Eq. (1.9) from the Laplace method. □

We end this section with the Proof of Proposition 2.1.

Proof of Proposition 2.1. Applying the Kac–Rice formula (Theorem 12.1.1 in Ref. 1) to the Hamiltonian (2.1), we have

$$\mathbb{E}[Crt_{N,\ell}(M,E)] = \int_{\sigma\cdot v_0\in M}\mathbb{E}\left[\left|\det \nabla^2 f(\sigma)\right|\cdot 1_{\{f\in\sqrt{N}E, i(\nabla^2 f)=l\}} \mid \nabla f = 0\right]\phi_{\nabla f(\sigma)}(0)d\sigma. \tag{2.11}$$

Set

$$A_{N,\ell}(\sigma) = \mathbb{E}\left[\left|\det \nabla^2 f(\sigma)\right|\cdot 1_{\{f\in\sqrt{N}E, i(\nabla^2 f)=l\}} \mid \nabla f = 0\right].$$

We now show that for any σ with $\sigma \cdot v_0 \in M \subseteq (m_\lambda, 1)$ and E satisfying (1.4), for $\ell \geq 1$, we have

$$\frac{1}{N}\log A_{N,\ell}(\sigma) = o\left(\frac{1}{N}\log A_{N,0}(\sigma)\right)$$

uniformly in σ. Looking at (2.5)–(2.7) and using Lemma A.1, it suffices to show there exists $\eta > 0$, independent of $y \in E$, such that

$$\frac{\mathbb{E}\left[|\det(M-\theta e_{N-1}e_{N-1}^T - yI_N)|\mathbf{1}\{\lambda_\ell \leq y\}\right]}{\mathbb{E}\left[|\det(M-\theta e_{N-1}e_{N-1}^T - yI_N)|\mathbf{1}\{\lambda_0 \leq y\}\right]} \leq \exp(-N\eta). \tag{2.12}$$

Let L_N be the empirical spectral measure of the matrix $M - \theta e_{N-1}e_{N-1}^T$, $\lambda_\ell(\theta)$ be its ℓth smallest eigenvalue, and μ denote the semi-circle law. For $\delta > 0$, consider the event

$$B_N(\delta) = \left\{\left|\int \log|x-y|dL_N(x) - \int \log|x-y|d\mu(x)\right| > \delta\right\}.$$

By Ref. 6 and an application of eigenvalue interlacement, there exist $\epsilon > 0$ so that for all N sufficiently large,

$$\mathbb{P}(B_N(\delta)) \leq e^{-\epsilon N^2}.$$

Now, writing

$$|\det(M-\theta e_{N-1}e_{N-1}^T - yI_N)| = \int \log|x-y|dL_N(x),$$

note that there exists $C > 0$ so that $\mathbb{E}\int \log|x-y|dL_N(x) \leq \exp(CN)$ and a positive constant C' such that for N large enough,

$$\begin{aligned}\mathbb{E}\left[|\det(M-\theta e_{N-1}e_{N-1}^T - yI_N)|\mathbf{1}\{\lambda_\ell \leq y\}\right] &= \mathbb{E}\left[|\det(M-\theta e_{N-1}e_{N-1}^T - yI_N)|\mathbf{1}\{\lambda_\ell \leq y\}\mathbf{1}\{B_N(\delta)\}\right] \\ &\quad + \mathbb{E}\left[|\det(M-\theta e_{N-1}e_{N-1}^T - yI_N)|\mathbf{1}\{\lambda_\ell \leq y\}\mathbf{1}\{B_N^c(\delta)\}\right] \\ &\leq e^{N(\delta+\int \log|x-y|d\mu)}\mathbb{P}(\lambda_\ell \leq y) + e^{-\epsilon N^2 + C'N}.\end{aligned} \tag{2.13}$$

At the same time, we also have the lower bound

$$\mathbb{E}\left[|\det(M-\theta e_{N-1}e_{N-1}^T - yI_N)|\mathbf{1}\{\lambda_0 \leq y\}\right] \geq e^{N(-\delta+\int \log|x-y|d\mu)}\mathbb{P}(\lambda_0 \leq y)(1-e^{-\epsilon N^2}). \tag{2.14}$$

Thus, for N large enough, we obtain for all $y \in E$,

$$\frac{\mathbb{E}\big[|\det(M-\theta e_{N-1}e_{N-1}^T - yI_N)|\mathbf{1}\{\lambda_\ell \le y\}\big]}{\mathbb{E}\big[|\det(M-\theta e_{N-1}e_{N-1}^T - yI_N)|\mathbf{1}\{\lambda_0 \le y\}\big]} \le \exp\Big(N2\delta - \frac{\epsilon}{2}N^2\Big)\frac{\mathbb{P}(\lambda_\ell \le y)}{\mathbb{P}(\lambda_0 \le y)}. \tag{2.15}$$

On the other hand, by our choice of E, there exists $\kappa > 0$ such that $y < -\sqrt{2} + \kappa$ for all $y \in E$. By an application of the large deviation principle for the extreme eigenvalues of rank one perturbation of GOE (Ref. 8, Theorem 2.13), there exists $\rho > 0$ depending only on κ so that

$$\frac{\mathbb{P}(\lambda_\ell \le y)}{\mathbb{P}(\lambda_0 \le y)} \le \exp(-N\rho). \tag{2.16}$$

Plugging (2.16) into (2.15), we find that there exists $\eta > 0$ so that for N large enough and all $y \in E$, bound (2.12) is satisfied. This completes the proof of the proposition. □

III. THE MEAN NUMBER OF DEEP MINIMA

In this section, we prove Theorems 1.1 and 1.3.

Proof of Theorem 1.1. If $x_*(\lambda) \in E$ and $m_*(\lambda) \in M$, we prove Theorem 1.1 by deriving an explicit formula for the constant $C(\lambda,p,k)$ as follows:

$$\lim_{N\to\infty}\mathbb{E}[Crt_N(M,E)] = C(\lambda,p,k) = \frac{\sqrt{2}\Big(\sqrt{p}(1-m_*)^{-\frac{3}{2}} - \lambda(k-1)m_*^{k-2}\Big)h(y_*)}{p\Big(\sqrt{y_*^2-2}-y_*\Big)\sqrt{|\partial_{yy}\tilde{S}_{p,k}(m_*,y_*)g''(m_*)|}}, \tag{3.1}$$

where

$$y_* := y_*(m_*), y_*(m) = \frac{\lambda m^k}{2\sqrt{p}}\frac{p-2}{\sqrt{2(p-1)}} - \frac{p}{\sqrt{2(p-1)}}\sqrt{\Big(\frac{\lambda m^k}{2\sqrt{p}}\Big)^2 + 1}, \tag{3.2}$$

$g(m) = \tilde{S}_{p,k}(m, y_*(m))$, and $h(\cdot)$ and $I_1(\cdot)$ are defined in Theorem A.7.

Otherwise, we show that $\tilde{S}_{p,k}(m_o,y_o) < 0$ in Eq. (1.9). Therefore,

$$\lim_{N\to\infty}\mathbb{E}[Crt_N(M,E)] = 0.$$

A direct computation gives

$$-\partial_y\tilde{S}_{p,k} = \frac{p-2}{p}y + \frac{\lambda m^k}{p}\sqrt{\frac{2(p-1)}{p}} - \sqrt{y^2-2}. \tag{3.3}$$

Let $A = \frac{p-2}{p}$ and $B = \frac{\lambda m^k}{p}\sqrt{\frac{2(p-1)}{p}}$; then, $-\partial_y\tilde{S}_{p,k} = Ay + B - \sqrt{y^2-2}$.

When $m \le m_\lambda$, $-\frac{B}{A} \ge -\sqrt{2}$. Therefore, $\partial_y\tilde{S}_{p,k} \ge 0$ on $(-\infty,-\sqrt{2})$ and $\tilde{S}_{p,k}(m,\cdot)$ is increasing.

When $m \ge m_\lambda$, $-\frac{B}{A} \le -\sqrt{2}$. Therefore, $\tilde{S}_{p,k}(m,\cdot)$ has a unique maximum in $(-\infty,-\sqrt{2})$ at

$$y_*(m) = \frac{AB - \sqrt{2+B^2-2A^2}}{1-A^2} = \frac{\lambda m^k}{2\sqrt{p}}\frac{p-2}{\sqrt{2(p-1)}} - \frac{p}{\sqrt{2(p-1)}}\sqrt{\Big(\frac{\lambda m^k}{2\sqrt{p}}\Big)^2 + 1}. \tag{3.4}$$

We then define $g(m) = \tilde{S}_{p,k}(m,y_*(m))$. Plugging Eq. (3.4) into $\tilde{S}_{p,k}(m,\cdot)$, we have

$$g(m) = l(v) = \frac{1}{2}\log(1-m^2) + (1-2/m^2)v^2 + v\sqrt{v^2+1} - \log\Big(v + \sqrt{v^2+1}\Big),$$

where $v = \frac{\lambda m^k}{2\sqrt{p}}$.

We compute $l'(v) = 2v\Big(1 - \frac{2}{m^2} + \frac{\sqrt{v^2+1}}{v}\Big)$. Since $\frac{\sqrt{v^2+1}}{v}$ is decreasing, $l'(v) = 0$ on $(0,\infty)$ if and only if $v = \frac{m^2}{2\sqrt{1-m^2}}$ and

$$l_{\max} = l\Big(\frac{m^2}{2\sqrt{1-m^2}}\Big) = 0.$$

The maximum is achieved if and only if Eq. (1.3) holds, i.e., $\frac{\lambda m^k}{\sqrt{p}} = \frac{m^2}{\sqrt{1-m^2}}$.

As mentioned in Lemma B.1, when $\lambda \geq \max\{\lambda^{(1)}(p,k), \lambda^{(2)}(p,k)\}$, there is a unique solution m_* of Eq. (1.3) such that $m_* \geq m_\lambda$, and the above computation implies that

$$\sup_{m_\lambda \leq m < 1, y < -\sqrt{2}} \tilde{S}_{p,k}(m,y) = \sup_{m_\lambda \leq m < 1} \tilde{S}_{p,k}(m, y_*(m)) = \tilde{S}_{p,k}(m_*, y_*(m_*)) = 0. \tag{3.5}$$

Recall that $\tilde{E}_m \subset (-\infty, -\sqrt{2})$ and $y_*(m) \in \tilde{E}_m$ for any m. By Laplace's method, as $N \to \infty$,

$$\begin{aligned} I &\sim \frac{N}{2\pi\sqrt{p}} \int_M \sqrt{\frac{2\pi}{N|\partial_{yy}\tilde{S}_{p,k}(m, y_*(m))|}} (1-m^2)^{-\frac{3}{2}} \frac{\sqrt{2}h(y_*(m))}{\sqrt{y_*(m)-2} - y_*(m)} \exp(Ng(m))dm \\ &\sim \frac{1}{\sqrt{p|\partial_{yy}\tilde{S}_{p,k}(m_*, y_*) g''(m_*)|}} (1-m_*^2)^{-\frac{3}{2}} \frac{\sqrt{2}h(y_*)}{\sqrt{y_*-2} - y_*} \exp(Ng(m_*)) \\ &= \frac{\sqrt{2}h(y_*)(1-m_*^2)^{-\frac{3}{2}}}{(\sqrt{y_*-2} - y_*)\sqrt{p|\partial_{yy}\tilde{S}_{p,k}(m_*, y_*) g''(m_*)|}}. \end{aligned} \tag{3.6}$$

Similarly, we apply the Laplace method to II and get

$$\begin{aligned} II &\sim -\frac{\lambda(N-1)(k-1)}{2p\pi} \int_M dm \times \sqrt{\frac{2\pi}{(N-1)|\partial_{yy}\tilde{S}_{p,k}(m, y_*(m))|}} m^{k-2} \frac{\sqrt{2}h(y_*(m))}{\sqrt{y_*(m)-2} - y_*(m)} \exp((N-1)g(m)) J(m, y_*(m)) \\ &\sim \frac{\lambda(k-1)}{p} \frac{m_*^{k-2}\sqrt{2}h(y_*)\exp((N-1)g(m_*))J(m_*, y_*)}{(\sqrt{y_*-2} - y_*)\sqrt{|\partial_{yy}\tilde{S}_{p,k}(m_*, y_*) g''(m_*)|}} \\ &= \frac{\sqrt{2}\lambda(k-1)m_*^{k-2}h(y_*)J(m_*, y_*)}{p(\sqrt{y_*-2} - y_*)\sqrt{|\partial_{yy}\tilde{S}_{p,k}(m_*, y_*) g''(m_*)|}}. \end{aligned} \tag{3.7}$$

It remains to show that $J(m_*, y_*) = 1$. This is obtained as follows. Using Eq. (1.3),

$$\begin{aligned} &-\frac{\lambda^2}{2p^2} m_*^{2k-2}(p(1-m_*^2) + m_*^2) + \frac{\lambda m_*^k y_*}{2p}\sqrt{\frac{2(p-1)}{p}} \\ &= \frac{\lambda m_*^k}{2p}(p(1-m_*^2) + m_*^2)\left(\frac{1}{\sqrt{p(1-m_*^2)}} - \frac{1}{\sqrt{p(1-m_*^2)}}\right) = 0. \end{aligned}$$

Combining this with Eqs. (3.6) and (3.7), we get Eq. (3.1). □

Proof of Theorem 1.3. Since m_* satisfies Eq. (1.3), as $\lambda \to \infty$, $m_* \to 1$. When $k = 1$ or 2, this can be obtained directly from Eq. (B1). When $k \geq 3$, by the implicit differentiation theorem, denoting $m_*' := \frac{d}{d\lambda} m_*(\lambda)$, we have

$$\begin{aligned} &\frac{m_*^k}{\sqrt{p}} + \frac{\lambda k m_*^{k-1}}{\sqrt{p}} m_*' = \left(2m_*(1-m_*^2)^{-\frac{1}{2}} + m_*^3(1-m_*^2)^{-\frac{3}{2}}\right) m_*' \\ &\Rightarrow \frac{m_*^k}{\sqrt{p}} = m_*'(1-m_*^2)^{-\frac{3}{2}} m_*((k-1)m_*^2 - (k-2)). \end{aligned}$$

Since $m_* > \sqrt{\frac{k-2}{k-1}}$, so $(k-1)m_*^2 - (k-2) > 0$ and thus $m_*' > 0$. Therefore, as $\lambda \to \infty$, $m_*(\lambda) \uparrow 1$.

We also have

$$y_*(m_*) = \frac{p-2}{2\sqrt{2(p-1)}} m_*^2 \left(1-m_*^2\right)^{-\frac{1}{2}} - \frac{p}{\sqrt{2(p-1)}} \sqrt{\frac{m_*^4}{4(1-m_*^2)}+1}$$
$$= \frac{1}{\sqrt{2(p-1)}} \left(1-m_*^2\right)^{-\frac{1}{2}} \left(\frac{p-2}{2} m_*^2 - \frac{p}{2}\left(2-m_*^2\right)\right)$$
$$= \frac{1}{\sqrt{2(p-1)}} \left(1-m_*^2\right)^{-\frac{1}{2}} \left((p-1)m_*^2 - p\right). \tag{3.8}$$

Therefore, as $\lambda \to \infty$,

$$y_*(m_*) \sim -\frac{1}{\sqrt{2(p-1)}} \left(1-m_*^2\right)^{-\frac{1}{2}}. \tag{3.9}$$

Note that

$$y_*{}'(m) = \frac{p-2}{2\sqrt{2p(p-1)}} \lambda k m^{k-1} - \frac{p}{\sqrt{2(p-1)}} \left(\left(\frac{\lambda m^k}{2\sqrt{p}}\right)^2 + 1\right)^{-\frac{1}{2}} \frac{\lambda m^k}{2\sqrt{p}} \cdot \frac{\lambda k m^{k-1}}{2\sqrt{p}}.$$

Using Eq. (1.3), we have as $\lambda \to \infty$,

$$y_*'(m_*) = \frac{k(p-2)}{2\sqrt{2(p-1)}} m_* \left(1-m_*^2\right)^{-\frac{1}{2}} - \frac{kp}{2\sqrt{2(p-1)}} \frac{m_*^3}{2-m_*^2} \left(1-m_*^2\right)^{-\frac{1}{2}}$$
$$= \frac{k}{\sqrt{2(p-1)}} \left(1-m_*^2\right)^{-\frac{1}{2}} \frac{(p-2)m_* - (p-1)m_*^3}{2-m_*^2}$$
$$\sim -\frac{k}{\sqrt{2(p-1)}} \left(1-m_*^2\right)^{-\frac{1}{2}} \tag{3.10}$$

and

$$\partial_{yy}\tilde{S}_{p,k}(m_*, y_*) = \frac{2-p}{p} + y_*\left(y_*^2 - 2\right)^{-\frac{1}{2}}$$
$$\sim -\frac{2(p-1)}{p}. \tag{3.11}$$

We also compute for $k \geq 1$,

$$\partial_m \tilde{S}_{p,k} = -\frac{m}{1-m^2} - \frac{\lambda k m^{k-1}}{p} \sqrt{\frac{2(p-1)}{p}} y - \frac{\lambda^2 (k-1)}{p} m^{2k-3} + \frac{\lambda^2 k(p-1)}{p^2} m^{2k-1}.$$

For $k \geq 2$,

$$\partial_{mm}\tilde{S}_{p,k} = -\frac{1+m^2}{(1-m^2)^2} - \frac{\lambda k(k-1)m^{k-2}}{p} \sqrt{\frac{2(p-1)}{p}} y - \frac{\lambda^2 (k-1)(2k-3)}{p} m^{2k-4}$$
$$+ \frac{\lambda^2 k(p-1)(2k-1)}{p^2} m^{2k-2},$$

$\partial_{mm}\tilde{S}_{p,1} = \frac{1+m^2}{-(1-m^2)^2} + \frac{(p-1)\lambda^2}{p^2}$, and for $k \geq 2$,

$$\partial_{my}\tilde{S}_{p,k} = -\frac{\lambda k m^{k-1}}{p} \sqrt{\frac{2(p-1)}{p}}.$$

Using Eqs. (1.3) and (3.9), we have as $\lambda \to \infty$,

$$\partial_{mm}\tilde{S}_{p,k}(m_*,y_*(m_*)) = -\left(1+m_*^2\right)\left(1-m_*^2\right)^{-2} - \frac{k(k-1)}{p}\left(1-m_*^2\right)^{-1}\left((p-1)m_*^2 - p\right)$$
$$- (k-1)(2k-3)\left(1-m_*^2\right)^{-1} - \frac{k(p-1)(2k-1)m_*^2}{p}\left(1-m_*^2\right)^{-1}$$
$$\sim -2\left(1-m_*^2\right)^{-2} \tag{3.12}$$

and

$$\partial_{my}\tilde{S}_{p,k}(m_*,y_*(m_*)) = \frac{k\sqrt{2(p-1)}}{p}m_*\left(1-m_*^2\right)^{-\frac{1}{2}} \sim \frac{k\sqrt{2(p-1)}}{p}\left(1-m_*^2\right)^{-\frac{1}{2}}. \tag{3.13}$$

Recall that $g(m) = \tilde{S}_{p,k}(m,y_*(m))$, so

$$g'' = \partial_{mm}\tilde{S}_{p,k} + 2\partial_{my}\tilde{S}_{p,k}\cdot y_*' + \partial_{yy}\left(y_*'\right)^2 + \partial_y\tilde{S}_{p,k}\cdot y''. \tag{3.14}$$

Note that $\partial_y\tilde{S}_{p,k}(m_*,y_*(m_*)) = 0$; using Eqs. (3.12), (3.13), (3.9), and (3.10), we know that as $\lambda\to\infty$,

$$g''(m_*) \sim -2\left(1-m_*^2\right)^{-2} + \frac{2k\sqrt{2(p-1)}}{p}\left(1-m_*^2\right)^{-\frac{1}{2}}\cdot\left(\frac{-k}{\sqrt{2(p-1)}}\right)\left(1-m_*^2\right)^{-\frac{1}{2}}$$
$$+ \frac{2(p-1)}{p}\frac{k^2}{2(p-1)}\left(1-m_*^2\right)^{-1}$$
$$\sim -2\left(1-m_*^2\right)^{-2}. \tag{3.15}$$

From the definition of $h(\cdot)$ in Theorem A.7 and Eq. (3.9), it is easy to see

$$\lim_{\lambda\to\infty} h(y_*(m_*)) = 2.$$

To sum up, as $\lambda\to\infty$,

$$C(\lambda,p,k) \sim 2\frac{\left(\sqrt{p}\left(1-m_*^2\right)^{-\frac{3}{2}} - \lambda(k-1)m_*^{k-2}\right)}{p\sqrt{2}y_*(m_*)\sqrt{\left|\partial_{yy}\tilde{S}_{p,k}(m_*,y_*)g''(m_*)\right|}}$$
$$\sim \frac{2\left(\sqrt{p}\left(1-m_*^2\right)^{-\frac{3}{2}} - (k-1)\sqrt{p}\left(1-m_*^2\right)^{-\frac{1}{2}}\right)}{-p\sqrt{2}y_*(m_*)\sqrt{2\frac{p-1}{p}\cdot 2\left(1-m_*^2\right)^{-2}}}$$
$$\sim \frac{1}{\sqrt{p-1}}\left(1-m_*^2\right)^{-\frac{1}{2}}\cdot\sqrt{(p-1)}\left(1-m_*^2\right)^{\frac{1}{2}} \tag{3.16}$$
$$= 1. \tag{3.17}$$

□

IV. LIMITING GROUND STATE ENERGY

In this section, we prove Theorem 1.4. The proof relies on the following two propositions whose proofs are presented after the Proof of Theorem 1.4.

Proposition 4.1. For any $m\in(0,1)$,

$$\limsup_{N\to\infty}\mathbb{E}\left[\frac{1}{N}\min_{\sigma\in S^{N-1}(\sqrt{N})}H_N(\sigma)\right] \le -\frac{\lambda m^k}{k} - \sqrt{p(1-m^2)}. \tag{4.1}$$

Proposition 4.2.

$$\liminf_{N\to\infty}\frac{1}{N}\min_{\sigma\in S^{N-1}(\sqrt{N})}H_N(\sigma) \ge \lambda m_*^k\left(\frac{1}{2}-\frac{1}{k}\right) - \sqrt{\frac{\lambda^2 m_*^{2k}}{4}+p}\ a.s. \tag{4.2}$$

Proof of Theorem 1.4 assuming Propositions 4.2 and 4.1. By the Gaussian concentration inequality and Borel–Canteli lemma,

$$\lim_{N\to\infty}\frac{1}{N}\min_{\sigma\in S^{N-1}(\sqrt{N})}H_N(\sigma)=\lim_{N\to\infty}\mathbb{E}\left[\frac{1}{N}\min_{\sigma\in S^{N-1}(\sqrt{N})}H_N(\sigma)\right]\ a.s.$$

Therefore, it suffices to show

$$-\frac{\lambda m_*^k}{k}-\sqrt{p(1-m_*^2)}=\lambda m_*^k\left(\frac{1}{2}-\frac{1}{k}\right)-\sqrt{\frac{\lambda^2 m_*^{2k}}{4}+p}. \tag{4.3}$$

Using Eq. (1.3),

$$\lambda m_*^k/2-\sqrt{\frac{\lambda^2 m_*^{2k}}{4}+p}=\frac{\sqrt{p}m_*^2}{2\sqrt{1-m_*^2}}-\sqrt{\frac{p(m_*^4-4m_*^2+4)}{4(1-m_*^2)}}$$
$$=\frac{\sqrt{p}\left(m_*^2+m_*^2-2\right)}{2\sqrt{1-m_*^2}}=-\sqrt{p(1-m_*^2)}.$$

Therefore, Eq. (4.3) holds. □

Now, we prove Propositions 4.1 and 4.2.

Proof of Proposition 4.1. For any $m\in(0,1)$,

$$\frac{1}{N}\min_{\sigma\in S^{N-1}(\sqrt{N}),\sigma\cdot\mathbf{v}_0=m}H_N(\sigma)\geq\frac{1}{N}\min_{\sigma\in S^{N-1}(\sqrt{N})}H_N(\sigma),$$

so

$$\mathbb{E}\left[\frac{1}{N}\min_{\sigma\in S^{N-1}(\sqrt{N}),\sigma\cdot\mathbf{v}_0=m}H_N(\sigma)\right]\geq\mathbb{E}\left[\frac{1}{N}\min_{\sigma\in S^{N-1}(\sqrt{N})}H_N(\sigma)\right].$$

Since H_N is isotropic, without loss of generality, we assume $\mathbf{v}_0=\sqrt{N}e_N$ and, then, conditional on $\sigma_N=\sqrt{N}m$,

$$H_N(\sigma)=-\lambda N\frac{m^k}{k}-\sqrt{N}J_{NN\ldots N}m^p-\frac{1}{N^{\frac{p-1-l}{2}}}\sum_{l=0}^{p-1}m^l\sum_{i_{k_j}=N,1\leq i_k\leq N-1,k\neq k_j,j\in[l]}J_{i_1,i_2,\ldots,i_p}\frac{\sigma_{i_1}\sigma_{i_2}\ldots\sigma_{i_p}}{\sigma_{i_{k_1}}\sigma_{i_{k_2}}\ldots\sigma_{i_{k_l}}}.$$

Since for different sets of $\left(i_{k_j}\right)_{j=1}^{l}$, $J_{i_1,i_2,\ldots,i_p}$ are i.i.d, so

$$H_N(\sigma)\overset{(d)}{=}-\lambda N\frac{m^k}{k}-\sqrt{N}J_{NN\ldots N}m^p$$
$$-\sqrt{N}\sum_{l=0}^{p-1}\binom{p}{l}^{\frac{1}{2}}m^l\left(1-m^2\right)^{\frac{p-l}{2}}\sum_{1\leq i_1,i_2,\ldots i_{p-l}\leq N-1}g_{i_1,i_2,\ldots,i_{p-1}}\hat{\sigma}_{i_1}\hat{\sigma}_{i_2}\ldots\hat{\sigma}_{i_{p-l}},$$

where $\hat{\sigma}_k=\sigma_k/\sqrt{N(1-m^2)}, k\in[N-1]$.

Note that $\sum_{k=1}^{N-1}\hat{\sigma}_k^2=1$; therefore,

$$-\sqrt{N}\sum_{l=0}^{p-1}\binom{p-1}{l}^{\frac{1}{2}}m^l\left(1-m^2\right)^{\frac{p-l}{2}}\sum_{1\leq i_1,i_2,\ldots i_{p-l}\leq N-1}g_{i_1,i_2,\ldots,i_{p-1}}\hat{\sigma}_{i_1}\hat{\sigma}_{i_2}\ldots\hat{\sigma}_{i_{p-l}}$$

is a spherical mixed p-spin model with mixture

$$\xi(x)=\sum_{l=0}^{p-1}\binom{p}{l}m^{2l}\left(1-m^2\right)^{p-l}x^{p-l}=\left(m^2+\left(1-m^2\right)x\right)^p-m^{2p}.$$

By Proposition 1 in Ref. 12 (see also Theorem 1.10 in Ref. 19),

$$\mathbb{E}\left[\frac{1}{N}\min_{\sigma\in S^{N-1}(\sqrt{N}),\sigma\cdot\mathbf{v}_0=m}H_N(\sigma)\right]=-\frac{\lambda m^k}{k}-\sqrt{\xi'(1)}=-\frac{\lambda m^k}{k}-\sqrt{p(1-m^2)}.$$ □

Recall that in this paper, we reserve the symbol x_* for the right-hand side of Eq. (4.2), i.e.,

$$x_* := \lambda m_*^k\left(\frac{1}{2}-\frac{1}{k}\right)-\sqrt{\frac{\lambda^2 m_*^{2k}}{4}+p}.$$

The key to proving Proposition 4.2 is to identify the point at which 0, the supremum of the complexity function $S_{p,k}$, is attained. The following proposition shows that the point lies in the high-latitude region of the sphere.

Proposition 4.3. There exists a constant $\tilde{\lambda}_c = \tilde{\lambda}_c(p,k)$ such that for any $\lambda \geq \tilde{\lambda}_c$, $M = (0, m_\lambda)$,

$$\sup_{m\in\tilde{M}} S_{p,k}(m, x_*) < 0.$$

Proof. Using the correspondence between y_* and x_* [see Eq. (1.8)], we have

$$f(m) := \tilde{S}_{p,k}(m, y_*) = S_{p,k}(m, x_*).$$

We will first show that $f(m)$ has at most one critical point on M, and if it exists, it must be a local minimum of f; then, we use the results on the pure p-spin model from Ref. 3 and Theorem 1.1 to show that $f(0) < 0$ and $f(m_\lambda) < 0$, thus deriving $\sup_{m\in\tilde{M}} S_{p,k}(m, x_*) = \sup_{m\in\tilde{M}} f(m) < 0$.

A direct computation shows that

$$\begin{aligned} f'(m) &= -\frac{m}{1-m^2} - \frac{\lambda k m^{k-1}}{p}\sqrt{\frac{2(p-1)}{p}}y_* - \frac{\lambda^2(k-1)m^{2k-3}}{p} + \frac{(p-1)\lambda^2 k m^{2k-1}}{p^2} \\ &= -\frac{m}{1-m^2} f_1(m), \end{aligned}$$

where

$$f_1(m) = 1 + \frac{\lambda k m^{k-2}(1-m^2)}{p}\sqrt{\frac{2(p-1)}{p}}y_* + \frac{\lambda^2(k-1)m^{2k-4}(1-m^2)}{p} - \frac{(p-1)\lambda^2 k m^{2k-2}(1-m^2)}{p^2}.$$

Let $u := u(\lambda, m) := \lambda m^{k-2}(1-m^2)$.

Case I: If $k \leq p$,

$$f_1(m) = 1 + \frac{k}{p}\sqrt{\frac{2(p-1)}{p}}y_* u + \frac{k-1}{p}u^2 + \frac{\lambda^2 m^{2k-2}(1-m^2)}{p^2}(k-p) \leq f_2(u),$$

where

$$f_2(u) = 1 + \frac{k}{p}\sqrt{\frac{2(p-1)}{p}}y_* u + \frac{k-1}{p}u^2.$$

The larger zero of f_2 is

$$u_* := \frac{1}{(k-1)\sqrt{2p}}\left(-k\sqrt{p-1}y_* + \sqrt{k^2(p-1)y_*^2 - 2(k-1)p^2}\right).$$

Recall that for fixed λ, $u = \lambda m^{k-2}(1-m^2)$, which is increasing with respect to m over $\left[0, \sqrt{\frac{k-2}{k}}\right]$, so when $\lambda \geq p\sqrt{\frac{p-2}{p-1}}\left(\frac{k}{k-2}\right)^{\frac{k}{2}}$,

$$u_{\max} = u(m_\lambda) = \lambda^{\frac{2}{k}}\left(\frac{p\sqrt{p-2}}{\sqrt{p-1}}\right)^{\frac{k-2}{k}}(1-m_\lambda^2).$$

Therefore,

$$u_{\max} \sim \lambda^{\frac{2}{k}}\left(\frac{p\sqrt{p-2}}{\sqrt{p-1}}\right)^{\frac{k-2}{k}} \quad \text{as} \quad \lambda\to\infty. \tag{4.4}$$

Note that

$$y_* = \sqrt{\frac{p}{2(p-1)}}\left(x_* - \lambda m_*^k\left(\frac{1}{p}-\frac{1}{k}\right)\right) = -\frac{\lambda m_*^k}{\sqrt{2p(p-1)}} - p\sqrt{\frac{1-m_*^2}{2(p-1)}}.$$

When $\lambda \to \infty$, it is observed from Eq. (1.3) that $\lim_{\lambda\to\infty} m_*(\lambda) = 1$, so

$$y_* \sim -\frac{\lambda}{\sqrt{2p(p-1)}} \quad \text{as} \quad \lambda \to \infty,$$

and thus,

$$u_* \sim -\frac{\lambda k}{p(k-1)} \quad \text{as} \quad \lambda \to \infty. \tag{4.5}$$

Combining Eqs. (4.4) and (4.5), there exists $\tilde{\lambda}_c > 0$ such that if $\lambda \geq \tilde{\lambda}_c$, $u_{\max} < u_*$, so $f_1(m)$ crosses the m-axis at most once over $[0, m_\lambda]$. Note that $f_1(0) = 1 > 0$, and it is continuous on $[0, m_\lambda]$, so $f'(m) < 0$ when m is small and it crosses the m-axis at most once over $[0, m_\lambda]$.

Case II: If $k > p$, then when $\lambda \geq 2^k(p-2)\sqrt{\frac{p}{p-1}}$, $m_\lambda \leq \frac{1}{2}$. Therefore, for any $m \leq m_\lambda$,

$$\lambda m^{2k-2}(1-m^2) \leq u^2.$$

Then, we have $f_1(m) \leq f_3(u)$, where

$$f_3(u) = 1 + \frac{k}{p}\sqrt{\frac{2(p-1)}{p}}y_* u + \left(\frac{k(p+1)-2p}{p^2}\right)u^2.$$

The same argument in case I also applies to case II, and we derive the same conclusion that $f'(m) < 0$ when m is small and it crosses the m-axis at most once over $[0, m_\lambda]$.

This implies that

$$\sup_{m\in\tilde{M}} f(m) = \max\{f(0), f(m_\lambda)\}. \tag{4.6}$$

Note that $f(0) = \Phi_p(y_*)$, where $\Phi_p(\cdot)$ is the annealed complexity of the p-spin spherical spin glass model; see Theorem 2.8 in Ref. 3. It is known that $\Phi_p(\cdot)$ is an increasing function on $\left(-\infty, -2\sqrt{\frac{p-1}{p}}\right)$ and $\lim_{y\to-\infty}\Phi_p(y) = -\infty$, so when λ is large enough so that y_* is smaller than the limiting ground state energy of the p-spin spherical spin glass model, which is the unique zero of $\Phi_p(\cdot)$ on $\left(-\infty, -2\sqrt{\frac{p-1}{p}}\right)$,

$$f(0) = \Phi_p(y_*) < 0. \tag{4.7}$$

As to $f(m_\lambda)$, we know from Theorem 1.1 [more specifically, Eq. (3.5)] that when $\lambda \geq \tilde{\lambda}$,

$$f(m_\lambda) \leq \sup_{m\geq m_\lambda} f(m) \leq 0.$$

Combining this with Eqs. (4.6) and (4.7), we prove this proposition. □

Proof of Proposition 4.2. For any $\epsilon > 0$, let $M = [0,1]$ and $E = (-\infty, x_* - \epsilon)$. It is shown in Theorem 1.1 that for fixed $m < 1$, $\tilde{S}_{p,k}(m,\cdot)$ is increasing on $(-\infty, -x_* - \epsilon)$. Combining this with Proposition 4.3, we see that

$$\begin{aligned}
\limsup_{N\to\infty}\frac{1}{N}\log\mathbb{E}[Crt_N(M,E)] &= \sup_{m\in\tilde{M},x\in\tilde{E}} S_{p,k}(m,x) \\
&\leq \sup_{m\in\tilde{M}}\tilde{S}_{p,k}(m,y_*) \\
&\leq \max\{\sup_{m\in[0,m_\lambda]}\tilde{S}_{p,k}(m,y_*), \sup_{m\in[m_\lambda,1]}\tilde{S}_{p,k}(m,y_*)\} \\
&< 0.
\end{aligned}$$

Therefore, by Markov inequality,

$$P\left(\frac{1}{N}\min_{\sigma\in S^{N-1}(\sqrt{N})} H_N(\sigma) \leq x_* - \epsilon\right) \leq P(Crt_N(M,E) \geq 1) \leq \mathbb{E}[Crt_N(M,E)].$$

Then, Eq. (4.2) follows from the Borel–Cantali lemma. □

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

DATA AVAILABILITY

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

APPENDIX A: COVARIANCE COMPUTATIONS AND SOME FORMULAS FROM RANDOM MATRIX THEORY

In this appendix, we derive the random matrices appearing in the Kac–Rice computation in Sec. II and summarize a series of tools that we use in random matrix theory.

Lemma A.1. Let $f : S^{N-1} \to \mathbb{R}$ be defined in Eq. (2.1). Without loss of generality, we set $\sigma = e_N, \hat{\mathbf{v}}_0 = me_N + \sqrt{1-m^2}e_{N-1}$; then,

$$\mathbb{E}[f(\sigma)] = -\lambda\sqrt{N}m^k/k, \quad \mathrm{Var}(f(\sigma)) = 1,$$

$$\mathbb{E}[\nabla f(\sigma)] = -\sqrt{N}\lambda m^{k-1}\sqrt{1-m^2}e_{N-1},$$

$$Cov(f(\sigma), \nabla_i f(\sigma)) = Cov\left(\nabla^2_{jk} f(\sigma), \nabla_i f(\sigma)\right) = 0 \quad \text{for} \quad i, j, \ k = 1, 2, \ldots, N-1,$$

$$Cov\left(\nabla^2 f, f\right) = -pI_{N-1},$$

$$Cov(\nabla f, \nabla f) = pI_{N-1},$$

$$Cov\left(\nabla^2_{ij} f, \nabla^2_{kl} f\right) = p(p-1)\left(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}\right) + p^2\delta_{ij}\delta_{kl} \quad \text{for} \quad i, j, k, \ l = 1, 2, \ldots, N-1.$$

Denote by $\mathbb{E}_A$ and Cov_A the expectation and covariance conditional on the event A; then,

$$\mathbb{E}_{\nabla f(\sigma)=0}[f(\sigma)] = \mathbb{E}[f(\sigma)],$$

$$\mathbb{E}_{\nabla f(\sigma)=0}\left[\nabla^2 f(\sigma)\right] = \mathbb{E}\left[\nabla^2 f(\sigma)\right],$$

$$\mathbb{E}\left[\nabla^2 f(\sigma)\right] = -\sqrt{N}\lambda(k-1)\left(1-m^2\right)m^{k-2}e_{N-1}e_{N-1}^T + \sqrt{N}\lambda m^k I_{N-1},$$

$$\mathbb{E}_{f=\sqrt{N}x}\left[\nabla^2 f(\sigma)\right] = -\lambda\sqrt{N}(k-1)m^{k-2}\left(1-m^2\right)e_{N-1}e_{N-1}^T - pI_{N-1}\left(\sqrt{N}x + \frac{\lambda\sqrt{N}m^k}{k}\right) + \lambda\sqrt{N}m^k I_{N-1},$$

$$Cov_{f=\sqrt{N}x}\left(\nabla^2_{ij} f(\sigma), \nabla^2_{kl} f(\sigma)\right) = p(p-1)\left(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}\right) \quad \text{for} \quad i, j, k, \ l = 1, 2, \ldots, N-1.$$

From Lemma A.1, conditional on $\nabla f(\sigma) = 0$, $f(\sigma) = \sqrt{N}x$,

$$\nabla^2 f(\sigma) \overset{d}{=} \sqrt{2(N-1)p(p-1)}W_{N-1} - \lambda\sqrt{N}(k-1)m^{k-2}\left(1-m^2\right)e_{N-1}e_{N-1}^T + \sqrt{N}I_{N-1}\left(-px + \left(1-\frac{p}{k}\right)\lambda m^k\right). \tag{A1}$$

Definition A.2. For $N \in \mathbb{N}$, denote the following:

- Hermite polynomials $h_N(x) = e^{x^2}\left(-\frac{d}{dx}\right)^N e^{-x^2}$.
- Hermite functions $\phi_N(x) = \left(2^N N!\sqrt{\pi}\right)^{-\frac{1}{2}} h_N(x) e^{-\frac{x^2}{2}}$.

Lemma A.3 (Lemma 3 in Ref. 2 and Corollary 11.6.3 in Ref. 1).

$$\mathbb{E}[\det(W_{N-1} - xI_{N-1})] = 2^{1-N}(N-1)^{\frac{1-N}{2}}(-1)^{N-1} h_{N-1}\left(\sqrt{N-1}x\right).$$

Using Eq. (1.8) in Ref. 13, we obtain the following proposition, which is useful for expressing determinants in terms of Hermite polynomials.

Lemma A.4.

$$\mathbb{E}\left[\det\left(W_{N-1} - f e_{N-1} e_{N-1}^T + sI_{N-1}\right)\right] = \left(\frac{-i}{\sqrt{N-1}}\right)^{N-1} \pi^{-\frac{1}{2}} e^{(N-1)s^2} \int_{\mathbb{R}} e^{-y^2}\left(y^{N-1} - i\sqrt{N-1} f y^{N-2}\right) e^{2\sqrt{N-1}iys} dy. \tag{A2}$$

Remark A.5. Setting $f = 0$, one can easily recover Lemma A.3 using the Fourier transform.

Lemma A.6.

$$\mathbb{E}\left[\det\left(W_{N-1} - f e_{N-1} e_{N-1}^T + sI_{N-1}\right)\right] = \mathbb{E}[\det(W_{N-1} + sI_{N-1})] - f\left(\frac{N-2}{N-1}\right)^{\frac{N-2}{2}} \mathbb{E}\left[\det\left(W_{N-2} + \sqrt{\frac{N-1}{N-2}} sI_{N-2}\right)\right].$$

Proof. Combine Lemmas A.3 and A.4. □

Theorem A.7 (Plancherel–Rotach asymptotics). There exists $\delta_0 > 0$ such that for any $\delta \in (0, \delta_0)$, we have uniformly in $x \in \left(-\infty, -\sqrt{2} - \delta\right)$,

$$\phi_N\left(\sqrt{N}x\right) = (-1)^{N-1} \frac{e^{-NI_1(-x)}}{\sqrt{4\pi\sqrt{2N}}} h(x)\left(1 + \mathcal{O}\left(N^{-1}\right)\right),$$

where

$$h(x) = \left|\frac{x - \sqrt{2}}{x + \sqrt{2}}\right|^{\frac{1}{4}} + \left|\frac{x + \sqrt{2}}{x - \sqrt{2}}\right|^{\frac{1}{4}}$$

and

$$I_1(x) = \int_{\sqrt{2}}^{x} \sqrt{t^2 - 2} dt.$$

Proof. This lemma is the same as Lemma 7.1 in Ref. 3 and Lemma 5 in Ref. 2. □

From Theorem A.7, we derive the following lemma that we need in the Proof of Theorem 1.5.

Lemma A.8. There exists $\delta_0 > 0$ such that for any $\delta \in (0, \delta_0)$, we have uniformly in $y \in \left(-\infty, -\sqrt{2} - \delta\right)$,

$$\phi_{N-1}\left(\sqrt{N}y\right) = (-1)^{N-2} \frac{e^{-NI_1(-y)}}{\sqrt{2\pi\sqrt{2N}}} \frac{h(y)}{\sqrt{y^2 - 2} - y}\left(1 + \mathcal{O}\left(N^{-1}\right)\right).$$

Proof. Note that $\lim_{N\to\infty} \sqrt{\frac{N}{N-1}} y = y$, so for N large enough and $y \in \left(-\infty, -\sqrt{2} - \delta\right)$, we use Theorem A.7 to derive

$$\begin{aligned}\phi_{N-1}\left(\sqrt{N}y\right) &= \phi_{N-1}\left(\sqrt{N-1}\frac{\sqrt{N}}{\sqrt{N-1}}y\right)\\ &= (-1)^{N-2}\frac{e^{-(N-1)I_1\left(-\frac{\sqrt{N}}{\sqrt{N-1}}y\right)}}{\sqrt{4\pi\sqrt{2N}}}h\left(\frac{\sqrt{N}}{\sqrt{N-1}}y\right)\left(1+\mathcal{O}\left(N^{-1}\right)\right)\\ &= (-1)^{N-2}\frac{e^{-NI_1(-y)}e^{I_1(-y)}e^{-(N-1)\int_{-y}^{-\sqrt{\frac{N}{N-1}}y}\sqrt{t^2-2}dt}}{\sqrt{4\pi\sqrt{2N}}}h\left(\frac{\sqrt{N}}{\sqrt{N-1}}y\right)\left(1+\mathcal{O}\left(N^{-1}\right)\right).\end{aligned} \tag{A3}$$

Since

$$h\left(\frac{\sqrt{N}}{\sqrt{N-1}}y\right) = h(y)\left(1+\mathcal{O}\left(N^{-1}\right)\right) \tag{A4}$$

and

$$\begin{aligned}e^{I_1(-y)}e^{-(N-1)\int_{-y}^{-\sqrt{\frac{N}{N-1}}y}\sqrt{t^2-2}dt} &= e^{I_1(-y)}e^{\frac{y\sqrt{y^2-2}}{2}\left(1+\mathcal{O}\left(N^{-1}\right)\right)}\\ &= \frac{\sqrt{2}}{\sqrt{y^2-2}-y}\left(1+\mathcal{O}\left(N^{-1}\right)\right)\end{aligned} \tag{A5}$$

uniformly for $y \in \left(-\infty, -\sqrt{2}-\delta\right)$, combining Eqs. (A3)–(A5), we prove Lemma A.8. □

APPENDIX B: MATHEMATICAL ANALYSIS ON THRESHOLDS

In this section, we discuss the existence and values of $\lambda_{tr}, \lambda^{(1)}(p,k)$ and $\lambda^{(2)}(p,k)$.

Lemma B.1. If $k \le 2$, then Eq. (1.3) has a unique solution on $(0,1]$ for any $\lambda > 0$. If $k > 2$, then Eq. (1.3) has a solution if and only if $\lambda \ge \sqrt{p\frac{(k-1)^{k-1}}{(k-2)^{k-2}}}$. In particular, when $k > 2$ and $\lambda \ge \sqrt{p\frac{(k-1)^{k-1}}{(k-2)^{k-2}}}$, the solution on $\left[\sqrt{\frac{k-2}{k-1}}, 1\right)$ is unique.

Proof. When $k = 1, 2$, m_* can be solved explicitly from Eq. (1.3) as follows:

$$m_*(\lambda) = \begin{cases}\sqrt{\dfrac{\lambda^2}{p}\Big/\left(1+\dfrac{\lambda^2}{p}\right)}, & k=1,\\ \sqrt{1-\dfrac{p}{\lambda^2}}, & k=2.\end{cases} \tag{B1}$$

When $k > 2$, let $g(m) = \frac{\lambda^2 m^{2k-4}(1-m^2)}{p}$. We compute

$$g'(m) = \frac{2(k-2)\lambda^2}{p}\left(1-\frac{k-1}{k-2}m^2\right)m^{2k-5}.$$

Therefore, Eq. (1.3) has a solution on $(0,1]$ if and only if $g_{\max} = g\left(\sqrt{\frac{k-2}{k-1}}\right) \ge 1$ if and only if $\lambda \ge \sqrt{p\frac{(k-1)^{k-1}}{(k-2)^{k-2}}}$. Moreover, when the solution exists on $\left[\sqrt{\frac{k-2}{k-1}}, 1\right]$, it is unique. □

In the next lemma, we study the values of $\lambda^{(2)}(p,k)$.

Lemma B.2. For any integers $p \ge 3, k \ge 1$, there exists $\lambda^{(2)} := \lambda^{(2)}(p,k) > 0$ such that $m_*(\lambda) < m_\lambda$ when $\lambda < \lambda^{(2)}$ and $m_*(\lambda) \ge m_\lambda$ when $\lambda \ge \lambda^{(2)}$.

Proof. When $k = 1$, from Eqs. (B1) and (1.2), we derive $m_*(\lambda) \le m_\lambda$ if and only if $\left(\lambda^2/p\right)^2 - \frac{(p-2)^2}{p-1}\frac{\lambda^2}{p} - \frac{(p-2)^2}{p-1} < 0$ if and only if

$$\lambda \le \sqrt{\frac{p\left(\frac{(p-2)^2}{p-1} + \sqrt{\frac{(p-2)^4}{(p-1)^2} + \frac{4(p-2)^2}{p-1}}\right)}{2}}.$$

When $k = 2$, again from Eqs. (B1) and (1.2), we derive $m_*(\lambda) \leq m_\lambda$ if and only if $\frac{p}{\lambda^2} + \frac{p-2}{\sqrt{p-1}}\frac{\sqrt{p}}{\lambda} - 1 > 0$ if and only if

$$\lambda \leq \frac{2\sqrt{p}}{\frac{2-p}{\sqrt{p-1}} + \sqrt{4 + \frac{(p-2)^2}{p-1}}}.$$

When $k > 2$, the existence of $\lambda^{(2)}$ is guaranteed by the fact that $m_*(\lambda)$ increases to 1 and m_λ decreases to 0. □

Lemma B.3. For any $p \geq 3$ and $k \geq 1$, $\lambda^{(1)}(p,k) = \lambda^{(2)}(p,k)$ if and only if $p \leq k$.

Proof. When $k = 1, 2$, using Lemmas B.1 and B.2, we have $\lambda^{(1)} = 0 < \lambda^{(2)}$. Therefore, from now on, we assume $p, k \geq 3$. When $p \leq k$,

$$\left(\frac{m_{\lambda^{(1)}(p,k)}}{\sqrt{\frac{k-2}{k-1}}}\right)^{2k} = \frac{(p-2)^2(k-1)}{(k-2)^2(p-1)} \leq 1.$$

By Lemma B.1, $m_{\lambda^{(1)}(p,k)} \leq \sqrt{\frac{k-2}{k-1}} \leq m_*$ and thus $\lambda^{(2)} = \lambda^{(1)}$.

When $p > k$, it suffices to show $g\big(m_{\lambda^{(1)}(p,k)}\big) < 1$. A direct computation gives

$$g\big(m_{\lambda^{(1)}(p,k)}\big) = \frac{(p-2)^2}{p-1} \cdot \frac{1 - m^2_{\lambda^{(1)}(p,k)}}{m^4_{\lambda^{(1)}(p,k)}}.$$

Since $f(x) = \frac{1-x}{x^2}$ is increasing on $[0,1]$ and for each fixed p,

$$\log\big(m_{\lambda^{(1)}(p,k)}\big) = \frac{1}{k}\left(\log\left(\frac{p-2}{\sqrt{p-1}}\right) - \log\left(\frac{k-2}{\sqrt{k-1}}\right)\right) + \log\left(\frac{k-2}{k-1}\right)$$

is decreasing for $k < p$.

Therefore,

$$g\big(m_{\lambda^{(1)}(p,k)}\big) < g\big(m_{\lambda^{(1)}(p,p)}\big) = \frac{(p-2)^2}{p-1} \cdot \frac{1 - \frac{p-2}{p-1}}{\left(\frac{p-2}{p-1}\right)^2} = 1.$$

□

REFERENCES

[1] Adler, R. J. and Taylor, J. E., *Random Fields and Geometry*, Springer Monographs in Mathematics (Springer-Verlag, New York, 2007).

[2] Auffinger, A. and Ben Arous, G., "Complexity of random smooth functions on the high-dimensional sphere," Ann. Probab. **41**(6), 4214–4247 (2013).

[3] Auffinger, A., Ben Arous, G., and Černý, J., "Random matrices and complexity of spin glasses," Commun. Pure Appl. Math. **66**(2), 165–201 (2013).

[4] Azaïs, J.-M. and Wschebor, M., *Level Sets and Extrema of Random Processes and Fields* (John Wiley & Sons, Inc., Hoboken, NJ, 2009).

[5] Belius, D., Černý, J., Nakajima, S., and Schmidt, M., "Triviality of the geometry of mixed p-spin spherical Hamiltonians with external field," J. Stat. Phys. **186**, 12 (2021).

[6] Ben Arous, G. and Guionnet, A., "Large deviations for Wigner's law and Voiculescu's non-commutative entropy," Probab. Theory Relat. Fields **108**(4), 517–542 (1997).

[7] Ben Arous, G., Mei, S., Montanari, A., and Nica, M., "The landscape of the spiked tensor model," Commun. Pure Appl. Math. **72**(11), 2282–2330 (2019).

[8] Benaych-Georges, F., Guionnet, A., and Maida, M., "Large deviations of the extreme eigenvalues of random deformations of matrices," Probab. Theory Relat. Fields **154**(3), 703–751 (2012).

[9] Cavagna, A., Garrahan, J. P., and Giardina, I., "Quenched complexity of the mean-field p-spin spherical model with external magnetic field," J. Phys. A: Math. Gen. **32**(5), 711–723 (1999).

[10] Cavagna, A., Giardina, I., and Parisi, G., "Stationary points of the Thouless-Anderson-Palmer free energy," Phys. Rev. B **57**, 11251–11257 (1998).

[11] Chen, W.-K., "Phase transition in the spiked random tensor with Rademacher prior," Ann. Stat. **47**(5), 2734–2756 (2019).

[12] Chen, W.-K. and Sen, A., "Parisi formula, disorder chaos and fluctuation for the ground state energy in the spherical mixed p-spin models," Commun. Math. Phys. **350**, 129 (2015).

[13] Desrosiers, P. and Liu, D.-Z., "Scaling limits of correlations of characteristic polynomials for the Gaussian β-ensemble with external source," Int. Math. Res. Not. **12**, 3751–3781 (2015).

[14] Fyodorov, Y. V., "High-dimensional random fields and random matrix theory," Markov Processes Relat. Fields **21**, 483–518 (2013).

[15] Fyodorov, Y. V., "Topology trivialization transition in random non-gradient autonomous ODEs on a sphere," J. Stat. Mech.: Theory Exp. **2016**(12), 124003.

[16] Fyodorov, Y. V. and Le Doussal, P., "Topology trivialization and large deviations for the minimum in the simplest random optimization," J. Stat. Phys. **154**(1), 466–490 (2014).

[17] Gillin, P. and Sherrington, D., "p > 2 spin glasses with first-order ferromagnetic transitions," J. Phys. A: Math. Gen. **33**(16), 3081–3091 (2000).
[18] Jagannath, A., Lopatto, P., and Miolane, L., "Statistical thresholds for tensor PCA," Ann. Appl. Probab. **30**(4), 1910–1933 (2020).
[19] Jagannath, A. and Tobasco, I., "Low temperature asymptotics of spherical mean field spin glasses," Commun. Math. Phys. **352**(3), 979–1017 (2017).
[20] Maillard, A., Ben Arous, G., and Biroli, G., "Landscape complexity for the empirical risk of generalized linear models," in *Proceedings of the First Mathematical and Scientific Machine Learning Conference, Proceedings of Machine Learning Research (PMLR)*, edited by Lu, J. and Ward, R. (Princeton University, Princeton, NJ, 2020), Vol. 107, pp. 287–327.
[21] Mannelli, S. S., Krzakala, F., Urbani, P., and Zdeborová, L., "Passed and spurious: Descent algorithms and local minima in spiked matrix-tensor models," in *PMLR*, edited by Chaudhuri, K. and Salakhutdinov, R. (Proceedings of Machine Learning Research, 2019), Vol. 97, pp. 4333–4342.
[22] Montanari, A. and Richard, E., "A statistical model for tensor PCA," in *Proceedings of the 27th International Conference on Neural Information Processing Systems* (online, 2014), Vol. 2, pp. 2897–2905.
[23] Perry, A., Wein, A. S., and Bandeira, A. S., "Statistical limits of spiked tensor models," Ann. Inst. Henri Poincare: Probab. Stat. **56**(1), 230–264 (2020).
[24] Ros, V., Ben Arous, G., Biroli, G., and Cammarota, C., "Complex energy landscapes in spiked-tensor and simple glassy models: Ruggedness, arrangements of local minima, and phase transitions," Phys. Rev. X **9**, 011003 (2019).
[25] Sarao Mannelli, S., Biroli, G., Cammarota, C., Krzakala, F., Urbani, P., and Zdeborová, L., "Marvels and pitfalls of the Langevin algorithm in noisy high-dimensional inference," Phys. Rev. X **10**, 011057 (2020).
[26] Sarao Mannelli, S., Biroli, G., Cammarota, C., Krzakala, F., and Zdeborová, L., "Who is afraid of big bad minima? Analysis of gradient-flow in spiked matrix-tensor models," in *Advances in Neural Information Processing Systems* (Curran Associates, Inc., 2019), Vol. 32, pp. 8679–8689.