Non-invariant random matrices and landscape complexity

by

Benjamin M^{c}Kenna

A dissertation submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

Department of Mathematics

Courant Institute of Mathematical Sciences

New York University

May 2021

Gérard Ben Arous

Paul Bourgade

Acknowledgements

First and foremost, I am extremely grateful to Gérard Ben Arous and Paul Bourgade for their joint supervision of this thesis. I am very fortunate to have first learned landscape complexity from Gérard, who asks the right questions years in advance, who gives powerful motivations for research questions, and who always warmly encourages me to prove a better result. I am equally fortunate to have first learned random matrices from Paul, who has been exceptionally generous with his time and his professional advice, on questions macroscopic, mesoscopic, and microscopic, and whose optimism is an inspiration.

I would like to thank Alice Guionnet, László Erdős, and Torben Krüger for stimulating mathematical discussions. I am also lucky to have a great thesis committee in S. R. S. Varadhan, Percy Deift, and Sylvia Serfaty.

My time at Courant has been extremely enjoyable, not least because of terrific topics classes in the hands of Yuri Bakhtin, Eyal Lubetzky, Chuck Newman, and Ofer Zeitouni, and those already mentioned. I'm grateful to Tim Austin and Sinan Güntürk for their always-cheerful advice on navigating graduate school. Marla Wolf helped teach me how to teach science, and Stephen S. Hall helped teach me how to write about it (in March/April 2020, no less). I'm also thankful for the administrative support of Brittany Shields, Michelle Shin, Adam Staszczuk, Betty Tsang, and Melissa Vacca, to all of whom I surely asked an above-average number of questions.

I am indebted to Tuca Auffinger, who has been enormously supportive of me since we met almost nine years ago: in particular for setting me on the path from my very first day of college; for suggesting that I go to Courant and that I talk to Paul and Gérard; and for inviting me to

Northwestern to talk about my work.

As an undergraduate I also benefitted from two extremely enjoyable summers at Peter May's REU, where I got my first exposure to real probability, and from the kind support of Robert Fefferman and John Boller.

I am grateful to the many friends and colleagues who have made the last five years go by in the blink of an eye: Lucas Benigni and my beautiful PCMI shirt, Jeanne Boursier, Hong-Bin Chen, Guillaume Dubach, Reza Gheissari, Lisa Hartung, Elias Hess-Childs, Tim Kunisky, Mihai Nica, Luke Peilen, and Zhe Wang, among others. I want to especially thank Krishnan Mody (equivalently, the baking team at Lafayette), who has been the best officemate and friend a guy could ask for, even or especially when "officemate" is more a mental designation than a physical one. The backend of this thesis was adapted from a template of Shannon M. Locke. Mark Shapiro has helped keep me from doing math all the time. My apartment-mates – at various times Ben Sulser, Jordan Hisel, Michael Ingram, and Kat Currier – have made going back to the apartment a pleasure. One day, I hope to leave the apartment again.

Finally, I could not have done any of this without the support and love of my family: my mother, the first person I ever met with a doctorate from NYU, for passing down her love of music and for encouraging me to do what makes me happy; my father, for his optimism, calm life advice, and food adventures; my brother Henry, for his cheerleading and his confidence in me; my grandmother, for her constant encouragement, her coffee, and her felicitous 1963 choice to live within walking distance of NYU; and my partner Ellen, for her impeccable professional and writing advice, for putting many miles on the road, and for her patience and kindness during the many hours I spend writing things in a notebook and crossing them out.

Abstract

This thesis considers Hermitian random matrices that are non-invariant, meaning they have few symmetries. First, we study the asymptotics of their determinants as the matrix size diverges, and the effects of these on the geometry of high-dimensional random functions. Second, we study large deviations of their extremal eigenvalues.

The classical Kac-Rice formula provides a bridge between random geometry and random matrices. It relates the expected number of critical points of a real-valued random function on $\mathbb{R}^N$, on the one hand, to the expected absolute value of the determinant of an $N \times N$ random matrix, on the other hand. In the large-N limit, it thus reduces counts of critical points to determinant asymptotics for large random matrices. We are especially interested in "non-invariant" random functions, meaning functions with few (distributional) symmetries. For such functions, the corresponding random matrices are also "non-invariant." In particular, large-deviations principles, crucial in past studies of highly symmetric random functions, are usually not available.

We start by identifying simple criteria that yield exponential asymptotics of these large determinants. These criteria are satisfied by a wide variety of matrix models, including Wigner matrices and sample covariance matrices with near-optimal $2+\varepsilon$ finite moments; Erdős-Rényi matrices with near-optimal sparsity $p \geqslant N^{\varepsilon}/N$; band matrices with any polynomial bandwidth $W \geqslant N^{\varepsilon}$; and Gaussian matrices with a variance profile.

Then we use our determinant asymptotics and the Kac-Rice formula to study the exponential count of critical points, called "landscape complexity," for three models of random functions. First, we consider the "elastic manifold," a classic model in statistical physics of particle configurations

with self-interactions in a disordered environment, where we confirm formulas of Fyodorov and Le Doussal on a phase transition between the simple and glassy regimes. Second, we introduce a new, general signal-plus-noise model, where we find a surprising threshold distinguishing positive vs. zero complexity, with universal near-critical behavior close to this threshold. Third, we characterize complexity of bipartite spherical spin glasses, a sandbox model of spin glasses beyond the classical mean-field setup.

Finally, we study additively deformed Wigner matrices with certain sub-Gaussian entries. We establish a large-deviations principle for their extremal eigenvalues, building on recent techniques of Guionnet-Husson and Guionnet-Maïda.

Contents

List of Figures

List of Appendices

Chapter 1

Introduction

1.1 Random determinants: Our results

In the first part of this thesis, we study the quantity $\mathbb{E}[|\det(H_N)|]$, where H_N is an $N \times N$ real-symmetric random matrix, and identify its leading-order exponential asymptotics for a wide variety of matrices H_N. These asymptotics can be guessed as follows: Writing $\hat{\mu}_{H_N} = \frac{1}{N}\sum_{i=1}^{N} \delta_{\lambda_i(H_N)}$ for the empirical spectral measure of H_N, one observes

$$\mathbb{E}[|\det(H_N)|] = \mathbb{E}[e^{N\int \log|\lambda| \hat{\mu}_{H_N}(\mathrm{d}\lambda)}]$$

(interpreted appropriately if an eigenvalue vanishes). If $\hat{\mu}_{H_N}$ concentrates about some deterministic limiting measure μ_∞, this suggests asymptotics of the form

$$\lim_{N\to\infty} \frac{1}{N} \log \mathbb{E}[|\det(H_N)|] = \int \log|\lambda| \mu_\infty(\mathrm{d}\lambda), \tag{1.1.1}$$

although to prove this, one needs to understand the logarithmic singularity and how it interacts with the concentration of the empirical spectral measure. We are able to overcome these issues, and obtain the following result.

Theorem 1.1.1. *(Chapter 2 in this thesis, from Ben Arous-Bourgade-M. [35]) We find*

$$\lim_{N\to\infty} \frac{1}{N} \log \mathbb{E}[|\det(H_N)|] = \int \log|\lambda| \mu_\infty(\lambda)\, \mathrm{d}\lambda$$

when H_N is one of the following:

- *a Wigner matrix with (near-optimal) $2+\varepsilon$ finite moments (and μ_∞ is the semicircle law ρ_{sc} with density*

  $$\rho_{\mathrm{sc}}(\mathrm{d}x) = \frac{\sqrt{(4-x^2)_+}}{2\pi}\, \mathrm{d}x$$

  with respect to Lebesgue measure),

- *a sample covariance matrix with (near-optimal) $2+\varepsilon$ finite moments and some regularity assumptions on the entries (and μ_∞ is the Marčenko-Pastur law),*

- *the adjacency matrix of an Erdős-Rényi random graph with parameter $p \geqslant N^\varepsilon/N$ (and μ_∞ is the semicircle law),*

- *a one-dimensional band matrix with any polynomial bandwidth $W \geqslant N^\varepsilon$ and some regularity assumptions on the entries (and μ_∞ is the semicircle law), or*

- *the free-addition model $A_N + O_N B_N O_N^T$, where A_N and B_N are real, deterministic, and diagonal, and O_N is Haar orthogonal (and μ_∞ is the free convolution of the limiting empirical spectral measures of A_N and B_N).*

In the Gaussian case, we find

$$\lim_{N\to\infty} \left(\frac{1}{N} \log \mathbb{E}[|\det(H_N)|] - \int \log|\lambda| \mu_N(\lambda)\, \mathrm{d}\lambda \right) = 0 \tag{1.1.2}$$

when H_N is one of the following:

- *Gaussian with a mean-field (co)variance profile and/or a mean, or*

- *Gaussian with zero blocks in special places (for example,* $A_N + \begin{pmatrix} W_1 & 0 \\ 0 & W_2 \end{pmatrix}$ *where* A_N *is deterministic and the* W_i*'s are independent matrices from the Gaussian Orthogonal Ensemble (GOE), meaning the entries* $(W_i)_{jk}$ *are independent up to symmetry with* $(W_i)_{jk} \sim \mathcal{N}(0, \frac{1+\delta_{jk}}{N})$*).*

In these Gaussian cases, the measures μ_N *arise from the theory of the so-called* Matrix Dyson Equation *(MDE), developed by Erdős and collaborators in papers such as [5, 6]. In most natural cases, they have a weak limit* μ_∞ *for which* (1.1.1) *holds.*

Versions of (1.1.1) have appeared in the literature, mostly in cases when H_N enjoys properties we could describe as "invariance" – a term more evocative than precise, but roughly meaning that H_N exhibits a high degree of *symmetry*, and therefore a high degree of *integrability*. For example, when H_N is a GOE matrix, one can give an exact expression at finite N for $\mathbb{E}[|\det(H_N)|]$ in terms of Hermite polynomials [84]. Other highly symmetric models include the classical compact groups, and in fact the whole family of *invariant ensembles*, or matrices whose law admits a density with respect to Lebesgue measure of the form $\frac{1}{Z_{N,\beta,V}} \exp(-\frac{\beta}{2} N \operatorname{Tr} V(H))$ for some "potential" function $V : \mathbb{R} \to \mathbb{R}$. (In my usage, "invariant ensembles" are a strict subset of "models enjoying properties we could describe as invariance.")

The novelty in our result is that the matrices are "non-invariant," meaning they exhibit few symmetries. One archetypal example is a full-rank additive deformation of GOE. Often, observables of non-invariant random matrices do not admit tractable finite-N formulas, but one can still hope for large-N asymptotics such as (1.1.1).

The goal of this thesis is to contribute to the general theory of non-invariant random matrices, finding both old and new phenomena, and to apply them to the study the geometry of high-dimensional random functions – a research program known as "landscape complexity," which we describe in Section 1.2. In Section 1.3 we give some classic results in this program, which motivate our new complexity results in Section 1.4. After circling back to a history of random determinants and a discussion of our new methods in Section 1.5, in Section 1.6 we give our results on large deviations for non-invariant random matrices, and remark on their possible future interplay with landscape complexity. Finally, in Section 1.7 we discuss open questions.

1.2 Landscape complexity: Overview

Consider a smooth, Gaussian random function $\mathcal{H}_N : \mathbb{R}^N \to \mathbb{R}$, perhaps a Hamiltonian from statistical physics or a loss function from data science, which we think of as a "landscape." We wish to understand the geometry of this function for large N, specifically through the random variable $\mathrm{Crt}(\mathcal{H}_N)$, which stores the (random) number of critical points of $\mathcal{H}_N$. This random variable is often exponential in N, so we consider the real number

$$\Sigma = \lim_{N\to\infty} \frac{1}{N} \log \mathbb{E}[\mathrm{Crt}(\mathcal{H}_N)],$$

which is known as the *(annealed) complexity* of the family $(\mathcal{H}_N)_{N=1}^{\infty}$. Much of this thesis is devoted to techniques to compute Σ and related quantities when the landscape $\mathcal{H}_N$ is non-invariant, meaning it exhibits few symmetries. One important variant is $\Sigma_{\min}$, which is the complexity specifically of local minima among all critical points.

Even though we lose information passing from $\mathcal{H}_N$ to $\mathrm{Crt}(\mathcal{H}_N)$, the complexity is still conjectured to be a good predictor of interesting phenomena about $\mathcal{H}_N$. For example:

- The sign of Σ is useful for predicting the *dynamics of optimization* on $\mathcal{H}_N$. For example, perhaps $\mathcal{H}_N$ is a likelihood function in some statistical problem, which is random because it depends on random samples. It might be the case that the maximum likelihood estimator (the MLE, which is the argmax of $\mathcal{H}_N$) is known to be a good, consistent estimator – but that we do not know a good algorithm with which to quickly compute it. This is known as a computational-to-statistical gap, and a variety of research in data science is concerned with studying such gaps. One reason local algorithms like Langevin dynamics might fail is if they get trapped in a large number of critical points of $\mathcal{H}_N$, and this suggests the following rule of thumb:
  - If $\Sigma \leqslant 0$, then optimizing $\mathcal{H}_N$ should be easier.
  - If $\Sigma > 0$, then optimizing $\mathcal{H}_N$ should be harder.

- Variants of Σ can be useful for locating ground states. Write $\Sigma(t)$ for the exponential asymptotics of the number of critical points of $\mathcal{H}_N$ at which $\mathcal{H}_N$ takes values below t (or Nt, depending on the scaling). Then $\Sigma(\cdot) : \mathbb{R} \to \mathbb{R}$ is a non-decreasing function, tending as $t \to +\infty$ to the complexity Σ of all critical points. By Markov's inequality, the quantity $\inf\{t : \Sigma(t) = 0\}$ is a lower bound for the ground state. (To find a matching upper bound, one has to show concentration of $\mathrm{Crt}(\mathcal{H}_N)$ about its mean.)

- Finally, if $\mathcal{H}_N$ is the Hamiltonian of some model in statistical physics, there should be some relationship between variants of Σ and replica symmetry/replica symmetry breaking. To the best of our knowledge, the precise relationship is still being worked out even in the physics literature, but here is a heuristic example: At very low temperature, the Gibbs measure should be dominated by (neighborhoods of) local minima with very low energy levels. If there are many of these arranged in some hierarchy (which could be studied via the quenched analogue of $\Sigma(t)$ above when t is close to the ground state), then the model might have broken replica symmetry; but if there are few of these, then the model might be replica symmetric. See Fyodorov and Williams [94] for one model where this connection can be proven.

We emphasize that these are not theorems, and indeed the extent to which they are true is still an area of active research.

The best way to compute Σ is through a classical result known as the *Kac-Rice formula*, which forms a bridge between random matrices and random geometry. In this case, the Kac-Rice formula reads

$$\mathbb{E}[\mathrm{Crt}(\mathcal{H}_N)] = \int_{\mathbb{R}^N} \mathbb{E}\Big[\Big|\det(\nabla^2 \mathcal{H}_N(x))\Big| \,\Big|\, \nabla \mathcal{H}_N(x) = 0\Big] \phi_x(0)\, \mathrm{d}x, \tag{1.2.1}$$

where $\phi_x(0)$ is the density of the (Gaussian) random vector $\nabla \mathcal{H}_N(x)$ evaluated at $0 \in \mathbb{R}^N$. (Kac-Rice is usually stated over a *compact* set, but for simplicity we assume the common situation where one can pass to all of $\mathbb{R}^N$ by monotone convergence.) The books [2, 17] give a thorough introduction to the Kac-Rice formula. We remark that, between a variety of differential-geometric conditions and the difficulty of conditioning beyond the Gaussian regime, Kac-Rice is effectively only available

when $\mathcal{H}_N$ is Gaussian (or a function of a Gaussian).

In many models, the inner expectation in (1.2.1) does not depend on $x \in \mathbb{R}^N$, and $\int_{\mathbb{R}^N} \phi_x(0)\,\mathrm{d}x$ is easy to understand, so all the difficulty in computing Σ lies in the asymptotics

$$\lim_{N\to\infty} \frac{1}{N} \log \mathbb{E}[|\det(H_N)|], \tag{1.2.2}$$

where H_N is a random matrix distributed as the Hessian of $\mathcal{H}_N$ (at, say, $x = 0$), conditioned on criticality. An analogue of Kac-Rice for local minima inserts an indicator $\mathbb{1}\{\nabla^2 \mathcal{H}_N(x) \geqslant 0\}$ inside the conditioned expectation, restricting the conditioned Hessian to be positive semi-definite, and this indicator persists in the analogue of (1.2.2).

In many of the landscape models studied to date, the random function $\mathcal{H}_N$ is invariant, and thus the random matrix H_N is invariant. In particular, large deviations principles (LDPs) are often available for its empirical spectral measure or extreme eigenvalues, and these LDPs have been an important input for past results on invariant models. Some of these models are described in Section 1.3.

But when the random function $\mathcal{H}_N$ is non-invariant, the random matrix H_N is also non-invariant. Theorem 1.1.1 therefore allows us to give new results in landscape complexity, which are described in Section 1.4.

1.3 Landscape complexity: History

Our models of study are inspired by three important results in landscape complexity, which we now present. A more extensive history is given in Chapter 3. In comparison with these models, our models will have fewer symmetries: for example we give a signal-plus-noise model where the signal is anisotropic rather than isotropic, and a spin-glass model where spins interact in multiple different groups rather than all on equal footing.

Soft spins in an isotropic well. The first model does not have a consistent name in the literature: Sometimes it is thought of as a "zero-dimensional elastic manifold," but we will refer to it as "soft spins in an isotropic well." It is the subject of Fyodorov's breakthrough 2004 paper [84], which is also the first paper in modern landscape complexity, and it is given as

$$\mathcal{H}_N(x) = \frac{\mu}{2}\|x\|^2 + V_N(x), \tag{1.3.1}$$

where $\mu > 0$ is a parameter and $V_N(x)$ is a centered, isotropic Gaussian field with covariance structure given by

$$\mathbb{E}[V_N(x)V_N(y)] = NB\left(\frac{\|x-y\|^2}{2N}\right)$$

for some function $B : \mathbb{R}_+ \to \mathbb{R}$ (think $B(r) = e^{-r}$, for example). Schoenberg classified all possible such functions B (see (3.2.1)), and we add a very mild regularity assumption. Typically V_N has many critical points, whereas the quadratic term only has one, and one should thus expect a phase transition in μ: When μ is quite small, the noise term should dominate, the model should be "disordered," and $\mathcal{H}_N$ should have many critical points. But when μ is quite large, the quadratic "signal" should dominate, the model should be "ordered," and $\mathcal{H}_N$ should have few critical points. This phenomenon is sometimes called "topology trivialization," and it is the intuition for the following theorem, which combines results of Fyodorov and Fyodorov-Williams.

Theorem 1.3.1. *[84, 94] Write $\Sigma^{\mathrm{tot}}(\mu, B)$ (respectively, $\Sigma^{\mathrm{min}}(\mu, B)$) for the complexity of total critical points (respectively, just local minima) of this model. These depend on B only through the scalar $B''(0)$, and*

$$\Sigma^{\mathrm{tot}}(\mu, B''(0)) = \begin{cases} \frac{1}{2}\left(\frac{\mu^2}{B''(0)} - 1\right) - \log\left(\frac{\mu}{\sqrt{B''(0)}}\right) & \text{if } \mu \leqslant \mu_c := \sqrt{B''(0)}, \\ 0 & \text{if } \mu \geqslant \mu_c, \end{cases}$$

$$\Sigma^{\mathrm{min}}(\mu, B''(0)) = \begin{cases} \frac{1}{2}\left[-3 - \log\left(\frac{\mu^2}{B''(0)}\right) + \frac{4\mu}{\sqrt{B''(0)}} - \frac{\mu^2}{B''(0)}\right] & \text{if } \mu \leqslant \mu_c, \\ 0 & \text{if } \mu \geqslant \mu_c. \end{cases}$$

(Actually, Fyodorov and Williams studied a more general model, replacing the radial quadratic term $\frac{\mu}{2}\|x\|^2$ with a radial term $NU(\frac{\|x\|^2}{2N})$ for some fixed, convex U.)

We make several observations about these functions that will reappear in the sequel. First, the phase transition is continuous, and the near-critical behavior is quadratic for total critical points but cubic for local minima. Second, it is not obvious that the same critical value μ_c should appear both for total critical points and for local minima; one might have guessed that, for fixed $B''(0)$, there was a range of μ values with zero complexity for local minima but positive complexity for total critical points (say from many saddle points), and this turns out to be wrong. Although the theorem is for the large-N limit, Figure 1.1 displays a visible phase transition when $N = 2$ and μ varies.

Spherical spin glasses. The second important result is the work of Auffinger-Ben Arous-Černý and Auffinger-Ben Arous on spherical spin glasses [10, 9]. For integer $p \geqslant 2$, the pure p-spin Hamiltonian $H_{N,p} : \mathbb{S}^{N-1} \to \mathbb{R}$ is given by

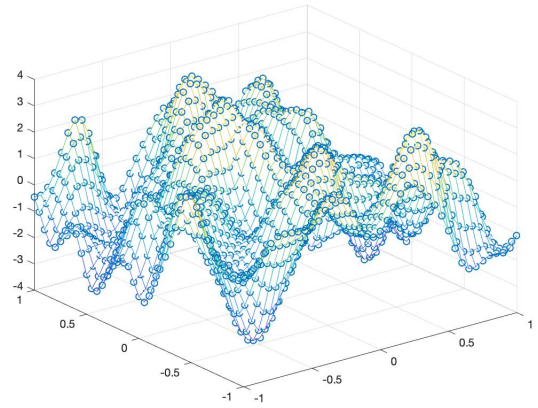
$$H_{N,p}(\sigma) = \frac{1}{N^{(p-1)/2}} \sum_{i_1,\dots,i_p=1}^{N} J_{i_1,\dots,i_p}\sigma_{i_1}\dots\sigma_{i_p}$$

where $\sigma = (\sigma_1, \dots, \sigma_N)$, the sum is over all p-tuples, and the $J_{i_1,\dots,i_p}$ are i.i.d. standard Gaussians. Spin-glass mixtures are prescribed by a sequence $\beta = (\beta_p)_{p=2}^{\infty}$ of positive numbers satisfying the decay condition $\sum_{p=2}^{\infty} 2^p\beta_p < \infty$, and are given by the Hamiltonian

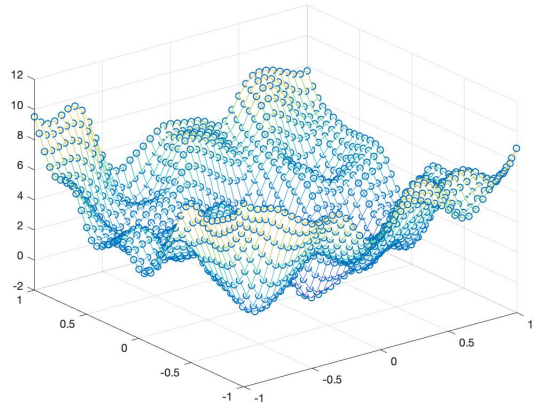
$$H_N(\sigma) = \sum_{p=2}^{\infty} \beta_p H_{N,p}(\sigma),$$

where the different pure-p spin Hamiltonians are independent. These exhibit markedly different phenomena from the soft-spins model: For any model that is not a pure 2-spin, the total complexity of local minima (and thus of critical points) is positive, so there is no order-disorder phase transition. Instead, the interesting question is about the location of critical points in *energy space*. That is, restricting ourselves to pure p-spin models for exposition, one studies the complexity $\Sigma_p(t)$

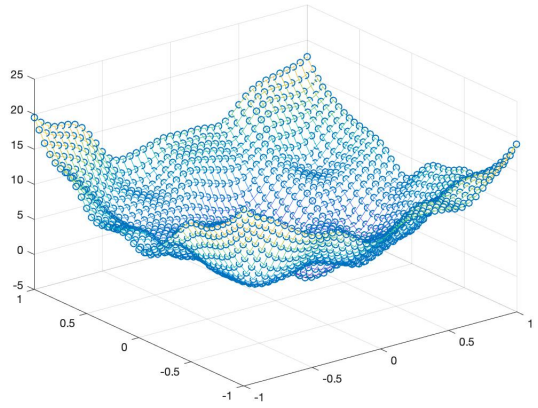
(respectively, $\Sigma_{p,k}(t)$) of critical points (respectively, critical points of fixed index k) at which the Hamiltonian takes values at most Nt. For pure p-spin models, among the many results of [10] are the following.



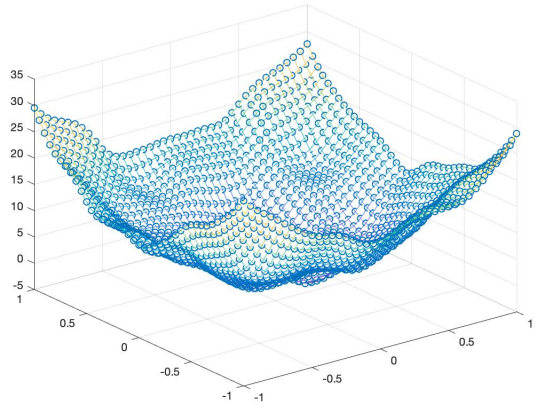
(a) The landscape appears rugged when $\mu = 0$.



(b) Fewer crit. points when $\mu = 10$.



(c) Many fewer crit. points when $\mu = 20$.



(d) Almost no crit. points when $\mu = 30$.

Figure 1.1: Numerical (discretized) samples of $\mathcal{H}_2$ on $[-1,1]^2$ with the same noise and four different choices of μ. Precisely, these are scatterplots of $\mathcal{H}_2(x)$ values for x on a 41×41 lattice, with an overlaid mesh fit, made with Matlab. Here $B(r) = \exp(-80r)$, meaning $\mathbb{E}[V_2(x)V_2(y)] = 2\exp(-20\|x-y\|^2)$.

Theorem 1.3.2. *[10] Fix $p \geqslant 3$ and consider the threshold*

$$E_\infty = E_\infty(p) = 2\sqrt{\frac{p-1}{p}}$$

and the function $I_1 : (-\infty, -E_\infty] \to \mathbb{R}$ given by

$$I_1(u) = \frac{2}{E_\infty^2} \int_u^{-E_\infty} \sqrt{y^2 - E_\infty^2}\, \mathrm{d}y.$$

The functions Σ_p and $\Sigma_{p,k}$ defined above are given explicitly as

$$\Sigma_p(t) = \begin{cases} \frac{1}{2}\log(p-1) - \frac{p-2}{4(p-1)}u^2 - I_1(u) & \text{if } u \leqslant -E_\infty, \\ \frac{1}{2}\log(p-1) - \frac{p-2}{4(p-1)}u^2, & \text{if } -E_\infty \leqslant u \leqslant 0, \\ \frac{1}{2}\log(p-1) & \text{if } 0 \leqslant u, \end{cases}$$

$$\Sigma_{p,k}(t) = \begin{cases} \frac{1}{2}\log(p-1) - \frac{p-2}{4(p-1)}u^2 - (k+1)I_1(u) & \text{if } u \leqslant -E_\infty, \\ \frac{1}{2}\log(p-1) - \frac{p-2}{p} & \text{if } u \geqslant -E_\infty. \end{cases}$$

Furthermore, for any $k \geqslant 0$ write $E_k = E_k(p)$ for the unique solution to $\Sigma_{p,k}(-E_k(p)) = 0$; then $E_0 > E_1 > E_2 > \cdots$ with $\lim_{k\to\infty} E_k(p) = E_\infty(p)$, and for any $\varepsilon > 0$ we have

$$\begin{aligned} &\limsup_{N\to\infty} \frac{1}{N^2} \log \mathbb{P}(\text{exists crit. point of index } k \text{ above level } -N(E_\infty(p) - \varepsilon)) < 0, \\ &\limsup_{N\to\infty} \frac{1}{N} \log \mathbb{P}(\text{exists crit. point of index } \geqslant k \text{ below level } -N(E_k(p) + \varepsilon)) < 0. \end{aligned} \tag{1.3.2}$$

This result establishes a layered structure, where most critical points of index k are found in the band $[-NE_k, -NE_\infty]$. So the critical points with lowest energy are primarily local minima; then in a higher-energy band one starts to see index-one saddle points; the index-two saddle points in the next higher band, and so on.

Spiked-tensor model. Finally, the third important result is the work of Ben Arous, Mei, Montanari, and Nica on the spiked-tensor model [43]. In this model, which depends on integer $k \geqslant 3$, one is given a sample of the random function $f : \mathbb{S}^{N-1} \to \mathbb{R}$ defined by

$$f(\sigma) = \lambda \langle u, \sigma \rangle^k + \frac{1}{\sqrt{2N}} \sum_{i_1,\dots,i_k=1}^{N} G_{i_1,\dots,i_k} \sigma_{i_1} \cdots \sigma_{i_k}.$$

Here $\lambda \geqslant 0$ is a signal-to-noise ratio; $u \in \mathbb{S}^{N-1}$ is an unknown signal we are trying to recover by maximizing our sample(s) of f; and $(G_{i_1,\dots,i_k})_{1 \leqslant i_1,\dots,i_k \leqslant N}$ are i.i.d. standard Gaussians. When $\lambda = 0$, we recover the pure spherical k-spin glass, which has positive complexity. But for large λ, one expects the landscape to trivialize close to u. The argmax of f is the maximum-likelihood estimator for u.

Given Borel $M \subset [-1, 1]$ and $E \subset \mathbb{R}$, write $\mathrm{Crt}_{N,*}(M, E)$ for the number of critical points σ of f at which $\langle \sigma, u \rangle \in M$ and $f(\sigma) \in E$. By rotational invariance, $\mathbb{E}[\mathrm{Crt}_{N,*}(M, E)]$ does not depend on the choice of u.

Theorem 1.3.3. *[43] For each fixed λ, there is an explicit, relatively simple function $S_* : [-1, 1] \times \mathbb{R} \to (\mathbb{R} \cup \{-\infty, +\infty\})$ such that, for any Borel M and E, we have*

$$\limsup_{N\to\infty} \left\{ \frac{1}{N} \log \mathbb{E}[\mathrm{Crt}_{N,*}(M, E)] - \sup_{m \in \overline{M}, e \in \overline{E}} S_*(m, e) \right\} \leqslant 0,$$
$$\liminf_{N\to\infty} \left\{ \frac{1}{N} \log \mathbb{E}[\mathrm{Crt}_{N,*}(M, E)] - \sup_{m \in M^\circ, e \in E^\circ} S_*(m, e) \right\} \geqslant 0.$$

There is another λ-dependent function S_0 satsfying an analogue for local maxima.

By analyzing S_* and S_0 as λ varies, Ben Arous *et al.* suggest the following qualitative picture for some values $\lambda_g > \lambda_c > 0$:

- For $0 \leqslant \lambda < \lambda_c$, most local maxima σ have small correlations $\langle \sigma, u \rangle$ with the planted signal and small function values $f(\sigma)$.
- For $\lambda \in (\lambda_c, \lambda_g)$, there are local maxima σ with large correlations $\langle \sigma, u \rangle$, but they have

smaller function value $f(\sigma)$ than other local maxima with small correlations, so the maximum-likelihood estimator is still bad.

- For $\lambda > \lambda_g$, there are local maxima σ with large correlations $\langle \sigma, u \rangle$ and large function value $f(\sigma)$.

They also note the following: On the one hand, the best known algorithm requires $\lambda \gtrsim N^{(k-2)/2}$ to succeed (meaning to achieve positive correlation with u). On the other hand, the complexity results suggest that there are an exponential number of local maxima in the annulus $\{|\langle \sigma, u \rangle| \lesssim \lambda^{-1/(k-2)}\}$, and a uniformly random initialization on the sphere lies *outside* of this annulus with positive probability in the *same* regime $\lambda \gtrsim N^{(k-2)/2}$. Although this is just an observation, it is consistent with a claim like "local algorithms fail if they are initialized in regions of positive complexity."

In general, the matrices that appear in these models are covered by our result on determinant concentration.

1.4 Landscape complexity: Our results

In this section we describe our results on landscape complexity for three models. The first two generalize the soft-spins model of [84], and the third is related to the spherical spin glasses of [10].

Elastic manifold. The elastic manifold is a classical model in statistical physics that assigns a random energy to deterministic configurations of L^d points lying in $\mathbb{R}^N$. There are two contributions to the energy: Each point behaves as the isotropic soft-spins model (1.3.1), plus there are nearest-neighbor interactions between points. More formally, given positive integers L ("length") and d ("internal dimension"), we write Ω for the lattice $[\![1, L]\!]^d$, understood periodically. A point configuration will be written as a deterministic $\mathbf{u} : \Omega \to \mathbb{R}^N$, and given positive numbers μ ("mass")

and t ("interaction strength"), we assign such a point configuration the random energy

$$\mathcal{H}[\mathbf{u}] = \sum_{x\in\Omega}\Big(\mu\|\mathbf{u}(x)\|^2 + V_N(\mathbf{u}(x),x)\Big) + \sum_{x,y\in\Omega} -t\Delta_{xy}\langle\mathbf{u}(x),\mathbf{u}(y)\rangle,$$

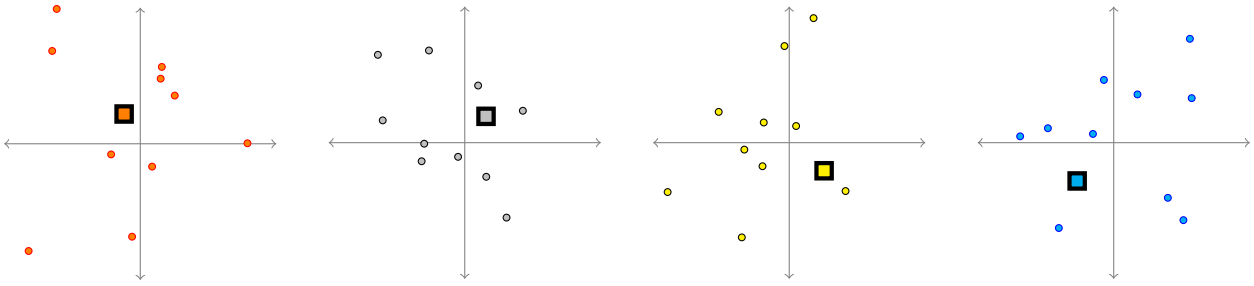
where Δ_{xy} is the (x,y) entry of the $L^d\times L^d$ matrix Δ, which is the periodic lattice Laplacian, and where the $(V_N(\cdot,x))_{x\in\Omega}$ are centered isotropic Gaussian fields, independent for different x values, each with covariance $\mathbb{E}[V_N(y_1,x)V_N(y_2,x)] = NB\Big(\frac{\|y_1-y_2\|^2}{N}\Big)$ for some function B satisfying (3.2.1) (scaled here, for exposition, so that certain factors in the following simplify).

Notice that $\mathcal{H}$ contains three competing influences: (i) a quadratic confining potential with strength μ that keeps points close to the origin; (ii) the elastic (Laplacian) term with strength t that prefers ordered point configurations; and (iii) random spatial impurities (the Gaussian fields V_N) with strength stored in B that prefer disordered point configurations. Figure 1.2 shows these three competing influences visually.

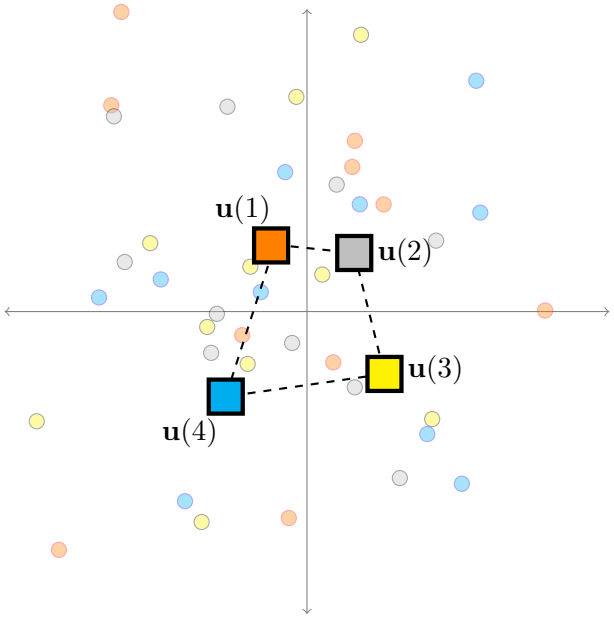
In this context a "critical point" is a configuration $\mathbf{u}$ that is critical for the Hamiltonian (meaning such that $\partial_{\mathbf{u}_i(x)}\mathcal{H}[\mathbf{u}] = 0$ for all i and all x), and a "local minimum" is a critical point that also locally minimizes the Hamiltonian against small perturbations of the points. Equivalently, one can think of $\mathcal{H}$ as a Gaussian random function defined on $(\mathbb{R}^N)^\Omega \cong \mathbb{R}^{NL^d}$, in which case "critical point" has its usual meaning. We want to count critical points in the regime when L and d are fixed but $N\to+\infty$. Notice that, if μ is very large, then point configurations get pushed towards zero, and there should be few critical point configurations; whereas if μ is very small, then point configurations have more freedom to roam about space, and there should be many critical point configurations. This is the content of the following theorem.

Write $\Sigma(\mu) = \Sigma(\mu,L,d,t,B)$ for the complexity of critical points, and $\Sigma_{\mathrm{st}}(\mu) = \Sigma_{\mathrm{st}}(\mu,L,d,t,B)$ for the complexity of local minima.

Theorem 1.4.1. *(Ben Arous-Bourgade-M. [36]) Write $\hat{\mu}_{-\Delta}$ for the empirical spectral measure of the Laplacian $-\Delta$, and for each t, let the "Larkin mass" $\mu_c = \mu_c(L,d,t)$ be the unique positive*



(a) $\mathbf{u}(1)$ contribution. (b) $\mathbf{u}(2)$ contribution. (c) $\mathbf{u}(3)$ contribution. (d) $\mathbf{u}(4)$ contribution.



(e) The elastic contribution to the configuration energy is small, because the sum $\langle \mathbf{u}(1), \mathbf{u}(2) \rangle + \langle \mathbf{u}(2), \mathbf{u}(3) \rangle + \langle \mathbf{u}(3), \mathbf{u}(4) \rangle + \langle \mathbf{u}(4), \mathbf{u}(1) \rangle$ of inner products between points whose indices are nearest neighbors in the underlying lattice is small.

Figure 1.2: Informal schematic of one low-energy elastic manifold configuration when $d = 1$, $L = 4$, and $N = 2$. The four manifold points $\mathbf{u}(1)$, $\mathbf{u}(2)$, $\mathbf{u}(3)$, and $\mathbf{u}(4)$ are indicated by squares. In the top four subfigures, which indicate the contributions to the total energy made by each manifold point on its own, each manifold point sees its own (independent) Gaussian environment and tries to avoid points of high energy cost (represented by circles of the same color) while staying close to the origin. In the bottom subfigure, these environments are overlaid, showing that the points have achieved their separate goals while also keeping their lattice-neighbor inner products small. The inner products $\langle \mathbf{u}(1), \mathbf{u}(3) \rangle$ and $\langle \mathbf{u}(2), \mathbf{u}(4) \rangle$ do not contribute, because $\{1, 3\}$ and $\{2, 4\}$ are not lattice neighbors. Perhaps this configuration is a local minimum, meaning the energy $\mathcal{H}[\mathbf{u}]$ increases if we slightly perturb any of the images $\mathbf{u}(i)$. We are trying to count such minima (and total critical points) in the $N \to +\infty$ limit, when these four points are immersed not in the plane but in a high-dimensional space. This figure is inspired by [95, Figure 2].

solution to

$$\int_{\mathbb{R}} \frac{\hat{\mu}_{-\Delta}}{(\mu_c + t\lambda)^2} = 1.$$

If $\mu \geqslant \mu_c$, then $\Sigma(\mu) = \Sigma_{\mathrm{st}}(\mu) = 0$. If $\mu < \mu_c$, then $\Sigma(\mu) > \Sigma_{\mathrm{st}}(\mu) > 0$, and these quantities are given by relatively explicit formulas involving the log-potential of a certain free-convolution measure.

Alternatively, for each fixed u and t one can rescale the noise B and find a similar phase transition; we show that the complexity of total critical points vanishes quadratically near this phase transition, while the complexity of local minima vanishes cubically.

This result solves a problem of Fyodorov and Le Doussal [88], who studied the complexity of this model and came to the same conclusion, *assuming* determinant asymptotics of the type (1.1.1) which we can now verify with Theorem 1.1.1.

The elastic manifold has long attracted interest in its own right, both for its *roughness exponent* and a phenomenon it displays called *(de)pinning*. The roughness exponent captures, for example, how rugged an elastic interface is at large spatial distance, and attempts to compute it inspired early technical developments of Fisher in functional renormalization group methods [80] and of Mézard and Parisi in replica symmetry breaking [122]. (De)pinning is a nonlinear response to applied force. Precisely, the point configurations have a preferred position in space, and they only move from this position if an applied force f is larger than the so-called *depinning threshold* f_c. Pinning is critical in applications: for example, we can effectively store information on a magnetic hard drive precisely because the data is "pinned" against, e.g., small temperature fluctuations.

Soft spins in an anisotropic well. Next we consider the Hamiltonian

$$\mathcal{H}_N(u) = \frac{1}{2}\langle u, D_N u\rangle + V_N(u),$$

where V_N is a centered isotropic Gaussian field with $\mathbb{E}[V_N(x)V_N(y)] = NB(\frac{\|x-y\|^2}{2N})$ as in (1.3.1), and where $(D_N)_{N=1}^{\infty}$ is a sequence of deterministic, diagonal matrices without outliers whose empirical spectral measures $\hat{\mu}_{D_N}$ tend weakly to some μ_D, a limiting probability measure which should be

compactly supported in $(0, \infty)$. That is, the confining potential is no longer radial, but has meaningfully different directions given by the entries of D_N, and the potential is ultimately characterized by a *probability measure* rather than by a scalar. For example, if $D_N = \mathrm{diag}(1, \dots, 1, 2, \dots, 2)$, then the confining potential has two directions, and $\mu_D = \frac{1}{2}(\delta_1 + \delta_2)$; but one can take general "signal measures" μ_D which are not combinations of delta masses. We call this model "soft spins in an anisotropic well." Of course, in the special case $D_N = \mu \,\mathrm{Id}$, we recover the original model (1.3.1) for soft spins in an isotropic well. Although our results are for the $N \to +\infty$ limit, Figure 1.3 displays how changing D_N can qualitatively change the number of critical points when $N = 2$.

It is natural to expect an order/disorder phase transition for this problem, depending on *some* scalar observable of μ_D. For example, consider the case $\mu_D = \frac{1}{2}(\delta_1 + \delta_2)$. One might guess that the complexity phase transition behaves as if μ_D were δ_1 (meaning worst-case complexity, or the left endpoint of μ_D), since perhaps there are many critical points "in the flat (one) direction." On the other hand, one might guess that the complexity phase transition behaves as if μ_D were δ_2 (meaning best-case complexity, or the right endpoints of μ_D), since perhaps there are few critical points "in the steep (two) direction." Finally, one might guess that the complexity phase transition behaves as if μ_D were $\delta_{3/2}$ (meaning average-case complexity, or the mean of μ_D), since perhaps these two directions just average out.

All of these guesses are wrong. In fact, the complexity phase transition behaves as if μ_D were $\delta_{\sqrt{8/5}}$. More generally, the right observable of μ_D is, surprisingly, the negative second moment.

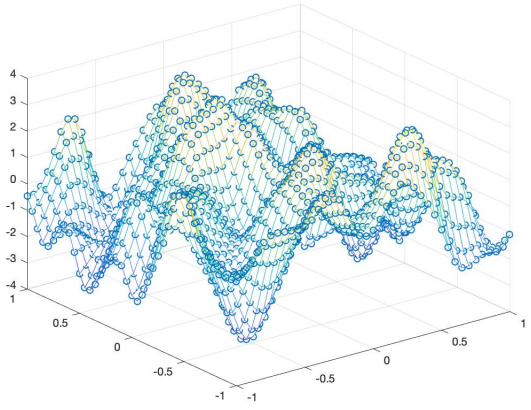
Theorem 1.4.2. *(Ben Arous-Bourgade-M. [36]) There exist relatively explicit functions*

$$\Sigma^{\mathrm{tot}}(\mu_D, t), \quad \Sigma^{\mathrm{min}}(\mu_D, t)$$

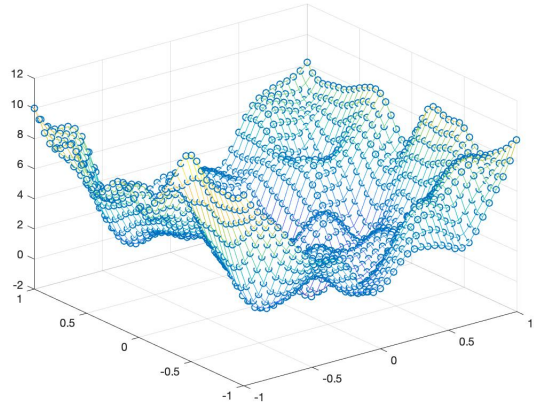
such that the complexity of total critical points of $\mathcal{H}_N$ is described by $\Sigma^{\mathrm{tot}}(\mu_D, B''(0))$ and the complexity of local minima of $\mathcal{H}_N$ is described by $\Sigma^{\mathrm{min}}(\mu_D, B''(0))$. Furthermore, these functions

are either both zero or both positive; we have

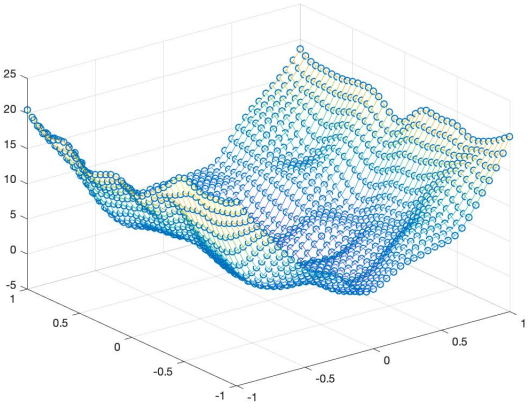
$$\Sigma^{\mathrm{tot}}(\mu_D, t) = \Sigma^{\mathrm{min}}(\mu_D, t) = 0 \quad \textit{if and only if} \quad t \leqslant t_c = t_c(\mu_D) = \left(\int_{\mathbb{R}} \frac{\mu_D(\mathrm{d}\lambda)}{\lambda^2} \right)^{-1};$$



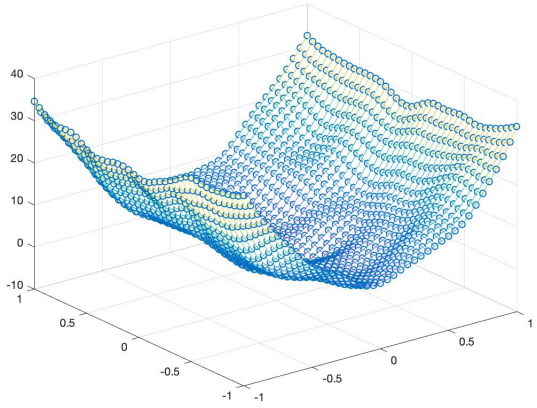
(a) The landscape appears rugged when $D_2 = 0$.



(b) Fewer crit. points when $D_2 = 3 \cdot \left(\begin{smallmatrix} 6 & 0 \\ 0 & 1 \end{smallmatrix}\right)$.



(c) Many fewer crit. points when $D_2 = 6 \cdot \left(\begin{smallmatrix} 6 & 0 \\ 0 & 1 \end{smallmatrix}\right)$.



(d) Almost no crit. points when $D_2 = 10 \cdot \left(\begin{smallmatrix} 6 & 0 \\ 0 & 1 \end{smallmatrix}\right)$.

Figure 1.3: Numerical (discretized) samples of $\mathcal{H}_2$ on $[-1, 1]^2$ with the same noise and four different choices of signal D_N. Precisely, these are scatterplots of $\mathcal{H}_2(x)$ values for x on a 41×41 lattice, with an overlaid mesh fit, made with Matlab. Here $B(r) = \exp(-80r)$, meaning $\mathbb{E}[V_2(x)V_2(y)] = 2\exp(-20\|x - y\|^2)$.

and for slightly supercritical t we have

$$\Sigma^{\mathrm{tot}}(\mu_D, t) = c_{\mathrm{tot}}(\mu_D) \cdot (t - t_c)^2 + O((t - t_c)^3),$$
$$\Sigma^{\mathrm{min}}(\mu_D, t) = c_{\mathrm{min}}(\mu_D) \cdot (t - t_c)^3 + O((t - t_c)^4),$$

with explicit prefactors $c_{\mathrm{tot}}(\mu_D)$, $c_{\mathrm{min}}(\mu_D)$.

While the negative-second moment criterion is new, one can think of this as a *universality* result for the quadratic near-critical behavior of total complexity, and the cubic near-critical behavior of complexity of local minima, as these powers already appeared in the isotropic case.

We mention briefly a technical result in free probability that we establish during the proof of Theorem 1.4.2, possibly of independent interest. Biane [49] gave a comprehensive study of measures of the form "free convolution with semicircle." These measures always admit a density, but in contrast to semicircle itself, they can have disconnected supports, whose components can merge at interesting cusps, typically with cube-root decay. Biane showed that all edges and cusps of such measures decay at least as quickly as a cube-root; we show in Appendix B that at the *extremal* edges the decay must be at least a square-root.

Bipartite spherical spin glasses. Finally, we consider a two-species spin glass, called the "bipartite spherical spin glass," generalizing the classical spherical spin glasses. Given integers $p, q \geqslant 1$ and $\gamma \in (0, 1)$, one defines the pure (p, q, γ) bipartite spin glass as the random function $\mathcal{H}_{N,p,q,\gamma} : \mathbb{S}^{\gamma N} \times \mathbb{S}^{(1-\gamma)N}$ given by

$$\mathcal{H}_{N,p,q,\gamma}(u, v) = \sum_{1 \leqslant i_1, \ldots, i_p \leqslant \gamma N} \sum_{1 \leqslant j_1, \ldots, j_q \leqslant (1-\gamma)N} J_{i_1, \ldots, i_p, j_1, \ldots, j_q} u_{i_1} \ldots u_{i_p} v_{j_1} \ldots v_{j_q},$$

where the J variables are i.i.d. Gaussians with variance $N/((\gamma N)^p((1-\gamma)N)^q)$. One can also define the "mixed" Hamiltonian

$$\mathcal{H}_N(u, v) = \sum_{p,q \geqslant 1} \beta_{p,q} \mathcal{H}_{N,p,q,\gamma}(u, v),$$

for some double sequence $(\beta_{p,q})_{p,q\geqslant 1}$ that decays fast enough.

Theorem 1.4.3. *(M. [120]) We find exact variational formulas for the complexity, both of total critical points and of local minima, for pure models as well as mixtures. We also find two interesting phenomena in the special case of pure models:*

- *There exists a constant $E_\infty(p,q,\gamma) > 0$ such that, for every $\varepsilon > 0$, all local minima have energy values at most $-N(E_\infty(p,q,\gamma) - \varepsilon)$ with all but exponentially small probability.*
- *If $\gamma = \frac{p}{p+q}$, then the complexity functions of a pure (p,q,γ) model are* exactly *those of a pure $p+q$ (single-species) spin glass as studied in [10].*

Notice that the first phenomenon – of local minima lying in a low-energy band – already appeared in pure p-spin models (see (1.3.2)). We also note that upper and lower bounds for the complexity of this model were previously given by Auffinger and Chen [11].

1.5 Random determinants: History

To the best of our knowledge, there is no previous systematic study of determinant asymptotics of the form (1.2.2). But there are previous works on the size of different random determinants, in two strands. A fuller history of the study of random determinants is given in Chapter 2.

First, starting in the 1950s, a variety of authors used combinatorics to find formulas for moments $\mathbb{E}[\det(H_N)^k]$, exact at finite N, for small k, when H_N is actually non-Hermitian. These are combined in the following theorem.

Theorem 1.5.1. *Let μ be a probability measure, symmetric about zero, with unit variance and fourth moment m_4. Let H_N be a non-Hermitian $N \times N$ random matrix with i.i.d. entries distributed according to μ, and let $U_{p,N}$ be a non-Hermitian $p \times N$ random matrix with i.i.d. entries distributed according to μ, with $p \leqslant N$.*

- *(Fortet [82])*

$$\mathbb{E}[\det(H_N)^2] = N!.$$

- *(Nyquist, Rice, Riordan [128])*

$$\mathbb{E}[\det(H_N)^4] = \frac{(N!)^2}{2} \sum_{k=0}^{N} \frac{(N-k+1)(N-k+2)}{k!} (m_4 - 3)^k,$$

and exact formulas for any higher moment when μ is Gaussian. (Another proof for the Gaussian case was later given in Prékopa [132].)

- *(Dembo [68])*

$$\mathbb{E}[\det(U_{p,N}(U_{p,N})^T)] = \frac{N!}{(N-p)!},$$
$$\mathbb{E}[\det(U_{p,N}(U_{p,N})^T)^2] = \frac{N!}{(N-p)!} \sum_{k=0}^{p} \binom{p}{j} \frac{(N+2-k)!}{(N+2-p)!} (m_4 - 3)^k,$$

and exact formulas for any higher moment when μ is Gaussian.

These explicit formulas do not admit extensions at other energy values (i.e., for the determinant of $H_N - E$, when $E \in \mathbb{R}$), or with absolute values, which are what appears in the Kac-Rice formula.

Second, in the last 20 years, different papers in landscape complexity have studied versions of (1.2.2) when H_N is the relevant random matrix for their particular landscapes, which is often invariant.

The earlier of these proofs often use the following clever trick, which allows for exact formulas at finite N: The joint density of eigenvalues of the GOE is proportional to

$$\prod_{1 \leqslant i < j \leqslant N} |\lambda_i - \lambda_j| \prod_{i=1}^{N} e^{-\frac{N}{4}\lambda_i^2} \, \mathrm{d}\lambda_i,$$

so we can write

$$\mathbb{E}[|\det(H_N - x\,\mathrm{Id})|] \propto \int \left(\prod_{i=1}^{N} |\lambda_i - x|\right) \left(\prod_{1 \leqslant i < j \leqslant N} |\lambda_i - \lambda_j|\right) \prod_{i=1}^{N} e^{-\frac{N}{4}\lambda_i^2}\, \mathrm{d}\lambda_i. \tag{1.5.1}$$

The idea is to recognize x as an eigenvalue of an $(N+1) \times (N+1)$ GOE matrix, i.e., to smuggle $\prod_{i=1}^{N} |\lambda_i - x|$ into the Vandermonde determinant. This allows one to relate the determinant to the density of states of GOE (i.e., $\mathbb{E}[\hat{\mu}_{H_N}](x)\,\mathrm{d}x$), which is explicit in terms of Hermite polynomials, allowing for asymptotic analysis. (In a variant that restricts the index as $\mathbb{E}[|\det(H_N - x\,\mathrm{Id})| \mathbb{1}\{i(H_N - x\,\mathrm{Id}) = k\}]$, one recognizes x as the kth smallest eigenvalue of an $(N+1) \times (N+1)$ GOE matrix.) Notice that this argument is specific both (i) to the GOE and its explicit joint density and (ii) to the perturbation $-x\,\mathrm{Id}$ (it would not work, for example, for the perturbation $-\mathrm{diag}(1, \dots, 1, 2, \dots, 2)$). That is, the trick (1.5.1) is an invariant argument.

Various results of this type are contained in the following theorem, which combines results of several authors (sometimes in less generality than given there, for conciseness). These were mostly originally stated as results about landscape complexity (and we discuss them as such in Section 1.3), but here we have rephrased them as results about determinants.

Theorem 1.5.2. *Let H_N be an $N \times N$ GOE matrix.*

- *(Fyodorov [84]) Let $\xi \sim \mathcal{N}(0, 1/N)$ be independent of H_N, and let $\mu > 0$ be deterministic. Writing $x = x(N, \mu, t) = \sqrt{\frac{N}{2}}(\mu + t)$, we have the finite-$N$ formula*

$$\begin{aligned}\mathbb{E}[|\det(H_N + (\xi + \mu)\,\mathrm{Id})|] = \frac{1}{2\pi}\left[\left(\frac{N-1}{2}\right)!\right] \int_{-\infty}^{\infty} \mathrm{d}t\, e^{-N\frac{t^2}{2}} \Bigg(e^{-\frac{x^2}{2}} \sum_{k=0}^{N} \frac{1}{2^k k!} H_k^2(x) \\ + \frac{1}{2^{N+1} N!} H_N(x) \int_{-\infty}^{\infty} e^{-\frac{u^2}{2}} H_{N+1}(u) \operatorname{sign}(x-u)\,\mathrm{d}u\Bigg)\end{aligned}$$

where H_k is the kth Hermite polynomial, and the asymptotics

$$\lim_{N\to\infty} \frac{1}{N} \log \mathbb{E}[|\det(H_N + (\xi+\mu)\,\mathrm{Id})|] = \begin{cases} \frac{\mu^2-1}{2} & \text{if } \mu \leqslant 1, \\ \log(\mu) & \text{if } \mu \geqslant 1 \end{cases}.$$

- *(Auffinger-Ben Arous-Černý [10]) Fix $c > 0$, and let $\xi_c \sim \mathcal{N}(0, c/N)$ be independent of H_N. Fix $k \in \mathbb{N}$, and write $\mathbb{E}_{\mathrm{GOE}}^{N+1}$ for the expectation when λ_k is the kth smallest eigenvalue of a GOE matrix of size $(N+1)\times(N+1)$. Then*

$$\mathbb{E}[|\det(H_N + \xi_c\,\mathrm{Id})|\mathbb{1}\{i(H_N + \xi_c\,\mathrm{Id}) = k\}] = \frac{\Gamma(\frac{N+1}{2})N^{-\frac{N}{2}}}{\sqrt{c\pi}} \mathbb{E}_{\mathrm{GOE}}^{N+1}\left[\exp\left\{(N+1)\left(\frac{1}{2} - \frac{1}{2c}\right)\lambda_k^2\right\}\right].$$

  Furthermore, there is an explicit function $I(\cdot)$ such that the kth smallest eigenvalue of a GOE matrix satisfies an LDP at speed N with the good rate function $kI(\cdot)$ (for the smallest eigenvalue, this dates back to Ben Arous-Dembo-Guionnet [37]; for $k > 1$ it was new in [10]). Thus if $c < 1$ (equivalently $\frac{1}{2} - \frac{1}{2c} < 0$) the asymptotics can be found using Varadhan's lemma.

Finally, in the last five or so years, papers in landscape complexity have started to consider models beyond the regime of the trick (1.5.1) – that is, non-integrable models where one can only give asymptotics rather than finite-N formulas.

For example, Ben Arous, Mei, Montanari, and Nica [43] consider a rank-one perturbation of GOE, using an LDP for the largest eigenvalue of such a matrix established by Maïda [118], and an LDP for the empirical spectral measure of a(n undeformed) GOE matrix due to Ben Arous and Guionnet [42].

Another example appears in two recent results of Baskerville, Keating, Mezzadri, and Najnudel, which cover finite-rank perturbations of GOE [29] and a specific ensemble of Gaussian matrices with a variance profile, inspired by a two-layer spin-glass model of neural networks [30]. In both cases the determinant analysis is performed through rigorous supersymmetric methods.

How do our methods differ? The model-specific results in our Theorem 1.1.1 are all corollaries

of a general theorem, which prove that we can obtain (1.1.2) after checking the following three general conditions which do not use invariance. Stated informally (see Theorems 2.1.1, 2.1.2 for complete statements):

(i) There is no problem caused by extremely large or small eigenvalues (at scale $e^{\pm N^{\varepsilon}}$).

(ii) The empirical measure $\hat{\mu}_{H_N}$ concentrates about its mean $\mathbb{E}[\hat{\mu}_{H_N}]$.

(iii) There exists a sequence $(\mu_N)_{N=1}^{\infty}$ of regular, deterministic probability measures which is a mildly good approximation for $(\mathbb{E}[\hat{\mu}_{H_N}])_{N=1}^{\infty}$.

We offer two interpretations for condition (ii). The first, more standard interpretation says it suffices for traces of Lipschitz functions of H_N to concentrate (which follows, for example, from log-Sobolev via [103], or from Gromov-Milman concentration of compact groups). The second, more novel interpretation holds when H_N is given as a Lipschitz, convex function of some independent random variables. As a simple example, a Wigner matrix is a *linear* function of $\frac{N(N+1)}{2}$ independent random variables – namely the upper-triangular entries, which are then arranged above and below the diagonal. As a more complicated example, one could start with some independent random variables and mix them differently in different entries, to make a random matrix with correlations. This case is designed to apply Talagrand's classical concentration-of-measure results, which say that Lipschitz, convex functions of bounded, independent random variables concentrate. By approximating the logarithm by some Lipschitz, convex functions, and truncating the input random variables, we can write the expected determinant almost as a Lipschitz, convex function of bounded, independent random variables, yielding concentration. Product-measure concentration is of course now very common in probability, but we remark that it is not so common in this corner of random matrix theory, and that it is pleasantly surprising that it gives almost-sharp results (for example, for Wigner matrices with $2 + \varepsilon$ finite moments).

1.6 Large deviations for extreme eigenvalues

As described above, one of the goals of this thesis is to describe random matrices for which LDPs are *not* available. However, some LDPs for non-invariant random matrices *have* recently been established, in the following breakthrough results of Guionnet and Husson [99] and Guionnet and Maïda [102]. We state only the real case, but all results are true in the complex case as well.

Theorem 1.6.1. *[99] Let μ be a centered probability measure on $\mathbb{R}$ with unit variance that is* sharp sub-Gaussian*, in the sense that*

$$A := 2 \sup_{t \in \mathbb{R}} \frac{1}{t^2} \log\left(\int e^{tx} \mu(\mathrm{d}x) \right) = 1. \tag{1.6.1}$$

(The condition $A < \infty$ is usually called sub-Gaussian*; notice this is asking for more. Examples of sharp sub-Gaussian measures include the Bernoulli and Uniform distributions, appropriately scaled.) Let W_N be a Wigner matrix associated with μ, meaning that W_N has independent entries up to symmetry with $(W_N)_{ij} \sim \sqrt{\frac{1+\delta_{ij}}{N}}\mu$. Then the largest eigenvalue of W_N satisfies an LDP at speed N with the same rate function as that of the GOE, namely*

$$I(x) = \begin{cases} +\infty & \text{if } x < 2, \\ \frac{1}{2}\int_2^x \sqrt{y^2 - 4}\,\mathrm{d}y & \text{if } x \geqslant 2. \end{cases}$$

For the case of non-sharp sub-Gaussian distributions (i.e., $A > 1$ in (1.6.1)), large-deviations estimates are given by Augeri, Guionnet, and Husson in [15]. These estimates show that the rate function is *not* the same as that of the GOE, since it is asymptotic to $\frac{1}{4A}x^2$ for large x.

Theorem 1.6.2. *[102] Let $(A_N)_{N=1}^{\infty}$ and $(B_N)_{N=1}^{\infty}$ be two sequences of deterministic real diagonal matrices whose empirical spectral measures $\hat{\mu}_{A_N}$ and $\hat{\mu}_{B_N}$ tend to compactly supported limiting measures μ_A and μ_B, respectively, as $N \to \infty$. Suppose that $\sup_N(\|A_N\| + \|B_N\|) < \infty$, and that the largest eigenvalues have limits $\lambda_{max}(A_N) \to \mathtt{r}(\mu_A)$ and $\lambda_{max}(B_N) \to \mathtt{r}(\mu_B)$, respectively (here*

$\mathtt{r}(\cdot)$ *is the right edge of a compactly supported probability measure). If* O_N *is a Haar orthogonal matrix of size* N*, then the largest eigenvalue of* $A_N + O_N B_N O_N^T$ *satisfies an LDP at speed* N *with the good rate function*

$$I(x) = \begin{cases} +\infty & \text{if } x < \mathtt{r}(\mu_A \boxplus \mu_B), \\ \sup_{\theta \geqslant 0}\{J(\mu_A \boxplus \mu_B, \theta, x) - J(\mu_A, \theta, \mathtt{r}(\mu_A)) - J(\mu_B, \theta, \mathtt{r}(\mu_B))\} & \text{if } x \geqslant \mathtt{r}(\mu_A \boxplus \mu_B). \end{cases}$$

Here the functions J *are explicit functions arising in the analysis of rank-one spherical integrals; see* (5.2.2) *for the exact form.*

Guionnet and Maïda give extensions to the case when A_N and B_N have outliers below the BBP threshold (meaning such that $\lambda_{\max}(A_N + O_N B_N O_N^T) \to \mathtt{r}(\mu_A \boxplus \mu_B)$), and for a specific model with more extreme outliers.

In Chapter 5, we use and extend the techniques of these papers to study large deviations of additively deformed Wigner matrices of the form $W_N + D_N$, where W_N is a sharp sub-Gaussian Wigner matrix and D_N is deterministic, usually with full rank. The results are stronger if W_N is in fact Gaussian, and even this special case was new:

Theorem 1.6.3. *(M. [121]) Let* W_N *be distributed according to the GOE, and let* $(D_N)_{N=1}^{\infty}$ *be a sequence of deterministic real symmetric matrices whose empirical spectral measures tend to some compactly supported limiting measure* μ_D *as* $N \to \infty$ *(with a mild speed-of-convergence assumption). Suppose that the largest and smallest eigenvalues of* D_N *tend to the right and left endpoints of* μ_D*, respectively. Then* $\lambda_{\max}(W_N + D_N)$ *satisfies an LDP at speed* N *with the good rate function*

$$I(x) = \begin{cases} +\infty & \text{if } x < \mathtt{r}(\rho_{\mathrm{sc}} \boxplus \mu_D), \\ \sup_{\theta \geqslant 0}\{J(\rho_{\mathrm{sc}} \boxplus \mu_D, \theta, x) - \theta^2 - J(\mu_D, \theta, \mathtt{r}(\mu_D))\} & \text{if } x \geqslant \mathtt{r}(\rho_{\mathrm{sc}} \boxplus \mu_D). \end{cases}$$

Surprisingly, the rate function here is not analytic: It has a second-order phase transition at some finite x_c, for reasons that remain mysterious, but are perhaps related to a localization-delocalization

phase transition.

All of these results are proved by tilting the measure by a Laplace transform, as in the classical proof of Cramér's theorem. But the appropriate Laplace transform here is the *(rank-one) spherical integral* defined for $\theta \geqslant 0$ and an $N \times N$ matrix A by

$$I_N(A, \theta) = \mathbb{E}_e[e^{N\theta\langle e, Ae\rangle}], \tag{1.6.2}$$

where $\mathbb{E}_e$ integrates over vectors e uniform on the unit sphere $\mathbb{S}^{N-1} \subset \mathbb{R}^N$. This is a special case of the famous Harish-Chandra/Itzykson/Zuber (HCIZ) integral, defined for two matrices A and B by integrating over Haar measure on the orthogonal group $\mathcal{O}_N$ as

$$\int e^{N \operatorname{Tr}(OAO^T B)} \, \mathrm{dHaar}(O). \tag{1.6.3}$$

If we take $B = \operatorname{diag}(\theta, 0, \ldots, 0)$, this reduces to (1.6.2). The large-N asymptotics of (1.6.2) were established by Guionnet and Maïda in 2005 [101], and involve the function J defined above; see Chapter 5 for complete formulas. We mention that the asymptotics of rank-one spherical integrals are remarkably concise, given the famous difficulty of describing the asymptotics of the full-rank HCIZ integral [104, 127].

Future research might be able to combine these LDPs with our techniques for random-matrix determinants to obtain new results in landscape complexity. Specifically, to study critical points of index k, a variant of the Kac-Rice formula reduces in all of our cases to

$$\int_{\mathbb{R}^m} e^{-N\frac{\|u\|^2}{2}} \mathbb{E}[|\det(H_N(u))| \mathbb{1}_{i(H_N(u))=k}] \, \mathrm{d}u,$$

where $i(\cdot)$ is the index of a matrix, m is independent of N, and $(H_N(u))_{u \in \mathbb{R}^m}$ is a field of non-invariant random matrices. We give results in this thesis for the case $k = 0$ (meaning local minima); we are about to describe these, after which we will consider possibilities for $k \geqslant 1$ (meaning saddle points). For pure spherical p-spin models, annealed (and quenched!) results for fixed k have been

established in [10, 141, 12]; here we have in mind a general model for which annealed estimates are not yet available.

If the matrix $H_N(u)$ has limiting empirical spectral measure $\mu_\infty(u)$, consider the set

$$\mathcal{G} = \{u \in \mathbb{R}^m : \mathrm{supp}(\mu_\infty(u)) \subseteq [0, \infty)\}$$

of good u values, and restrict momentarily to the case $k = 0$. If the matrices $H_N(u)$ have asymptotically no outliers, which is true for all of our models, then $\mathcal{G}$ is the same (up to boundary issues) as the set of u values for which $\{i(H_N(u)) = 0\}$ is a likely event. Indeed, in Chapter 2 we show that

$$\mathbb{E}[|\det(H_N(u))|\mathbb{1}_{i(H_N(u))=0}] \approx \begin{cases} \mathbb{E}[|\det(H_N(u))|] & \text{if } u \in \mathcal{G}, \\ 0 & \text{if } u \notin \mathcal{G} \end{cases}$$

at exponential scale, and that consequently

$$\lim_{N\to\infty} \frac{1}{N} \log \int_{\mathbb{R}^m} e^{-N\frac{\|u\|^2}{2}} \mathbb{E}[|\det(H_N(u))|\mathbb{1}_{i(H_N(u))=0}]\, \mathrm{d}u = \sup_{u\in\mathcal{G}} \left\{ \int \log|\lambda| \mu_\infty(u)(\mathrm{d}\lambda) - \frac{\|u\|^2}{2} \right\}.$$

What happens when $k = 1$? Suppose one can show, perhaps through tilting by spherical integrals, that $\lambda_{\min}(H_N(u))$ satisfies an LDP at speed N with the good rate function $I(u, \cdot)$. Then one naturally guesses

$$\frac{1}{N} \log \mathbb{E}[|\det(H_N(u))|\mathbb{1}_{i(H_N(u))=1}] \approx \begin{cases} \int \log|\lambda| \mu_\infty(u)(\mathrm{d}\lambda) - I(u, 0) & \text{if } u \in \mathcal{G}, \\ -\infty & \text{otherwise,} \end{cases}$$

and thus one guesses

$$\begin{aligned} &\lim_{N\to\infty} \frac{1}{N} \log \int_{\mathbb{R}^m} e^{-N\frac{\|u\|^2}{2}} \mathbb{E}[|\det(H_N(u))|\mathbb{1}_{i(H_N(u))=1}]\, \mathrm{d}u \\ &= \sup_{u\in\mathcal{G}} \left\{ \int \log|\lambda| \mu_\infty(u)(\mathrm{d}\lambda) - I(u, 0) - \frac{\|u\|^2}{2} \right\}. \end{aligned}$$

Then the complexity of index-one saddle points would be described as this variational problem, plus simpler terms. In fact, if one could prove this, then index-k saddle points for fixed k would likely follow in the same way: Recently, Guionnet and Husson proved that rank-k spherical integrals approximately factor as the product of k rank-one spherical integrals [100]. As they showed, this implies that, for several random matrix ensembles where the largest eigenvalue was already known to satisfy an LDP at speed N with some rate function $I(x)$, in fact the kth largest eigenvalue satisfies an LDP at speed N with the rate function $kI(x)$. (This phenomenon was already known for the GOE, from [10].) This might allow one to give the complexity of index-k saddle points as

$$\sup_{u\in\mathcal{G}}\left\{\int \log|\lambda|\mu_\infty(u)(\mathrm{d}\lambda) - kI(u,0) - \frac{\|u\|^2}{2}\right\}$$

plus simpler terms. For example, for the model described in Theorem 1.3.1 above, this would give the complexity $\Sigma^k(\mu,1)$ of index-k critical points as

$$\Sigma^k(\mu,1) = \frac{1}{2}[-3+4\mu-\mu^2-\log(\mu^2)] = \Sigma^{\min}(\mu,1),$$

meaning that index-k critical points have the same total complexity for every fixed $k \geqslant 0$ (notice this is the same phenomenon as exhibited in pure spherical spin glasses [10]).

Finally, we consider the case $k = \alpha N$ for $\alpha \in (0,1)$. Suppose that the empirical spectral measure $\hat{\mu}_{H_N(u)}$ satisfies an LDP at speed N^2 (or even $N^{1+\varepsilon}$). If we consider the set

$$\mathcal{G}_\alpha = \{u \in \mathbb{R}^m : \mu_\infty(u)((-\infty,0)) = 1-\mu_\infty(u)((0,\infty)) = \alpha\}$$

(which may well be a singleton set), then one guesses

$$\frac{1}{N}\log\mathbb{E}[|\det(H_N(u))|\mathbb{1}_{i(H_N(u))=\alpha N}] \approx \begin{cases} \int \log|\lambda|\mu_\infty(u)(\mathrm{d}\lambda) & \text{if } u\in\mathcal{G}_\alpha, \\ -\infty & \text{otherwise,} \end{cases}$$

and thus one guesses

$$\lim_{N\to\infty}\frac{1}{N}\log\int_{\mathbb{R}^m}e^{-N\frac{\|u\|^2}{2}}\mathbb{E}[|\det(H_N(u))|\mathbb{1}_{i(H_N(u))=\alpha N}]\,\mathrm{d}u=\sup_{u\in\mathcal{G}_\alpha}\left\{\int\log|\lambda|\mu_\infty(u)(\mathrm{d}\lambda)-\frac{\|u\|^2}{2}\right\}.$$

1.7 Open questions

We end with a selection of open questions.

Question 1.7.1. *Establish determinant concentration (1.1.1) when H_N is the adjacency matrix of a random d-regular graph, in any range of parameters d (either fixed or tending to infinity). Via our theorem, it suffices to either (a) prove some version of discrete log-Sobolev for random regular graphs, or (b) find an equality in distribution describing the adjacency matrix of a random d-regular graph as a Lipschitz, convex function of some independent random variables (Bernoullis seem natural, although any description would suffice).*

Question 1.7.2. *Describe the quenched picture for any of the models we study, i.e., compute not* $\lim_{N\to\infty}\frac{1}{N}\log\mathbb{E}[\mathrm{Crt}]$ *but* $\lim_{N\to\infty}\frac{1}{N}\mathbb{E}[\log\mathrm{Crt}]$.

Versions of the Kac-Rice formula allow for computation of finite moments $\mathbb{E}[\mathrm{Crt}^k]$, $k\in\mathbb{N}$, in terms determinants of the form $\mathbb{E}[\prod_{i=1}^k|\det(H^{(i)})|]$, where the $H^{(i)}$'s may be correlated. Our determinant asymptotics extend to such products; this is the motivation of Appendix A. If the second moment matches the first squared, then one can show that the quenched picture is the same as the annealed one (this is the approach, and the result, of Subag [141] and Auffinger-Gold [12] for pure p-spins). But if the second moment does not match the first squared, the situation seems largely intractable: It is not even clear that the distribution is determined by its moments. In the physics literature, Ros, Ben Arous, Biroli, and Cammarota [136] have proposed a method they call replicated Kac-Rice*, which claims to compute the quenched asymptotics even when they do not match the annealed. Can one make this method mathematically rigorous?*

Question 1.7.3. *We have suggested that zero/positive complexity is a good predictor for the*

success/failure, respectively, of local optimization algorithms like Langevin dynamics. Can one study Langevin dynamics directly and prove this? See the work of Ben Arous-Gheissari-Jagannath [40, 39, 41].

Question 1.7.4. *The Kac-Rice formula is technically not restricted to Gaussian processes, but the versions for non-Gaussian processes require many conditions. More importantly, the process of* conditioning *the Hessian on criticality, which is of course routine in the Gaussian case, looks prohibitively difficult for non-Gaussian processes. Can one establish an easy-to-use non-Gaussian Kac-Rice?*

Question 1.7.5. *Can one study complexity of serious machine-learning models (which are often non-Gaussian), and relate the complexity to the performance of these models?*

Question 1.7.6. *A technical random-matrix question: To apply our theorem on determinant concentration to a particular random matrix H_N, we need, as input, an estimate like*

$$\mathbb{P}(H_N \text{ has no eigenvalues in } [E - e^{-N^\varepsilon}, E + e^{-N^\varepsilon}]) = 1 - o(1) \tag{1.7.1}$$

for fixed $E \in \mathbb{R}$. Results of this type should be true under almost no assumptions on H_N (in fact, there should usually be a gap around E of size $o(N^{-1})$ with high probability, and (1.7.1) is asking for much less). But it is not clear what such a minimal proof would be; if this result were known, we could remove many of the regularity assumptions in our results.

Can one find techniques to prove (1.7.1) with minimal assumptions? To give a concrete example, does it hold if H_N is a sample covariance matrix with $2 + \varepsilon$ finite moments and E is in the bulk of the Marčenko-Pastur law?

Question 1.7.7. *For the elastic manifold, we studied the mean-field scaling where L and d are fixed and N tends to infinity. What can be said for non-mean-field scalings where L and d tend to infinity with N? (Or other scalings – for example, Fyodorov, Le Doussal, Rosso, and Texier studied the scaling $d = 1$, $N = 1$, and $L \to +\infty$ [90].) When does the Larkin phase transition persist?*

Question 1.7.8. *For the soft-spins model $\frac{1}{2}\langle u, D_N u\rangle + V_N(u)$, we find the phase transition between positive and zero complexity, and establish the critical exponents governing this phase transition. Can one find more fine-grained information? For example, in the physics literature, Fyodorov and Nadal [91] studied the special case $D_N = \mu\,\mathrm{Id}$ using left- and right-tail asymptotics of the Tracy-Widom function, and argued that (i) for* every *supercritical μ (not "large enough"), the expected number of local minima in the large-N limit is not just subexponential but actually almost* one*; (ii) the non-trivial phase transition occurs at the scale $|\mu - \mu_c| \approx N^{-1/3}$; and (iii) when $\left(\frac{\mu}{\mu_c} - 1\right)N^{1/3} = \delta > 0$, the expected number of local minima as a function of δ, in the large-N limit, approaches some limit shape that can be written using (but is not exactly) the Tracy-Widom distribution.*

What happens in the anisotropic case? In particular (i) does there exist an environment D_N such that, in the trivial phase, the expected number of local minima is subexponential but does actually grow with N; and (ii) what happens when the environment D_N has outliers beyond the BBP phase transition (meaning fluctuations that are not Tracy-Widom)?

Question 1.7.9. *For bipartite spherical spin glasses, the complexity results we can establish are essentially all of the form "phenomena already present in usual (single-species) spin glasses, due to [10], also occur for the bipartite case." Are there other complexity questions for which new phenomena appear in the bipartite case?*

Related: Consider the case of p-partite spherical spin glasses (where spins interact in p groups, instead of the $p = 2$ model studied in this thesis). This is defined on the product of p spheres of dimension order N. Likely our arguments extend to any fixed p. What happens in the non-mean-field scaling $p = p_N \to +\infty$?

Question 1.7.10. *Another question in large deviations: The rank-one spherical integral (1.6.2), which has asymptotics at scale e^N per [101], can establish LDPs for extreme eigenvalues at speed N. Can the full-rank HCIZ integral (1.6.3), which has asymptotics at scale e^{N^2} per [104], be useful for establishing LDPs for empirical spectral measures at speed N^2? Belinschi, Guionnet, and Huang made recent notable progress in this direction, for some matrix models arising in free probability*

[32]. It might also be helpful to interpret the HCIZ integral differently; for example see Novak's recent proof [127] of a longstanding physics conjecture on the relationship between the HCIZ integral, combinatorics, and representation theory.

For concreteness, take the case of symmetric Bernoulli matrices B_N. Here is a tricky observation due to Guionnet. Although the largest eigenvalue of B_N satisfies an LDP with the same rate function as GOE (as a special case of [99]), the empirical spectral measure of B_N cannot *satisfy an LDP at speed N^2 with the same rate function as GOE: The rate function for GOE vanishes at δ_0, but the rate function for B_N cannot, since $\mathbb{P}(B_N = 0) = (\frac{1}{2})^{\frac{N(N+1)}{2}} \approx e^{-N^2 \frac{\log 2}{2}}$. So we do not even have a guess for the rate function.*

Question 1.7.11. *The least well-formed question: What would a discrete complexity theory look like? For example, consider the Sherrington-Kirkpatrick Hamiltonian $\mathcal{H}_N : \{-1, +1\}^N \to \mathbb{R}$. There* is *a natural definition of "local minimum" – namely, "switching each coordinate does not decrease the Hamiltonian" – but it is not clear how to define an index-one saddle point, for example. Worse, there is no Kac-Rice formula, since there are no classical derivatives. But one can still study dynamics on the Sherrington-Kirkpatrick model; is there some scalar observable of the geometry that should predict when these dynamics succeed or fail?*

Roadmap

The organization of the disseration is as follows. In Chapter 2, based on [35], we study the determinants of non-invariant random matrices and prove Theorem 1.1.1. In Chapter 3, based on [36], we study the complexity of the elastic manifold and of soft spins in an anisotropic well, proving Theorems 1.4.1 and 1.4.2. In Chapter 4, based on [120], we study the complexity of bipartite spin glasses and prove Theorem 1.4.3. Finally, in Chapter 5, based on [121], we study large deviations and prove Theorem 1.6.3.

Chapter 2

Exponential growth of random determinants beyond invariance

This chapter is essentially borrowed from [35], joint with Gérard Ben Arous and Paul Bourgade, which will appear on the arXiv soon.

2.1 Introduction

2.1.1 Overview. In this paper, our goal is to study the expected absolute values of the determinants of general $N \times N$ real symmetric random matrices H_N, specifically at exponential scale in the large-N limit:

$$\lim_{N\to\infty} \frac{1}{N} \log \mathbb{E}[|\det(H_N)|]. \tag{2.1.1}$$

We identify two sets of simple criteria that lead to asymptotics of this type (Theorems 2.1.1 and 2.1.2), and apply them to a wide variety of matrix models.

Initiated in the 1930s, and developed early on by Turán, Fortet, Tukey, Nyquist, Rice, Riordan, Prékopa, and others, the study of random determinants has focused on three distinct questions: the singularity probability (that the determinant of a discrete random matrix vanishes), Gaussian

fluctuations, and asymptotics of the type (2.1.1). We will describe this history below in greater detail. The third direction is useful for the topological "landscape complexity" program, which studies the geometry of high-dimensional random functions via the Kac-Rice formula, and which motivates our present work.

Most studies in this direction have focused on the invariant Gaussian ensembles. We study random determinants in contexts where the distribution of the matrix H_N is not necessarily invariant by orthogonal conjugacy, evaluating (2.1.1) for matrix models including Gaussian matrices with variance profiles, large zero blocks, or even correlations; Wigner matrices and sample covariance matrices with near-optimal $2 + \varepsilon$ finite moments; Erdős-Rényi graphs with parameter $p \geqslant N^{\varepsilon}/N$; band matrices with any bandwidth $W \geqslant N^{\varepsilon}$; and the classical free-convolution model $A + OBO^T$ with O uniform on the orthogonal group. For example, denoting ρ_{sc} the semicircle density on $[-2, 2]$, for any E we prove that

$$\lim_{N\to\infty} \frac{1}{N} \log \mathbb{E}[|\det(W_N - E)|] = \int \log|\lambda - E| \rho_{\mathrm{sc}}(\lambda)\, \mathrm{d}\lambda,$$

whenever W_N is a Wigner matrix (Corollary 2.1.3) or a random band matrix (Corollary 2.1.6), under the above moments and bandwidth assumptions.

In the companion papers [36, 120], we use these results to study the landscape complexity of non-invariant random functions. There, we prove formulas of Fyodorov and Le Doussal [88] on the classical "elastic manifold" from statistical physics, which models a point configuration with local self-interactions in a disordered environment. We also find a new phase transition, with universal near-critical behavior, for a certain anisotropic signal-plus-noise model.

In fact, for these geometric applications we need to understand asymptotics like (2.1.1) when the matrix H_N has long-range correlations, for example when all the diagonal entries are correlated with each other. In the last section of this paper, we show how to give exact variational formulas for asymptotics of this type, based on the (simpler) formulas for matrices with short-range correlations.

Theorem 2.1.1 and 2.1.2 below prove that we can obtain the asymptotics (2.1.1) under three

general conditions which do not use invariance, stated informally as follows.

(1) We can discard the contribution of extremely large and small eigenvalues (at scales e^{N^ε} and e^{-N^ε}).

(2) Some form of concentration of the empirical spectral measure $\hat{\mu}_{H_N}$ about its mean $\mathbb{E}[\hat{\mu}_{H_N}]$ holds.

(3) There exists a deterministic sequence $(\mu_N)_{N=1}^\infty$ of probability measures, sufficiently regular, that are mildly good approximations for the mean spectral measure $\mathbb{E}[\hat{\mu}_{H_N}]$.

Overall, our proof strategy is to write the determinant as an almost-continuous test function integrated against $\hat{\mu}_{H_N}$, regularize the logarithm using (1), prove concentration of this test statistic about its mean using (2), and relate this mean to something more recognizable using (3). Checking condition (1) is typically model-specific, but conditions (2) and (3) can be discussed in general.

To prove condition (2) on concentration of $\hat{\mu}_{H_N}$, we identify two distinct criteria, corresponding to our general theorems:

– Either (the *convexity-preserving functional* case, Theorem 2.1.1) H_N is built in a convexity-preserving and Lipschitz way from arbitrary independent random variables,

– or (the *concentrated input* case, Theorem 2.1.2) linear statistics of H_N are already known to concentrate. This is meant to be applied if, e.g., H_N satisfies log-Sobolev, or Gromov-Milman concentration on compact groups.

To prove condition (3) regarding convergence of $\mathbb{E}[\hat{\mu}_{H_N}]$, in the case of classical random matrices the approximating sequence $(\mu_N)_{N=1}^\infty$ is well-known (and in fact constant): For example, one should choose the semicircle law for Wigner matrices, or the Marčenko-Pastur law for sample covariance models. But the good choice of μ_N for non-invariant Gaussian ensembles, which are the most important matrices for applications to complexity, has only been understood recently, a consequence of the theory of the Matrix Dyson Equation (MDE) as developed in [5, 6]. Given nice H_N, the

MDE produces a probability measure μ_N found by solving a constrained problem over matrices. The existence, uniqueness, and regularity theory of the MDE is an important input for our work.

The organization of the paper is as follows: In the rest of this section, we give some history on determinants of random matrices, then state our main results. We prove our general results, Theorems 2.1.1 and 2.1.2, in Section 2.2, then prove our applications to matrix models in Section 2.3. In Section 2.4, we discuss determinants in the presence of long-range correlations. Finally, in Appendix A we extend our results to product of determinants, showing

$$\lim_{N\to\infty} \frac{1}{N} \log \mathbb{E}\left[\prod_{i=1}^{\ell} |\det(H_N^{(i)})|\right] = \sum_{i=1}^{\ell} \left(\lim_{N\to\infty} \frac{1}{N} \log \mathbb{E}[|\det(H_N^{(i)})|]\right) \tag{2.1.2}$$

for any fixed ℓ and random matrices $H_N^{(1)}, \ldots, H_N^{(\ell)}$ which may be correlated with each other. This asymptotic factoring holds regardless of the correlation structure between the $H_N^{(i)}$'s.

2.1.2 History. The earliest research on random-matrix determinants covered non-Hermitian matrices with i.i.d. entries, discussing an extremal problem on the determinant of Bernoulli matrices [142] (extended in [151]) and exact formulas at finite N for small moments of determinants [82, 81, 128, 132] (see also Girko's book [97]). Later in the literature, we identify three main strands of research on determinants.

First, one can ask for the probability that an $N \times N$ discrete matrix (Bernoulli, say) is singular, i.e., that its determinant is zero, for large N. In the non-Hermitian case, Komlós showed that this probability is $o(1)$ [110, 111]. Recently K. Tikhomirov established the long-standing conjecture that this probability is $(\frac{1}{2} + o(1))^N$ [149]; earlier exponential estimates in this direction include [108, 144, 145, 61].

Second, one can show that the determinant, appropriately normalized, has Gaussian fluctuations. In the non-Hermitian case, if the entries are Gaussian this follows from work of Goodman [98]. Gaussianity was replaced by an exponential-tails assumption in [126] and a fourth-moment assumption in [23]. In the Hermitian case, Gaussian matrices were studied in [67]. Gaussianity was

relaxed to a four-moment-matching assumption in [146], then to a two-moment-matching assumption in [59, 60]. Some other ensembles were treated in [63, 138], and more about determinants for Gaussian ensembles was discussed in [56, 72].

Third, one can study the same question we do here, namely the asymptotics of $\mathbb{E}[|\det(H_N)|]$, usually in the same context of studying complexity for high-dimensional random fields. Here we just discuss the types of random matrices that have appeared; for a discussion of what these prior results mean for complexity, we refer to the companion paper [36]. Fyodorov [84] studied Gaussian matrices of type GOE $+ \mathcal{N}(0, 1/N)\,\mathrm{Id}$ using supersymmetry, and a similar model was addressed in Auffinger *et al.* [10] using known large-deviations principles (LDPs) [42, 37]. Rank-one perturbations of GOE appeared in [43], using an LDP of Maïda [118]. An upper bound for full-rank perturbations of GOE appeared in [78], based on free probability and large deviations. Upper and lower bounds for Gaussian matrices with a certain covariance structure were given in [11]. The (Gaussian) real elliptic ensemble was discussed in [38], based on a new result on large deviations for its spectral measure. Baskerville *et al.* cover finite-rank perturbations of GOE in [29] and a specific ensemble of Gaussian matrices with a variance profile, inspired by a two-layer spin-glass model, in [30]. In both cases the determinant analysis is performed through supersymmetry, for the asymptotic spectral density and for Wegner estimates. Our corollaries 2.1.8.A, 2.1.8.B, and 2.1.9 about general Gaussian ensembles provide alternative derivations for all these results about Hermitian matrices. These corollaries also make rigorous the analysis of random determinants by Fyodorov and Le Doussal [88] (see [36] for corresponding complexity results).

Finally, asymptotics for a *pair* of determinants, in the style of (2.1.2) with $\ell = 2$, appeared for a particular pair of random matrices from spin glasses, closely related to correlated GOE matrices, in [141, 12, 44]. These arguments were based on known LDPs for Gaussian ensembles.

Notations. We write $\|\cdot\|$ for the operator norm on elements of $\mathbb{C}^{N\times N}$ induced by the L^2 distance on $\mathbb{C}^N$. We let $\|f\|_{\mathrm{Lip}} = \sup_{x\neq y} \frac{|f(x)-f(y)|_{\mathrm{L}^2}}{|x-y|_{\mathrm{L}^2}}$ for functions $f : \mathbb{R}^m \to \mathbb{R}^n$, and consider the following three distances on probability measures on the real line (called bounded-Lipschitz, Wasserstein-1,

and Kolmogorov-Smirnov, respectively):

$$d_{\mathrm{BL}}(\mu,\nu) = \sup\left\{\left|\int_{\mathbb{R}} f \,\mathrm{d}(\mu-\nu)\right| : \|f\|_{\mathrm{Lip}} + \|f\|_{L^\infty} \leqslant 1\right\},$$
$$\mathrm{W}_1(\mu,\nu) = \sup\left\{\left|\int_{\mathbb{R}} f \,\mathrm{d}(\mu-\nu)\right| : \|f\|_{\mathrm{Lip}} \leqslant 1\right\},$$
$$d_{\mathrm{KS}}(\mu,\nu) = \sup\{|\mu((-\infty,x]) - \nu((-\infty,x])| : x \in \mathbb{R}\}.$$

We normalize the semicircle law as $\rho_{\mathrm{sc}}(\mathrm{d}x) = \frac{\sqrt{4-x^2}}{2\pi}\,\mathbf{1}_{x\in[-2,2]}\,\mathrm{d}x$. We write $\mathtt{l}(\mu)$ for the left edge (respectively, $\mathtt{r}(\mu)$ for the right edge) of a compactly supported measure μ. For an $N\times N$ Hermitian matrix M, we write $\lambda_{\min}(M) = \lambda_1(M) \leqslant \cdots \leqslant \lambda_N(M) = \lambda_{\max}(M)$ for its eigenvalues and $\hat{\mu}_M = \frac{1}{N}\sum_{i=1}^N \delta_{\lambda_i(M)}$ for its empirical measure. We write $\mathscr{S}_N$ for the set of all $N \times N$ realy symmetric matrices, and $\boxplus$ for the free (additive) convolution of probability measures.

We write B_R for the ball of radius R around 0 in the relevant Euclidean space. We use $(\cdot)^T$ for the matrix transpose, which is distinguished both from the matrix conjugate transpose $(\cdot)^*$, and from the matrix trace $\mathrm{Tr}(\cdot)$.

2.1.3 General theorem for convexity-preserving functional. The following Theorem 2.1.1 is our first general result, it applies to random matrices without any a priori concentration hypothesis, but requires the tools of convex analysis, in particular results of Talagrand.

To state the hypotheses, we denote $\kappa > 0$ an arbitrarily small control parameter which does not depend on N. Let $M = M_N \geqslant 1$. Consider $X = (X_1, \ldots, X_M)$ a random vector. We now consider the following set of assumptions.

(I) The X_i's are independent and real-valued.

(M) Matrix model. Let $H = H_N = \Phi(X)$ where $\Phi : \mathbb{R}^M \to \mathscr{S}_N$ is deterministic and Lipschitz and $\Phi^{-1}(A)$ is convex for any convex set A.

(E) Expectation. A sequence of probability measures μ_N exists satisfying the following properties.

First,

$$d_{\mathrm{BL}}(\mathbb{E}\hat{\mu}_{\Phi(X)}, \mu_N) \leqslant N^{-\kappa}. \tag{2.1.3}$$

Moreover, the μ_N's are supported in a common compact set, and each has a density $\mu_N(\cdot)$ in the same neighborhood $(-\kappa, \kappa)$ around 0, which satisfies $\mu_N(x) < \kappa^{-1}|x|^{-1+\kappa}$ for all $|x| < \kappa$ and all N.

(C) Coarse bounds. Write $(\lambda_i)_{i=1}^N$ for the eigenvalues of $\Phi(X)$. For every $\varepsilon > 0$,

$$\lim_{N\to\infty} \frac{1}{N} \log \mathbb{E}\left[\prod_{i=1}^{N} (1 + |\lambda_i| \mathbb{1}_{|\lambda_i| > e^{N^\varepsilon}})\right] = 0, \tag{2.1.4}$$

$$\lim_{N\to\infty} \mathbb{P}(\Phi(X) \text{ has no eigenvalues in } [-e^{-N^\varepsilon}, e^{-N^\varepsilon}]) = 1. \tag{2.1.5}$$

In addition, there exists $\delta > 0$ such that

$$\limsup_{N\to\infty} \frac{1}{N \log N} \log \mathbb{E}[|\det(H_N)|^{1+\delta}] < \infty. \tag{2.1.6}$$

(S) Spectral stability. Let $(X_{\mathrm{cut}})_i = X_i \mathbb{1}_{|X_i| < N^{-\kappa}/\|\Phi\|_{\mathrm{Lip}}}$. We have

$$\lim_{N\to\infty} \frac{1}{N \log N} \log \mathbb{P}\Big(d_{\mathrm{KS}}(\hat{\mu}_{\Phi(X)}, \hat{\mu}_{\Phi(X_{\mathrm{cut}})}) > N^{-\kappa}\Big) = -\infty. \tag{2.1.7}$$

Theorem 2.1.1. ***(Convexity-preserving functional)*** *Under the assumptions (I), (M), (E), (C), (S), we have*

$$\lim_{N\to\infty} \left(\frac{1}{N} \log \mathbb{E}[|\det(H_N)|] - \int_{\mathbb{R}} \log|\lambda| \mu_N(\mathrm{d}\lambda)\right) = 0. \tag{2.1.8}$$

Comments on the result. (i) A polynomial rate in (2.1.4) is enough to give a polynomial rate of convergence

$$\left|\frac{1}{N} \log \mathbb{E}[|\det(H_N)|] - \int_{\mathbb{R}} \log|\lambda| \mu_N(\mathrm{d}\lambda)\right| \leqslant N^{-\varepsilon}$$

for some $\varepsilon > 0$ and $N \geqslant N_0(\varepsilon)$. Indeed, an examination of the proof shows that ε depends only on κ and the polynomial rate in (2.1.4), but $N_0(\varepsilon)$ also depends on the rates of convergence in (2.1.5) and (2.1.7), and on the permissible values of δ and the value of the lim sup in (2.1.6).

(ii) The matrix H_N does not need to be centered. As an elementary example, we can choose $H_N = W_N - E$ for W_N a Wigner matrix and obtain concentration around $\int_{\mathbb{R}} \log|\lambda - E| \rho_{\mathrm{sc}}(\lambda)\, \mathrm{d}\lambda$; see Corollary 2.1.3 below.

(iii) The proof uses Talagrand's concentration inequality for product measures. We want to recognize the determinant almost as a Lipschitz, convex function of independent, bounded random variables. Ideally these would be the X_i's, but they are not bounded; however, we truncate them using assumption (S). The functional $H = \Phi(X)$ gives the Lipschitz, convex condition, after regularizing the logarithm using assumption (C).

Comments on the assumptions. We discuss briefly why our assumptions are reasonable and close to optimal. In our applications, Φ is linear so Assumption (M) is trivially satisfied, but Φ is also allowed to create correlations between the entries in a nonlinear fashion. Equation (2.1.5) avoids a non-trivial kernel, an obviously necessary condition for (2.1.8). Equation (2.1.6) asks for slightly more integrability than finiteness of $\limsup N^{-1} \log \mathbb{E}[|\det H_N|]$ which is implied by the result and assumption (E). In Section 2.3.10 we show the importance (2.1.4) (which is a constraint on large eigenvalues) and Assumption (S) (which essentially states that the spectrum should not depend too much on a small number of X_i's): for each of these, we give an example of a distribution on matrices satisfying every *other* assumption but not this one, for which the result of the theorem fails.

2.1.4 General theorem for concentrated input. Here we consider the problem of exponential growth for random matrices H_N that already satisfy some concentration property directly, without having to cut the tails and apply a result of Talagrand as in (the proof of) Theorem 2.1.1. For example, in applications we will take matrices whose upper triangles satisfy a log-Sobolev inequality (even if correlated), or Gromov-Milman-type concentration. We remark that the dichotomy

in Theorems 2.1.1 and 2.1.2 – namely, proving the similar results, once under product-measure assumptions and once under log-Sobolev-style assumptions – first appeared in the classic concentration paper of Guionnet-Zeitouni [103]. We have termed these models "concentrated input," to contrast with the previous section's "convexity-preserving functional" where H_N is written as $\Phi(X)$ and concentration is provided by convexity-preserving properties of Φ (and tail bounds). In this section, we will therefore consider H_N directly. We will also replace some of the assumptions above with the following.

(W) Wasserstein-1. A sequence of probability measures μ_N exists satisfying the following properties. First,

$$\mathrm{W}_1(\mathbb{E}\hat{\mu}_{H_N}, \mu_N) \leqslant N^{-\kappa}. \tag{2.1.9}$$

Moreover, the μ_N's are supported in a common compact set, and each has a density $\mu_N(\cdot)$ in the same neighborhood $(-\kappa, \kappa)$ around 0, which satisfies $\mu_N(x) < \kappa^{-1}|x|^{-1+\kappa}$ for all $|x| < \kappa$ and all N.

(L) Concentration for Lipschitz traces. There exists $\varepsilon_0 > 0$ with the following property: For every $\zeta > 0$, there exists $c_\zeta > 0$ such that, whenever $f : \mathbb{R} \to \mathbb{R}$ is Lipschitz, we have for every $\delta > 0$

$$\mathbb{P}\left(\left|\frac{1}{N}\operatorname{Tr}(f(H_N)) - \frac{1}{N}\mathbb{E}[\operatorname{Tr}(f(H_N))]\right| > \delta\right) \leqslant \exp\left(-\frac{c_\zeta}{N^\zeta}\min\left\{\left(\frac{N\delta}{\|f\|_{\mathrm{Lip}}}\right)^2, \left(\frac{N\delta}{\|f\|_{\mathrm{Lip}}}\right)^{1+\varepsilon_0}\right\}\right). \tag{2.1.10}$$

On a first pass readers can drop the $N^{-\zeta}$ factor in (2.1.10). It is included because, for Gaussian matrices as in Section 2.1.9, our assumption on the correlation structure implies (2.1.10) for every $\zeta > 0$ but not necessarily for $\zeta = 0$.

Theorem 2.1.2. ***(Concentrated input)*** *Under the assumptions (W), (L), and the gap assumption (2.1.5), we have*

$$\lim_{N\to\infty}\left(\frac{1}{N}\log\mathbb{E}[|\det(H_N)|] - \int_{\mathbb{R}}\log|\lambda|\mu_N(\mathrm{d}\lambda)\right) = 0.$$

As in Theorem 2.1.1, by examining the proof one can find a small polynomial rate $N^{-\varepsilon}$ in Theorem 2.1.2.

Compared to [103], we do not require bounded entries in Theorem 2.1.1, our matrix models are more general, and we consider logarithmic singularities. On the other hand, [103] identifies the correct *scale* of fluctuations, analogous to a rate of convergence of order N^{-1} in (2.1.8), for test functions without singularities.

2.1.5 Wigner matrices. We now discuss determinant asymptotics for Wigner matrices W_N with $2+\varepsilon$ finite moments. This is almost optimal, up to the ε, in the sense that $\mathbb{E}(|W_{12}|^2) = +\infty$ implies $\mathbb{E}[|\det(W_N)|] = +\infty$ (we give a short proof of this fact in Section 2.3.3 below). It would be interesting to study the case of Wigner matrices with exactly two finite moments, or the intermediate regime of an $N \times N$ matrix with $2+\varepsilon_N$ finite moments as $\varepsilon_N \to 0$.

Fix some $\varepsilon > 0$, and let μ be a centered probability measure on $\mathbb{R}$ with $2+\varepsilon$ finite moments and unit variance. Let W_N be a real symmetric $N \times N$ Wigner matrix associated with μ, by which we mean that the entries of $\sqrt{N}W_N$ are independent up to symmetry and each distributed according to μ. The following corollary uses Theorem 2.1.1.

Corollary 2.1.3. ***(Wigner matrices with*** $2+\varepsilon$ ***moments)*** *For every $E \in \mathbb{R}$ we have*

$$\lim_{N\to\infty} \frac{1}{N} \log \mathbb{E}[|\det(W_N - E)|] = \int_{\mathbb{R}} \log|\lambda - E| \rho_{sc}(\lambda)\, \mathrm{d}\lambda.$$

An examination of the proof shows local uniformity in E, meaning that for every compact $K \subset \mathbb{R}$ we have

$$\lim_{N\to\infty} \sup_{E\in K} \left(\frac{1}{N} \log \mathbb{E}[|\det(W_N - E)|] - \int_{\mathbb{R}} \log|\lambda - E| \rho_{\mathrm{sc}}(\lambda)\, \mathrm{d}\lambda \right) = 0.$$

Remark 2.1.4. *One would also be interested in results of the form*

$$\lim_{N\to\infty} \frac{1}{N} \log \mathbb{E}[|\det(W_N + D_N)|] = \int_{\mathbb{R}} \log|\lambda| (\rho_{sc} \boxplus \mu_D)(\mathrm{d}\lambda), \tag{2.1.11}$$

where $(D_N)_{N=1}^{\infty}$ is a sequence of deterministic matrices whose empirical measures tend to some compactly-supported μ_D (at some polynomial speed and without outliers, say). Our techniques could likely be extended to prove such a result under the assumption of $2+\varepsilon$ moments on the Wigner matrices. We do not pursue this direction further here; however, in the companion paper [36], we prove (2.1.11) with a different approach when W_N is a GOE matrix. For a related problem, see the free-addition model below, in Corollary 2.1.10.

2.1.6 Erdős-Rényi matrices. We now consider Erdős-Rényi matrices with near-optimal sparsity parameter $p \geqslant N^{\varepsilon}/N$, i.e., when each vertex has expected degree N^{ε}. It is classical that the limiting spectral distribution of such matrices is semicircular as long as $p = \omega(1/N)$ (see, e.g., [150]), but not semicircular anymore if $p = \alpha/N$ for α fixed (see, e.g., [31]).

Fix some $\varepsilon > 0$, and let H_N be an $N \times N$ Erdős-Rényi random matrix with parameter $1-\varepsilon \geqslant p_N \geqslant \frac{N^{\varepsilon}}{N}$. scaled so that the bulk eigenvalues are order one. This means that the entries are independent up to symmetry and

$$(H_N)_{ij} = \frac{1}{\sqrt{Np_N(1-p_N)}} \begin{cases} 1 & \text{with probability } p_N, \\ 0 & \text{with probability } 1-p_N. \end{cases}$$

The following corollary uses Theorem 2.1.1.

Corollary 2.1.5. ***(Erdős-Rényi matrices with $p \geqslant N^{\varepsilon}/N$)*** *For any $E \in \mathbb{R}$ with $|E| \neq 2$ we have*

$$\lim_{N\to\infty} \frac{1}{N} \log \mathbb{E}[|\det(H_N - E)|] = \int_{\mathbb{R}} \log|\lambda - E| \rho_{sc}(\lambda)\, \mathrm{d}\lambda.$$

This result is locally uniform for E away from the edges, meaning E in any compact subset of $\mathbb{R} \setminus \{-2, 2\}$.

2.1.7 Band matrices. In this section we consider random band matrices H_N, i.e., matrices whose (i,j)th entry is zero unless i and j are less than some W apart. Many statistics of H_N are

believed to undergo a phase transition at $W \sim N^{1/2}$. For example, the eigenvectors are supposed to be localized on $o(N)$ sites for $W \ll N^{1/2}$ and delocalized for $W \gg N^{1/2}$. However, we establish that the determinant asymptotics do not see this phase transition: They are the same as long as $W \to +\infty$ polynomially in N. For a full discussion, we direct the reader to [57].

Let μ be a centered probability measure with unit variance that has subexponential tails, in the sense that there exist constants $\alpha, \beta > 0$ such that, if $X \sim \mu$, then

$$\mathbb{P}(|X| \geqslant t^{\alpha}) \leqslant \beta e^{-t}$$

for all $t > 0$. Suppose also that μ has a bounded density $\mu(\cdot)$. Fix any $\varepsilon > 0$. Let H_N be an $N \times N$ band matrix with bandwidth $W = W_N \geqslant N^{\varepsilon}$ corresponding to μ. This means that H_N has independent entries up to symmetry with

$$(H_N)_{ij} \begin{cases} = 0 & \text{if } \|i-j\| > W, \\ \sim \frac{X}{\sqrt{2W+1}} & \text{if } \|i-j\| \leqslant W. \end{cases}$$

(Here we take periodic distance $\|i-j\| = \min(|i-j|, N-|i-j|)$.) The following corollary uses Theorem 2.1.1.

Corollary 2.1.6. ***(One-dimensional band matrices with bandwidth*** $W \geqslant N^{\varepsilon}$***)*** *Under the above assumptions,*

$$\lim_{N\to\infty} \frac{1}{N} \log \mathbb{E}[|\det(H_N - E)|] = \int_{\mathbb{R}} \log|\lambda - E| \rho_{sc}(\lambda)\, \mathrm{d}\lambda.$$

This result is locally uniform in E.

We now comment on the significance of this result. In the companion paper [36], we solve a problem of Fyodorov-Le Doussal [88] on a model called the "elastic manifold." They consider a mean-field version of this model, corresponding to block-banded random matrices with bandwidth order N, and find the "Larkin mass" separating ordered and disordered phases. An important open

problem is the behavior of the elastic manifold beyond mean field, when the corresponding random matrix is block-banded with sublinear bandwidth. It does not seem to be clear in which regimes this Larkin transition should persist, but Corollary 2.1.6 may suggest that the transition remains for any polynomial bandwidth.

Comment on the assumptions. We require subexponential tails in order to use the bulk local law of Erdős *et al.* [77] and the extreme-eigenvalue estimates of Benaych-Georges/Péché [46]. The bounded density lets us prove the Wegner estimate (2.1.5) to control eigenvalues close to 0, with the Schur complement formula. We believe that the conclusion holds under weaker assumptions.

2.1.8 Sample covariance matrices. Let μ be a centered probability measure on $\mathbb{R}$ with unit variance, finite moment of order $2+\varepsilon$ for some $\varepsilon > 0$. We assume μ has density $f = e^{-g}$ with f smooth enough in the sense that, for any $a \geqslant 1$, there exists $C_a > 0$ such that for any $s \in \mathbb{R}$

$$|\widehat{f}(s)| + |\widehat{fg''}(s)| \leqslant \frac{C_a}{(1+s^2)^a}. \tag{2.1.12}$$

Let $Y_{p,N}$ be a $p \times N = p_N \times N$ matrix whose entries are independent copies of μ. Suppose that

$$\gamma = \lim_{N\to\infty} \frac{p_N}{N} \in (0,1].$$

If $\gamma < 1$, we require a mild speed-of-convergence assumption

$$\left|\gamma - \frac{p_N}{N}\right| \leqslant N^{-\varepsilon} \tag{2.1.13}$$

for some $\varepsilon > 0$; if $\gamma = 1$, then for technical reasons we require $p_N = N$, i.e., we require the matrices to be exactly square rather than asymptotically square. Write $\mu_{\mathrm{MP},\gamma}$ for the Marčenko-Pastur distribution

$$\mu_{\mathrm{MP},\gamma}(\mathrm{d}x) = \frac{\sqrt{(b_\gamma - x)(x - a_\gamma)}}{2\pi\gamma x} \mathbb{1}_{[a_\gamma, b_\gamma]}\,\mathrm{d}x \tag{2.1.14}$$

where $a_\gamma = (1-\sqrt{\gamma})^2$, $b_\gamma = (1+\sqrt{\gamma})^2$.

Proposition 2.1.7. ***(Sample covariance matrices with*** $2+\varepsilon$ ***moments)*** *Under the above assumptions, for every* $E \in \mathbb{R}$*, we have*

$$\lim_{N\to\infty} \frac{1}{p_N} \log \mathbb{E}\left[\left|\det\left(\frac{1}{N} Y_{p,N}(Y_{p,N})^T - E\right)\right|\right] = \int \log|\lambda - E| \mu_{\mathrm{MP},\gamma}(\lambda)\, \mathrm{d}\lambda.$$

We call this a "proposition" instead of a "corollary" because it is not a direct consequence of our theorems, but rather can be proved in a similar way. We give details in Section 2.3.6. The proof also shows, as usual, that the limit holds locally uniformly in E.

Proposition 2.1.7 complements a 1989 result of Dembo [68], who gave an exact formula at finite N for the averaged determinant in the special case $E = 0$, without requiring the assumption of a bounded density. In our normalization, he showed by a combinatorial method that

$$\mathbb{E}\left[\left|\det\left(\frac{1}{N} Y_{p,N}(Y_{p,N})^T\right)\right|\right] = \mathbb{E}\left[\det\left(\frac{1}{N} Y_{p,N}(Y_{p,N})^T\right)\right] = \frac{N!}{N^p (N-p)!},$$

and one can check from the known log-potential of the Marčenko-Pastur law that

$$\lim_{N\to\infty} \frac{1}{N} \log\left(\frac{N!}{N^p (N-p)!}\right)$$

is the same as given by our proposition.

2.1.9 Gaussian matrices with a (co)variance profile. Let H_N be an $N \times N$ real symmetric Gaussian matrix, possibly with a mean, a variance profile, and/or correlated entries, satisfying the technical assumptions below. These are essentially the assumptions needed for the local law of Erdős *et al.* [75] which we will use in the proof. We first give an easier statement for matrices with independent entries up to symmetry (Corollary 2.1.8.A), then a more involved statement for matrices with correlations (Corollary 2.1.8.B). In the statement, we decompose $H_N = A_N + W_N$ where $A_N = \mathbb{E}[H_N]$. These corollaries use Theorem 2.1.2.

In the following mean-field conditions, the arbitrary parameter $p > 0$ is fixed.

(B) Bounded mean. We have $\sup_N \|A_N\| < \infty$.

(F) Flatness. For each N,

$$T \in \mathbb{C}^{N\times N}, \quad T \text{ positive semi-definite} \quad \Longrightarrow \quad \frac{1}{p}\frac{\operatorname{Tr}(T)}{N} \leqslant \mathbb{E}[W_N T W_N] \leqslant p\frac{\operatorname{Tr}(T)}{N}.$$

Let μ_N be the measure from the size-N Matrix Dyson Equation, that is, the measure with density $\mu_N(\cdot)$ whose Stieltjes transform at $z \in \mathbb{H}$ is $\frac{1}{N}\operatorname{Tr}(M_N(z))$, where $M_N(z)$ is the (unique, deterministic) solution to the following constrained equation over $\mathbb{C}^{N\times N}$:

$$\begin{aligned} &\mathrm{Id}_{N\times N} + (z\,\mathrm{Id}_{N\times N} - A_N + \mathbb{E}[W_N M_N(z) W_N])M_N(z) = 0 \\ &\text{subject to} \quad \operatorname{Im} M_N(z) = \frac{M_N(z) - M_N(z)^*}{2\mathrm{i}} > 0 \quad \text{in the sense of quadratic forms.} \end{aligned}$$

Corollary 2.1.8.A. ***(Gaussian matrices with a variance profile)*** *If H_N has independent entries up to symmetry, then under assumptions (B) and (F) we have*

$$\lim_{N\to\infty}\left(\frac{1}{N}\log\mathbb{E}[|\det(H_N)|] - \int_{\mathbb{R}}\log|\lambda|\mu_N(\lambda)\,\mathrm{d}\lambda\right) = 0.$$

The following assumptions are needed if H_N has correlations among its entries beyond the symmetry constraints.

(wF) Weak fullness. Whenever $T \in \mathbb{R}^{N\times N}$ is real symmetric,

$$\mathbb{E}\Big[(\operatorname{Tr}(BW))^2\Big] \geqslant N^{-1-p}\operatorname{Tr}(B^2).$$

(The $p = 0$ case is called "fullness" in [5].)

(D) Decay of correlations. Write κ for multivariate cumulants (for any number of arguments), and consider the distance on subsets of $[\![1, N]\!]^2$ given by $d(A, B) = \min\{\min\{|\alpha - \beta|, |\alpha^t - \beta|\} : \alpha \in A, \beta \in B\}$ where $(\cdot)^t$ switches the elements of an ordered pair. For the order-two

cumulants we assume

$$|\kappa(f_1(W_N), f_2(W_N))| \leqslant \frac{C}{1 + d(\operatorname{supp} f_1, \operatorname{supp} f_2)^s} \|f_1\|_2 \|f_2\|_2$$

for some $s > 12$ and all L^2 functions f_1, f_2 on $N \times N$ matrices. For order-k cumulants, $k \geqslant 3$, we consider, for any L^2 functions $f_1, \dots, f_k$, the complete graph on $\{1, \dots, k\}$ with the edge-weights $d(\{i,j\}) = d(\operatorname{supp} f_i, \operatorname{supp} f_j)$. Writing $T_{\min}$ for the minimal spanning tree on this graph (i.e., smallest sum of edge weights) and lifting covariance to edges as $\kappa(\{i,j\}) = \kappa(f_i, f_j)$, we assume

$$|\kappa(f_1(W_N), \dots, f_k(W_N))| \leqslant C_k \prod_{e \in E(T_{\min})} |\kappa(e)|.$$

(In fact, our results hold under some weaker correlation-decay conditions that are longer to state; see [75, Example 2.12].)

Corollary 2.1.8.B. ***(Gaussian matrices with a (co)variance profile)*** *Under assumptions (B), (F), (wF), and (D), we have*

$$\lim_{N \to \infty} \left(\frac{1}{N} \log \mathbb{E}[|\det(H_N)|] - \int_{\mathbb{R}} \log|\lambda| \mu_N(\lambda)\, \mathrm{d}\lambda \right) = 0.$$

Corollary 2.1.8.A is an immediate consequence of Corollary 2.1.8.B, because it is easy to check that (F) implies both (wF) and (D) if H_N has independent entries up to symmetry. In Section 2.3.7 we therefore only prove Corollary 2.1.8.B.

In some cases one can show that the sequence $(\mu_N)_{N=1}^{\infty}$ has a limit μ_∞, and obtain

$$\lim_{N \to \infty} \frac{1}{N} \log \mathbb{E}[|\det(H_N)|] = \int \log|\lambda| \mu_\infty(\mathrm{d}\lambda).$$

Notice this does not follow from our assumptions, because we do not assume any consistency in N. For example, this corollary applies to the contrived example $H_N = \mathrm{GOE} + (-1)^N \mathrm{Id}$. In the

companion paper [36], we show how to use the (well-established) stability theory of the Matrix Dyson Equation to find a limit μ_∞ when it exists.

2.1.10 Block-diagonal Gaussian matrices. In this section, we are interested in Gaussian random matrices with large zero blocks. These are *not* covered by Corollary 2.1.8.B, since the "flatness" assumption there implies that all entries have variance in some $[\frac{c}{N}, \frac{C}{N}]$. In the landscape complexity program, such block-diagonal matrices describe random functions whose components in certain directions are *independent* of those in other directions. In the companion paper [36], we study one such random function from statistical physics, called the "elastic manifold."

Consider matrices $H_N = A_N + W_N$, with $A_N = \mathbb{E}[H_N]$, that have the following special form. Fix once and for all some $K \in \mathbb{N}$ (the number of blocks), and consider matrices in $\mathbb{R}^{K\times K} \otimes \mathbb{R}^{N\times N}$, i.e., matrices with K^2 blocks each of size $N \times N$. Write E_{ii} for the matrix with a one in the (i,i)th entry and zeros otherwise; depending on the context this will be either an $N \times N$ matrix or a $K \times K$ matrix.

(MS) Bounded mean structure. Consider a deterministic triangular array $(a_i)_{i=1}^N = (a_{i,N})_{i=1}^N$ with each $a_i \in \mathbb{R}^{K\times K}$, and define

$$A_N = \sum_{i=1}^{N} a_i \otimes E_{ii}.$$

In particular A_N can only have nonzero entries on the diagonals of each block. Assume

$$\sup_N \|A_N\| < \infty.$$

(MF) Mean-field randomness in diagonal blocks. The Gaussian random matrix W_N has the form

$$W_N = \sum_{i=1}^{K} E_{ii} \otimes X_i = \begin{pmatrix} X_1 & 0 & \cdots & 0 \\ 0 & X_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & X_K \end{pmatrix},$$

where the X_i's are independent $N \times N$ Gaussian random matrices, each of which has centered independent entries up to symmetry. Write $x_{jk}^{(i)}$ for the (j,k)th entry of X_i and $s_{jk}^{(i)}$ for its variance. For some parameter p, each $i \in [\![1, K]\!]$, and each $j, k \in [\![1, N]\!]$, we have

$$s_{jk}^{(i)} \leqslant \frac{p}{N}, \quad s_{jj}^{(i)} \geqslant \frac{1}{pN}.$$

Notice the lower bound is only along the diagonal.

(R) Regularity of MDE solution. Given $\mathbf{r} = (r_1, \ldots, r_N) \in (\mathbb{C}^{K \times K})^N$, define

$$\mathscr{S}_i[\mathbf{r}] = \sum_{k=1}^{N} \sum_{j=1}^{K} s_{ik}^{(j)} E_{jj} r_k E_{jj} \in \mathbb{C}^{K \times K} \tag{2.1.15}$$

for each $i \in [\![1, N]\!]$. The MDE in this context is a system of N coupled equations over $K \times K$ matrices; we seek the unique solution $\mathbf{m}(z) = \mathbf{m}^{(N)}(z) = (m_1(z), \ldots, m_N(z)) = (m_1^{(N)}(z), \ldots, m_N^{(N)}(z)) \in (\mathbb{C}^{K \times K})^N$ to

$$\begin{aligned} &\mathrm{Id}_{K \times K} + (z \,\mathrm{Id}_{K \times K} - a_i + \mathscr{S}_i[\mathbf{m}(z)]) m_i(z) = 0 \\ &\text{subject to} \quad \operatorname{Im} m_i(z) > 0 \quad \text{as a quadratic form.} \end{aligned} \tag{2.1.16}$$

Consider the probability measure μ_N on $\mathbb{R}$ whose Stieltjes transform is given at the point z by $\frac{1}{NK} \sum_{j=1}^{N} \operatorname{Tr} m_j(z)$.

Assume that each μ_N admits a density with respect to Lebesgue measure, and that these densities are bounded in L^∞, uniformly over N.

The following corollary uses Theorem 2.1.2.

Corollary 2.1.9. ***(Block-diagonal Gaussian matrices)*** *Under assumptions (MS), (MF), and (R), we have*

$$\lim_{N \to \infty} \left(\frac{1}{NK} \log \mathbb{E}[|\det(H_N)|] - \int_{\mathbb{R}} \log|\lambda| \mu_N(\lambda) \, \mathrm{d}\lambda \right) = 0.$$

(The normalization is $\frac{1}{NK}$ because H_N is an $NK \times NK$ matrix.)

In applications to landscape complexity, the description of these measures μ_N via the MDE is very important to prove properties of the limit measures μ_∞. For example, in our companion paper [36], we use this MDE description to identify a crucial convexity property in a variational problem.

2.1.11 Free addition. Let $(A_N)_{N=1}^\infty$, $(B_N)_{N=1}^\infty$ be a sequence of deterministic, $N \times N$, real diagonal matrices, whose empirical measures tend to some μ_A, μ_B respectively. We will be interested in the random matrix $A_N + O_N B_N O_N^T$, where O_N is sampled from Haar measure on the orthogonal group $\mathcal{O}_N$.

We require the following assumptions.

- The measures μ_A and μ_B admit densities ρ_A and ρ_B, respectively. These densities have single nonempty interval supports $[E_-^A, E_+^A]$ and $[E_-^B, E_+^B]$, and each density is strictly positive on the interior of its support.

- Each measure μ_A and μ_B has a power-law behavior with exponent in $(-1, 1)$ at each of its edges; that is, there exist $\delta > 0$ and exponents $-1 < t_-^A, t_-^B, t_+^A, t_+^B < 1$ such that, for some $C > 1$,

$$\begin{aligned}
C^{-1} \leqslant \frac{\rho_A(x)}{(x - E_-^A)^{t_-^A}} \leqslant C \quad &\text{for all } x \in [E_-^A, E_-^A + \delta], \\
C^{-1} \leqslant \frac{\rho_B(x)}{(x - E_-^B)^{t_-^B}} \leqslant C \quad &\text{for all } x \in [E_-^B, E_-^B + \delta], \\
C^{-1} \leqslant \frac{\rho_A(x)}{(E_+^A - x)^{t_+^A}} \leqslant C \quad &\text{for all } x \in [E_+^A - \delta, E_+^A], \\
C^{-1} \leqslant \frac{\rho_B(x)}{(E_+^B - x)^{t_+^B}} \leqslant C \quad &\text{for all } x \in [E_+^B - \delta, E_+^B].
\end{aligned}$$

- One of the measures μ_A and μ_B has a bounded Stieltjes transform.

- The eigenvalues $(a_i)_{i=1}^N = (a_i^{(N)})_{i=1}^N$ of A_N, ordered increasingly, are close to the classical

particle locations a_i^* defined by

$$a_i^* = \inf\left\{ s : \int_{-\infty}^{s} \mu_A(\mathrm{d}y) = i/N \right\}$$

in the sense that for any $c > 0$, $\sup_{1 \leqslant i \leqslant N} |a_i - a_i^*| \leqslant N^{-1+c}$ for N sufficiently large. The analogous condition also holds for the eigenvalues of B_N.

For example, all of these assumptions are satisfied if μ_A is the semicircle law and μ_B is either a uniform measure, the Marčenko-Pastur law, or the semicircle law; and if A_N and B_N store the relevant $\frac{1}{N}$-quantiles.

The following corollary uses Theorem 2.1.2.

Corollary 2.1.10. ***(Free addition)*** *If O_N is chosen randomly from the Haar measure on the orthogonal group $\mathcal{O}_N$, then whenever E is not an edge of $\mu_A \boxplus \mu_B$, we have*

$$\lim_{N\to\infty} \frac{1}{N} \log \mathbb{E}[\left|\det(A_N + O_N B_N O_N^T - E)\right|] = \int_{\mathbb{R}} \log|\lambda - E| (\mu_A \boxplus \mu_B)(\lambda)\, \mathrm{d}\lambda.$$

This result is locally uniform in E away from the edge, meaning uniform in any compact subset of $\mathbb{R} \setminus \{\mathtt{l}(\mu_A \boxplus \mu_B), \mathtt{r}(\mu_A \boxplus \mu_B)\}$.

Comment on the assumptions. For the proof, we check the assumptions of the concentrated-input Theorem 2.1.2 using the local law of Bao-Erdős-Schnelli [22] and the fixed-energy universality of Che-Landon [66]. For concise writing, the assumptions we state here are a bit stronger than "the union of the assumptions of these two papers," but in fact this union suffices for Corollary 2.1.10. In fact, our result likely holds under even weaker assumptions than required in these papers, which handle more fine-grained questions.

Acknowledgements. We wish to thank László Erdős and Torben Krüger for many helpful discussions about the Matrix Dyson Equation. BM also thanks Krishnan Mody for helpful discussions. GBA acknowledges support by the Simons Foundation collaboration Cracking the Glass Problem,

PB was supported by NSF grant DMS-1812114 and a Poincaré chair, and BM was supported by NSF grant DMS-1812114.

2.2 Proofs of determinant asymptotics

2.2.1 Proof of Theorem 2.1.1. The proof depends on a careful tuning of many N-dependent parameters; in the next section we define these parameters and prove some estimates that are common to both the upper and lower bounds. In the following subsections we then prove these upper and lower bounds in order.

2.2.1.1 Definitions and common estimates. Let κ be as in the assumptions (i.e., given to us), and write K, η, t, w_b, p_b for some N-dependent parameters. In fact we will choose

$$\begin{cases} K = e^{N^\varepsilon} \quad \text{for some } \varepsilon \text{ small enough } (\varepsilon = \kappa^2/16 \text{ suffices}), \\ \eta = N^{-\kappa/2}, \\ t = N^{-\kappa/4}, \\ w_b = N^{-\kappa/4}, \\ p_b = N^{-\kappa^2/8}, \end{cases} \tag{2.2.1}$$

but we find it more transparent to work with the names K, η, and so on for the bulk of the proof, checking only at the end that these specific choices make the error estimates useful. We will work with the following regularizations of the logarithm:

$$\log_\eta(\lambda) = \log|\lambda + \mathrm{i}\eta|,$$
$$\log_\eta^K(\lambda) = \min(\log_\eta(\lambda), \log_\eta(K)).$$

Let $b = b_N : \mathbb{R} \to \mathbb{R}$ be some smooth, even, nonnegative function that is identically one on $[-w_b, w_b]$, vanishes outside of $[-2w_b, 2w_b]$, and is $\frac{1}{w_b}$-Lipschitz. Consider the following events:

$$\begin{aligned}
\mathcal{E}_{\text{gap}} &= \{\Phi(X) \text{ has no eigenvalues in } [-e^{-N^\varepsilon}, e^{-N^\varepsilon}]\}, \\
\mathcal{E}_{\text{ss}} &= \{d_{\text{KS}}(\hat{\mu}_{\Phi(X)}, \hat{\mu}_{\Phi(X_{\text{cut}})}) \leqslant N^{-\kappa}\}, \\
\mathcal{E}_{\text{conc}} &= \left\{\left|\int \log_\eta^K \mathrm{d}(\hat{\mu}_{\Phi(X_{\text{cut}})} - \mathbb{E}[\hat{\mu}_{\Phi(X_{\text{cut}})}])\right| \leqslant t\right\}, \\
\mathcal{E}_b &= \left\{\int b \, \mathrm{d}\hat{\mu}_{\Phi(X)} \leqslant p_b\right\}.
\end{aligned} \tag{2.2.2}$$

It turns out that all of these events are likely. For $\mathcal{E}_{\text{gap}}$ and $\mathcal{E}_{\text{ss}}$ this is by assumption; we will prove that $\mathcal{E}_{\text{conc}}$ and $\mathcal{E}_b$ are likely below.

Now we collect some estimates which will be useful for both the upper and lower bounds.

Lemma 2.2.1. *We have*

$$\left|\int \log_\eta^K \mathrm{d}(\hat{\mu}_{\Phi(X)} - \hat{\mu}_{\Phi(X_{\text{cut}})})\right| \mathbb{1}_{\mathcal{E}_{\text{ss}}} \leqslant N^{-\kappa} \log\left(1 + \frac{K^2}{\eta^2}\right).$$

Proof. The proof of [55, Lemma C.2] shows that, if $\hat{\mu}_A$ and $\hat{\mu}_B$ are empirical measures of matrices A and B (which have the same size as each other) and if f is a test function of bounded variation, then

$$\left|\int f \, \mathrm{d}\hat{\mu}_A - \int f \, \mathrm{d}\hat{\mu}_B\right| \leqslant \|f\|_{\text{TV}} \cdot d_{\text{KS}}(\hat{\mu}_A, \hat{\mu}_B).$$

Then the result follows from the computation $\| \log_\eta^K \|_{\text{TV}} = \log\left(1 + \frac{K^2}{\eta^2}\right)$ and the definition of $\mathcal{E}_{\text{ss}}$. □

Lemma 2.2.2. *With*

$$\varepsilon_1(N) := N^{-\kappa} \log\left(1 + \frac{K^2}{\eta^2}\right) + 2\| \log_\eta^K \|_\infty \mathbb{P}((\mathcal{E}_{\text{ss}})^c) + \left(\frac{1}{2\eta} + \| \log_\eta^K \|_\infty\right) N^{-\kappa},$$

we have

$$\left|\int \log_\eta^K \mathrm{d}(\mathbb{E}[\hat{\mu}_{\Phi(X_{\text{cut}})}] - \mu_N)\right| \leqslant \varepsilon_1(N).$$

Proof. First, by inserting $\mathbb{1}_{\mathcal{E}_{\mathrm{ss}}}$ and using Lemma 2.2.1, we find

$$\left|\int \log_\eta^K \mathrm{d}(\mathbb{E}[\hat{\mu}_{\Phi(X_{\mathrm{cut}})}] - \mathbb{E}[\hat{\mu}_{\Phi(X)}])\right| \leqslant N^{-\kappa} \log\left(1 + \frac{K^2}{\eta^2}\right) + 2\| \log_\eta^K \|_\infty \mathbb{P}((\mathcal{E}_{\mathrm{ss}})^c).$$

Next, since $\log_\eta^K$ is $\frac{1}{2\eta}$-Lipschitz, (2.1.3) yields

$$\left|\int \log_\eta^K \mathrm{d}(\mathbb{E}[\hat{\mu}_{\Phi(X)}] - \mu_N)\right| \leqslant \left(\frac{1}{2\eta} + \| \log_\eta^K \|_\infty\right) d_{\mathrm{BL}}(\mathbb{E}[\hat{\mu}_{\Phi(X)}], \mu_N) \leqslant \left(\frac{1}{2\eta} + \| \log_\eta^K \|_\infty\right) N^{-\kappa}.$$

Both equations above conclude the proof. □

Lemma 2.2.3. *Let* $t_0(N) = 24\sqrt{2\pi}/(\eta N^{\frac{1}{2}+\kappa})$. *If* $t \geqslant t_0(N)$, *then*

$$\mathbb{P}((\mathcal{E}_{\mathrm{conc}})^c) \leqslant 12 \exp\left(-\frac{(t - t_0(N))^2 \eta^2 N^{1+2\kappa}}{288}\right).$$

Proof. The function $\log_\eta^K$ is not convex (it is convex on $[-\eta, \eta]$ and concave outside this interval). But it is a linear combination of three convex functions. Indeed, for $i = 1, 2, 3$, consider $\log_i = \log_{i,\eta,K} : \mathbb{R} \to \mathbb{R}$ given by

$$\log_1(x) = \begin{cases} -\frac{x}{2\eta} - \frac{1}{2} + \log_\eta(\eta) & \text{if } x \leqslant -\eta, \\ \log_\eta(x) & \text{if } -\eta \leqslant x \leqslant \eta, \\ \frac{x}{2\eta} - \frac{1}{2} + \log_\eta(\eta) & \text{if } x \geqslant \eta, \end{cases}$$

$$\log_2(x) = \begin{cases} \frac{x}{2\eta} & x \leqslant \eta, \\ \log_\eta^K(x) + \frac{1}{2} - \log_\eta(\eta) & \text{if } x \geqslant \eta, \end{cases}$$

$$\log_3(x) = \begin{cases} -\frac{x}{2\eta} & \text{if } x \geqslant -\eta, \\ \log_\eta^K(x) + \frac{1}{2} - \log_\eta(\eta) & \text{if } x \leqslant -\eta. \end{cases}$$

Notice that $\log_\eta^K = \sum_{i=1}^3 \log_i$, that $\log_1$ is convex while $\log_2$ and $\log_3$ are concave, and that each

$\log_i$ is $\frac{1}{2\eta}$-Lipschitz. For each i, consider the function $f_i : [-\frac{N^{-\kappa}}{\|\Phi\|_{\mathrm{Lip}}}, \frac{N^{-\kappa}}{\|\Phi\|_{\mathrm{Lip}}}]^M \to \mathbb{R}$ given by

$$f_i(X) = (-1)^{\mathbb{1}_{i\neq 1}} \frac{1}{N} \operatorname{tr}(\log_i(\Phi(X))) = (-1)^{\mathbb{1}_{i\neq 1}} \int_{\mathbb{R}} \log_i(\lambda) \hat{\mu}_{\Phi(X)}(\mathrm{d}\lambda).$$

The factors of -1 are for convenience, so that each f_i will be convex. Notice that

$$\mathbb{P}((\mathcal{E}_{\mathrm{conc}})^c) = \mathbb{P}\left(\left|\sum_{i=1}^{3} (-1)^{\mathbb{1}_{i\neq 1}} (f_i(X) - \mathbb{E}[f_i(X)])\right| > t\right) \leqslant \sum_{i=1}^{3} \mathbb{P}\left(|f_i(X) - \mathbb{E}[f_i(X)]| > \frac{t}{3}\right). \quad (2.2.3)$$

Each f_i is a Lipschitz, convex function of the many independent compactly supported variables $X_1, \dots, X_M$. Thus we can apply concentration-of-measure results of Talagrand. It will be useful to factor $f_i = g_i \circ \Phi$, where $g_i : \mathscr{S}_N \to \mathbb{R}$ is given by $g_i(T) = (-1)^{\mathbb{1}_{i\neq 1}} \frac{1}{N} \operatorname{tr}(\log_i(T))$.

Indeed, since $\log_i$ is $(2\eta)^{-1}$-Lipschitz, we know that g_i is $(\eta\sqrt{2N})^{-1}$-Lipschitz (see, e.g., [7, Lemma 2.3.1], and thus f_i is $\|\Phi\|_{\mathrm{Lip}}/(\eta\sqrt{2N})$-Lipschitz. Furthermore, since $(-1)^{\mathbb{1}_{i\neq 1}} \log_i$ is convex, by Klein's lemma (see, e.g., [103, Lemma 1.2]) g_i is also convex; since we assumed that Φ pulls back convex sets to convex sets, we conclude that $\{X : f_i(X) \leqslant a\}$ is a convex set of $[-\frac{N^{-\kappa}}{\|\Phi\|_{\mathrm{Lip}}}, \frac{N^{-\kappa}}{\|\Phi\|_{\mathrm{Lip}}}]^M$ for every $a \in \mathbb{R}$. Then [143, Theorem 6.6] implies that

$$\mathbb{P}(|f_i(X) - \mathfrak{M}_{f_i}| \geqslant t) \leqslant 4 \exp\left(-\frac{t^2\eta^2 N^{1+2\kappa}}{32}\right)$$

where $\mathfrak{M}_{f_i}$ is a median of $f_i(X)$. We conclude using (2.2.3) and the estimate

$$|\mathbb{E}(f_i(X)) - \mathfrak{M}_{f_i}| \leqslant \mathbb{E}|f_i(X) - \mathfrak{M}_{f_i}| \leqslant 4 \int_0^\infty \exp\left(-\frac{t^2\eta^2 N^{1+2\kappa}}{32}\right) \mathrm{d}t = \frac{8\sqrt{2\pi}}{\eta N^{\frac{1}{2}+\kappa}} = \frac{1}{3} t_0(N)$$

to substitute the median with the mean. □

2.2.1.2 Upper bound. After establishing one more estimate, we prove the upper bound of Theorem 2.1.1.

Lemma 2.2.4. *With the parameter choices (2.2.1), we have*

$$\lim_{N\to\infty} \frac{1}{N} \log \mathbb{E}[|\det(H_N)|(1 - \mathbb{1}_{\mathcal{E}_{\mathrm{ss}}} \mathbb{1}_{\mathcal{E}_{\mathrm{conc}}})] = -\infty.$$

Proof. Writing $\mathcal{E} = \mathcal{E}_{\mathrm{ss}} \cap \mathcal{E}_{\mathrm{conc}}$, for any $\delta > 0$ Hölder's inequality gives

$$\frac{1}{N} \log \mathbb{E}[|\det(H_N)|\mathbb{1}_{\mathcal{E}^c}] \leqslant \frac{1}{(1+\delta)N} \log \mathbb{E}[|\det(H_N)|^{1+\delta}] + \frac{\delta}{(1+\delta)N} \log \mathbb{P}(\mathcal{E}^c).$$

For δ satisfying (2.1.6), the first term is $\mathrm{O}(\log N)$. Concerning the second term, we have

$$\frac{1}{N} \log \mathbb{P}(\mathcal{E}^c) \leqslant \frac{1}{N} \log[\mathbb{P}((\mathcal{E}_{\mathrm{conc}})^c) + \mathbb{P}((\mathcal{E}_{\mathrm{ss}})^c)] \leqslant -C \log N,$$

for any $C > 0$ and $N \geqslant N_0(C)$, where the last inequality follows from Lemma 2.2.3, our parameter choices (2.2.1), and our assumption (2.1.7). □

Proof of upper bound. From our assumptions on μ_N we have $\liminf_{N\to\infty} \int \log|\lambda| \mu_N(\mathrm{d}\lambda) > -\infty$. Thus, by Lemma 2.2.4, it suffices to prove

$$\limsup_{N\to\infty} \left(\frac{1}{N} \log \mathbb{E}[|\det(H_N)|\mathbb{1}_{\mathcal{E}_{\mathrm{ss}}} \mathbb{1}_{\mathcal{E}_{\mathrm{conc}}}] - \int \log|\lambda| \mu_N(\mathrm{d}\lambda) \right) \leqslant 0. \tag{2.2.4}$$

On the events $\mathcal{E}_{\mathrm{ss}}$ and $\mathcal{E}_{\mathrm{conc}}$, Lemmas 2.2.1 and 2.2.2 give us

$$\begin{aligned}
&\int \log_\eta^K \mathrm{d}\hat{\mu}_{\Phi(X)} \\
&= \int \log_\eta^K \mathrm{d}(\hat{\mu}_{\Phi(X)} - \hat{\mu}_{\Phi(X_{\mathrm{cut}})}) + \int \log_\eta^K \mathrm{d}(\hat{\mu}_{\Phi(X_{\mathrm{cut}})} - \mathbb{E}[\hat{\mu}_{\Phi(X_{\mathrm{cut}})}]) + \int \log_\eta^K \mathbb{E}[\hat{\mu}_{\Phi(X_{\mathrm{cut}})}] \\
&\leqslant N^{-\kappa} \log\left(1 + \frac{K^2}{\eta^2}\right) + t + \int \log_\eta^K \mathbb{E}[\hat{\mu}_{\Phi(X_{\mathrm{cut}})}] \leqslant 2\varepsilon_1(N) + t + \int \log_\eta^K(\lambda) \mu_N(\mathrm{d}\lambda).
\end{aligned}$$

We use this estimate to obtain

$$\begin{aligned}\frac{1}{N}\log\mathbb{E}[|\det(H_N)|\mathbb{1}_{\mathcal{E}_{\mathrm{ss}}}\mathbb{1}_{\mathcal{E}_{\mathrm{conc}}}] &= \frac{1}{N}\log\mathbb{E}\left[\left(\prod_{i=1}^{N}|\lambda_i|\mathbb{1}_{|\lambda_i|\leqslant K}\right)\left(\prod_{i=1}^{N}|\lambda_i|\mathbb{1}_{|\lambda_i|>K}\right)\mathbb{1}_{\mathcal{E}_{\mathrm{ss}}}\mathbb{1}_{\mathcal{E}_{\mathrm{conc}}}\right]\\ &\leqslant \frac{1}{N}\log\mathbb{E}\left[e^{N\int\log_\eta^K \mathrm{d}\hat{\mu}_{\Phi(X)}}\left(\prod_{i=1}^{N}(1+|\lambda_i|\mathbb{1}_{|\lambda_i|>K})\right)\mathbb{1}_{\mathcal{E}_{\mathrm{ss}}}\mathbb{1}_{\mathcal{E}_{\mathrm{conc}}}\right]\\ &\leqslant 2\varepsilon_1(N)+t+\frac{1}{N}\log\mathbb{E}\left[\prod_{i=1}^{N}(1+|\lambda_i|\mathbb{1}_{|\lambda_i|>K})\right]+\int\log_\eta^K \mathrm{d}\mu_N.\end{aligned}$$

From our choice of parameters (2.2.1) and the assumption (2.1.4), this last term is $\int\log_\eta^K \mathrm{d}\mu_N+o(1)$. Furthermore, since the μ_N's are supported on a common compact set and K increases with N, we have $\int\log_\eta^K \mathrm{d}\mu_N=\int\log_\eta \mathrm{d}\mu_N$ for N large enough. Thus to prove (2.2.4) we need only show

$$\limsup_{N\to\infty}\int(\log_\eta(\lambda)-\log|\lambda|)\mu_N(\mathrm{d}\lambda)\leqslant 0. \tag{2.2.5}$$

To show this, we use

$$\int_\kappa^\infty(\log_\eta(\lambda)-\log|\lambda|)\mu_N(\mathrm{d}\lambda)\leqslant\frac{1}{2}\log\left(1+\frac{\eta^2}{\kappa^2}\right)$$

which tends to zero since η does, and

$$\left|\int_{-\kappa}^{\kappa}(\log_\eta(\lambda)-\log|\lambda|)\mu_N(\mathrm{d}\lambda)\right|\leqslant\kappa^{-1}\int_{-\kappa}^{\kappa}(\log|\lambda|-\log_\eta(\lambda))|\lambda|^{-1+\kappa}\,\mathrm{d}\lambda,$$

which tends to zero by dominated convergence. This completes the proof of (2.2.5) and thus of the upper bound. □

2.2.1.3 Lower bound. We first collect some estimates.

Lemma 2.2.5. *We have*

$$\frac{1}{N}\log\mathbb{E}[e^{N\int(\log|\lambda|-\log_\eta(\lambda))\hat{\mu}_{\Phi(X)}(\mathrm{d}\lambda)}\mathbb{1}_{\mathcal{E}_{\mathrm{gap}}}\mathbb{1}_{\mathcal{E}_{\mathrm{ss}}}\mathbb{1}_{\mathcal{E}_{\mathrm{conc}}}]\geqslant-\varepsilon_2(N),$$

where

$$\varepsilon_2(N) = \frac{p_b}{2}\log(1+e^{2N^\varepsilon}\eta^2) + \frac{\eta^2}{2w_b^2} - \frac{1}{N}\log \mathbb{P}(\mathcal{E}_{\mathrm{gap}}, \mathcal{E}_{\mathrm{ss}}, \mathcal{E}_{\mathrm{conc}}, \mathcal{E}_b).$$

Proof. On $\mathcal{E}_{\mathrm{gap}}$, for any eigenvalue λ of $\Phi(X)$ we have

$$\log|\lambda| - \log_\eta(\lambda) = -\frac{1}{2}\log\left(1+\frac{\eta^2}{\lambda^2}\right) \geqslant -\frac{1}{2}\log(1+e^{2N^\varepsilon}\eta^2).$$

Similarly, since $1-b(\lambda) \leqslant \mathbb{1}_{|\lambda|\geqslant w_b}$ and $\log(1+x) \leqslant x$ for $x>0$, we have

$$\int (\log|\lambda| - \log_\eta(\lambda))(1-b(\lambda))\hat{\mu}_{\Phi(X)}(\mathrm{d}\lambda) \geqslant -\frac{1}{2}\log\left(1+\frac{\eta^2}{w_b^2}\right) \geqslant -\frac{\eta^2}{2w_b^2}.$$

Thus

$$\begin{aligned}
&\mathbb{E}[e^{N\int(\log|\cdot|-\log_\eta)\mathrm{d}\hat{\mu}_{\Phi(X)}}\mathbb{1}_{\mathcal{E}_{\mathrm{gap}}}\mathbb{1}_{\mathcal{E}_{\mathrm{ss}}}\mathbb{1}_{\mathcal{E}_{\mathrm{conc}}}\mathbb{1}_{\mathcal{E}_b}]\\
&\geqslant e^{-\frac{Np_b}{2}\log(1+e^{2N^\varepsilon}\eta^2)}\mathbb{E}[e^{N\int(\log|\cdot|-\log_\eta)(1-b)\mathrm{d}\hat{\mu}_{\Phi(X)}}\mathbb{1}_{\mathcal{E}_{\mathrm{gap}}}\mathbb{1}_{\mathcal{E}_{\mathrm{ss}}}\mathbb{1}_{\mathcal{E}_{\mathrm{conc}}}\mathbb{1}_{\mathcal{E}_b}]\\
&\geqslant e^{-\frac{Np_b}{2}\log(1+e^{2N^\varepsilon}\eta^2)}e^{-\frac{N\eta^2}{2w_b^2}}\mathbb{P}(\mathcal{E}_{\mathrm{gap}}, \mathcal{E}_{\mathrm{ss}}, \mathcal{E}_{\mathrm{conc}}, \mathcal{E}_b),
\end{aligned}$$

which concludes the proof. □

Lemma 2.2.6. *For N large enough we have*

$$\mathbb{P}((\mathcal{E}_b)^c) \leqslant \frac{2}{p_b}\left(\frac{N^{-\kappa}}{w_b} + \frac{(2w_b)^\kappa}{\kappa^2}\right).$$

Proof. By our choice (2.2.1) of w_b tending to zero, μ_N admits a density on $[-2w_b, 2w_b]$ for N large enough. Since $b(\lambda)$ is $\frac{1}{w_b}$-Lipschitz and bounded above by $\mathbb{1}_{|\lambda|\leqslant 2w_b}$, we use (2.1.3) to find

$$\mathbb{E}\left[\int b\ \mathrm{d}\hat{\mu}_{\Phi(X)}\right] \leqslant \left(\frac{1}{w_b}+1\right)d_{\mathrm{BL}}(\mathbb{E}[\hat{\mu}_{\Phi(X)}], \mu_N) + \mu_N([-2w_b, 2w_b]) \leqslant \frac{2N^{-\kappa}}{w_b} + \frac{1}{\kappa}\int_{-2w_b}^{2w_b}|x|^{-1+\kappa}\,\mathrm{d}x.$$

The conclusion follows by evaluating this integral and applying Markov's inequality. □

Proof of lower bound. Lemmas 2.2.1, 2.2.2, and 2.2.5 show that $N^{-1}\log\mathbb{E}[|\det(H_N)|]$ is larger than

$$\begin{aligned}
&\frac{1}{N}\log\mathbb{E}\Big[e^{N\left(\int(\log|\cdot|-\log_\eta)\mathrm{d}\hat{\mu}_{\Phi(X)}+\int\log_\eta^K\mathrm{d}(\hat{\mu}_{\Phi(X)}-\hat{\mu}_{\Phi(X_{\mathrm{cut}})}+\hat{\mu}_{\Phi(X_{\mathrm{cut}})}-\mathbb{E}[\hat{\mu}_{\Phi(X_{\mathrm{cut}})}])\right)}\mathbb{1}_{\mathcal{E}_{\mathrm{gap}}}\mathbb{1}_{\mathcal{E}_{\mathrm{ss}}}\mathbb{1}_{\mathcal{E}_{\mathrm{conc}}}\Big]\\
&+\int\log_\eta^K\mathrm{d}\mathbb{E}[\hat{\mu}_{\Phi(X_{\mathrm{cut}})}]\\
\geqslant\ &\frac{1}{N}\log\mathbb{E}[e^{N\int(\log|\cdot|-\log_\eta)\mathrm{d}\hat{\mu}_{\Phi(X)}}\mathbb{1}_{\mathcal{E}_{\mathrm{gap}}}\mathbb{1}_{\mathcal{E}_{\mathrm{ss}}}\mathbb{1}_{\mathcal{E}_{\mathrm{conc}}}]-N^{-\kappa}\log\left(1+\frac{K^2}{\eta^2}\right)-t+\int\log_\eta^K\mathrm{d}\mathbb{E}[\hat{\mu}_{\Phi(X_{\mathrm{cut}})}]\\
\geqslant\ &\int\log|\cdot|\mathrm{d}\mu_N-\varepsilon(N),
\end{aligned}\tag{2.2.6}$$

where $\varepsilon(N)=\varepsilon_1(N)+\varepsilon_2(N)+N^{-\kappa}\log\left(1+\frac{K^2}{\eta^2}\right)+t$ and we have used

$$\int\log_\eta^K(\lambda)\mu_N(\mathrm{d}\lambda)\geqslant\int\log(\min(|\lambda|,K))\mu_N(\mathrm{d}\lambda)=\int\log|\lambda|\mu_N(\mathrm{d}\lambda)\tag{2.2.7}$$

for N large enough in the last inequality (2.2.6), as the μ_N's are supported on a common compact set and K grows with N. It remains to check that $\varepsilon(N)\to 0$. This follows immediately from our parameter choices (2.2.1), except possibly for the term $\varepsilon_2(N)$. For this term, we note that $\mathbb{P}(\mathcal{E}_{\mathrm{ss}})\to 1$ and $\mathbb{P}(\mathcal{E}_{\mathrm{gap}})\to 1$ by assumption ((2.1.7) and (2.1.5), respectively), then use Lemmas 2.2.3 and 2.2.6 to show that $\mathbb{P}(\mathcal{E}_{\mathrm{conc}})\to 1$ and $\mathbb{P}(\mathcal{E}_b)\to 1$. This shows that $\varepsilon_2(N)\to 0$, which concludes the proof of the lower bound and thus of (2.1.8). □

2.2.2 *Proof of Theorem 2.1.2.*

In this subsection we prove Theorem 2.1.2. The proof is largely similar to that of Theorem 2.1.1, so we will omit some steps.

We make the same parameter choices as in (2.2.1). We also work with the events $\mathcal{E}_{\mathrm{gap}}$ and $\mathcal{E}_b$ from (2.2.2), but $\mathcal{E}_{\mathrm{ss}}$ is no longer relevant, and $\mathcal{E}_{\mathrm{conc}}$ is replaced by

$$\mathcal{E}_{\mathrm{Lip}}=\left\{\left|\int\log_\eta(\lambda)(\hat{\mu}_{H_N}-\mathbb{E}[\hat{\mu}_{H_N}])(\mathrm{d}\lambda)\right|\leqslant t\right\}.$$

Proof of upper bound of Theorem 2.1.2. From (2.1.10) and some elementary estimates, there exists

a universal constant c_{ε_0} such that, for N large enough, we have

$$\mathbb{E}[e^{N\int \log_\eta(\lambda)(\hat{\mu}_{H_N}-\mathbb{E}[\hat{\mu}_{H_N}])(\mathrm{d}\lambda)}] \leqslant c_{\varepsilon_0}\exp\left[\left(\frac{2N^\zeta}{c_\zeta}\right)^{1/\varepsilon_0}\left(\frac{1}{2\eta}\right)^2\right].$$

Hence

$$\begin{aligned}\frac{1}{N}\log\mathbb{E}[|\det(H_N)|] &\leqslant \frac{1}{N}\log\mathbb{E}[e^{N\int\log_\eta(\lambda)\hat{\mu}_{H_N}(\mathrm{d}\lambda)}] \leqslant \left(\frac{2}{4^{\varepsilon_0}c_\zeta}\right)^{1/\varepsilon_0}\frac{N^{\zeta/\varepsilon_0-1}}{\eta^2}+\int\log_\eta(\lambda)\mathbb{E}[\hat{\mu}_{H_N}](\mathrm{d}\lambda)\\ &\leqslant \left(\frac{2}{4^{\varepsilon_0}c_\zeta}\right)^{1/\varepsilon_0}\frac{N^{\zeta/\varepsilon_0-1}}{\eta^2}+\frac{1}{2\eta}\mathrm{W}_1(\mathbb{E}[\hat{\mu}_{H_N}],\mu_N)+\int\log_\eta(\lambda)\mu_N(\mathrm{d}\lambda).\end{aligned}$$

For ζ small enough, the first term decays with N. We complete the proof by applying (2.1.9) and (2.2.5). □

Proof of lower bound of Theorem (2.1.2). Arguing as in (2.2.6), $N^{-1}\log\mathbb{E}[|\det(H_N)|]$ is larger than

$$\begin{aligned}&\frac{1}{N}\log\mathbb{E}[e^{N\left(\int(\log|\cdot|-\log_\eta)\mathrm{d}\hat{\mu}_{H_N}+\int\log_\eta\mathrm{d}(\hat{\mu}_{H_N}-\mathbb{E}[\hat{\mu}_{H_N}])\right)}\mathbb{1}_{\mathcal{E}_{\mathrm{Lip}}}\mathbb{1}_{\mathcal{E}_{\mathrm{gap}}}]+\int\log_\eta\mathrm{d}\mathbb{E}[\hat{\mu}_{H_N}]\\ &\geqslant \frac{1}{N}\log\mathbb{E}\Big[e^{N\int(\log|\cdot|-\log_\eta)\hat{\mu}_{H_N}}\mathbb{1}_{\mathcal{E}_{\mathrm{Lip}}}\mathbb{1}_{\mathcal{E}_{\mathrm{gap}}}\Big]-t-\frac{1}{2\eta}\mathrm{W}_1(\mathbb{E}[\hat{\mu}_{H_N}],\mu_N)+\int\log|\cdot|\mu_N.\end{aligned}$$

As in Lemma 2.2.5, we have

$$\begin{aligned}&\frac{1}{N}\log\mathbb{E}\Big[e^{N\int(\log|\lambda|-\log_\eta(\lambda))\hat{\mu}_{H_N}(\mathrm{d}\lambda)}\mathbb{1}_{\mathcal{E}_{\mathrm{Lip}}}\mathbb{1}_{\mathcal{E}_{\mathrm{gap}}}\mathbb{1}_{\mathcal{E}_b}\Big]\\ &\geqslant -\frac{p_b}{2}\log(1+e^{2N^\varepsilon}\eta^2)-\frac{\eta^2}{2w_b^2}-\frac{1}{N}\log\mathbb{P}(\mathcal{E}_{\mathrm{Lip}},\mathcal{E}_{\mathrm{gap}},\mathcal{E}_b),\end{aligned}$$

so by our parameter choices (2.2.1) it suffices to show $\mathbb{P}(\mathcal{E}_{\mathrm{Lip}},\mathcal{E}_{\mathrm{gap}},\mathcal{E}_b)\to 1$. The event $\mathcal{E}_{\mathrm{gap}}$ is handled by assumption (2.1.5); the event $\mathcal{E}_b$ is handled by Lemma 2.2.6 (replacing d_{BL} there with W_1 here); and the event $\mathcal{E}_{\mathrm{Lip}}$ is handled by assumption (L), since (2.1.10) gives $\mathbb{P}(\mathcal{E}_{\mathrm{Lip}}^c)\leqslant \exp\Big(-\frac{c_\zeta}{N^\zeta}\min\{(2Nt\eta)^2,(2Nt\eta)^{1+\varepsilon_0}\}\Big)$. □

2.3 Applications to matrix models

In this section, we check the assumptions of our general theorems, 2.1.1 and 2.1.2, for our different matrix models. First we present two general and classical techniques that will help us check these assumptions. Informally speaking, the first technique shows how local laws for the Stieltjes transform along lines of the form $\{E + \mathrm{i}N^{-\varepsilon} : E \in [-C, C]\}$ give polynomial convergence rates of the averaged empirical spectral measure, corresponding to assumptions (E) and (W). The second technique proves Wegner estimates of the form (2.1.5) using the Schur complement formula.

In the last Section 2.3.10, we prove the claims made just after Theorem 2.1.1 about the necessity of its assumptions.

2.3.1 General technique: Convergence rates via local laws. In this subsection, we summarize the general technique for using local laws to derive estimates like (2.1.3) and (2.1.9). We will use this technique repeatedly for specific matrix models. This idea is classical; see for instance [19] for the specific estimates we need.

Write $s_N(z) = \int \mathrm{d}\hat{\mu}_N(\lambda)/(\lambda - z)$ the Stieltjes transform of $\hat{\mu}_{H_N}$, and $m_N(z) = \int \mathrm{d}\mu_N(\lambda)/(\lambda - z)$ the Stieltjes transform of μ_N. Define the distribution functions $F_{\mathbb{E}\hat{\mu}}(x) = \mathbb{E}[\hat{\mu}_{H_N}]((-\infty, x])$, $F_{\mu_N}(x) = \mu_N((-\infty, x])$.

Proposition 2.3.1. *Suppose the measures μ_N have densities $\mu_N(\cdot)$ on all of $\mathbb{R}$, not just near the origin, and $\sup_N \|\mu_N(\cdot)\|_{L^\infty} < \infty$. Assume also that there exist fixed (N-independent) constants $A, \varepsilon_1, \varepsilon_2 > 0$ such that*

$$\int_{-3A}^{3A} \left|\mathbb{E}[s_N(E + \mathrm{i}N^{-\varepsilon_1})] - m_N(E + \mathrm{i}N^{-\varepsilon_1})\right| \leqslant N^{-\varepsilon_2}, \tag{2.3.1}$$

$$\int_{|x|>A} \left|F_{\mathbb{E}[\hat{\mu}]}(x) - F_{\mu_N}(x)\right| \mathrm{d}x \leqslant N^{-\varepsilon_1-\varepsilon_2}. \tag{2.3.2}$$

Then there exists $\gamma > 0$ with $d_{\mathrm{KS}}(\mathbb{E}[\hat{\mu}_{H_N}], \mu_N) = \mathrm{O}(N^{-\gamma})$. If in addition $\mathrm{supp}(\mu_N) \subset (-A, A)$ for

each N, and

$$\left|F_{\mathbb{E}[\hat{\mu}]}(x) - F_{\mu_N}(x)\right| = o_{|x|\to\infty}\left(\frac{1}{|x|}\right), \tag{2.3.3}$$

then there exists $\gamma' > 0$ with $d_{\mathrm{BL}}(\mathbb{E}[\hat{\mu}_{H_N}], \mu_N) \leqslant \mathrm{W}_1(\mathbb{E}[\hat{\mu}_{H_N}], \mu_N) = \mathrm{O}(N^{-\gamma'})$.

Proof. From [19, Theorem 2.2], we have

$$\begin{aligned} d_{\mathrm{KS}}(\mathbb{E}[\hat{\mu}_{H_N}], \mu_N) \leqslant \eta^{-1} \sup_x \int_{|y|\leqslant 10\eta} |F_{\mu_N}(x+y) - F_{\mu_N}(x)|\,\mathrm{d}y \\ + \frac{2\pi}{\eta}\int_{|x|>A}\left|F_{\mathbb{E}[\hat{\mu}]}(x) - F_{\mu_N}(x)\right|\mathrm{d}x + \int_{-3A}^{3A} |\mathbb{E}[s_N(E+\mathrm{i}\eta)] - m_N(E+\mathrm{i}\eta)|\,\mathrm{d}E. \end{aligned}$$

Since the measures μ_N have densities bounded by S, say, the function F_{μ_N} is S-Lipschitz; hence the first term is at most $100S\eta$. With the choice $\eta = N^{-\varepsilon_1}$, the second and third terms are handled by assumption.

For the Wasserstein distance, let f be a test function with $\|f\|_{\mathrm{Lip}} \leqslant 1$. We integrate by parts (notice (2.3.3) gives us the decay at infinity necessary to do this) to find

$$\begin{aligned} \left|\int_{2A}^{\infty} f(x)(\mathbb{E}[\hat{\mu}_{H_N}] - \mu_N)(\mathrm{d}x)\right| = \left|\int_{2A}^{\infty} f(x)\mathbb{E}[\hat{\mu}_{H_N}](\mathrm{d}x)\right| \leqslant \int_{2A}^{\infty}(x - (2A-1))\mathbb{E}[\hat{\mu}_{H_N}](\mathrm{d}x) \\ \leqslant d_{\mathrm{KS}}(\mathbb{E}[\hat{\mu}_{H_N}], \mu_N) + \int_{2A}^{\infty}\left|F_{\mathbb{E}[\hat{\mu}]}(x) - F_{\mu_N}(x)\right|\mathrm{d}x \leqslant N^{-\gamma} + N^{-\varepsilon_1-\varepsilon_2} \end{aligned}$$

and similarly for the left tail. For the bulk, we approximate f on $[-2A, 2A]$ with test functions smooth enough to integrate by parts on f directly, which gives

$$\left|\int_{-2A}^{2A} f(x)(\mathbb{E}[\hat{\mu}_{H_N}] - \mu_N)(\mathrm{d}x)\right| \leqslant (8A+4)d_{\mathrm{KS}}(\mathbb{E}[\hat{\mu}_{H_N}], \mu_N).$$

This completes the proof. □

2.3.2 General technique: Wegner estimates via Schur complements. In this subsection, we summarize the classical idea of using the Schur-complement formula to derive Wegner estimates on the probability that there are no eigenvalues in a small gap around energy level E. These will

be used to check (2.1.5) for a wide variety of models.

For compactness, we temporarily drop the N-dependence from the notation H_N. For any j in $[\![1, N]\!]$, write $H^{(j)}$ for the matrix obtained by erasing the jth column and row from H, write h_j for the $(N-1)$-vector consisting of the jth column of H with the entry H_{jj} removed, and write $H_{\widehat{jj}}$ for the collection of every entry of H except for H_{jj}.

Proposition 2.3.2. *Fix $E \in \mathbb{R}$ and suppose there exists a sequence $\eta = \eta_N$ tending to zero such that*

$$\sup_{j \in [\![1,N]\!]} \mathbb{E}\left[\mathbb{E}\left[\operatorname{Im}\left(\frac{1}{H_{jj} - (E + \mathrm{i}\eta + h_j^T (H^{(j)} - (E+\mathrm{i}\eta))^{-1} h_j)}\right) \middle| H_{\widehat{jj}}\right]\right] = o\left(\frac{1}{N\eta}\right). \tag{2.3.4}$$

Then

$$\lim_{N\to\infty} \mathbb{P}(H_N \text{ has no eigenvalues in } [E-\eta, E+\eta]) = 1.$$

Proof. We have

$$\begin{aligned}
&\mathbb{P}(H_N \text{ has an eigenvalue in } [E-\eta, E+\eta]) \leqslant \mathbb{E}[\#\{j : |\lambda_j - E| \leqslant \eta\}] \leqslant \mathbb{E}\left[2\sum_{j=1}^N \frac{\eta^2}{\eta^2 + (\lambda_j - E)^2}\right] \\
&= 2\eta \mathbb{E}\left[\operatorname{Im}\left(\sum_{j=1}^N \frac{1}{\lambda_j - E - \mathrm{i}\eta}\right)\right] = 2\eta \mathbb{E}\left[\operatorname{Im}\left(\operatorname{Tr} \frac{1}{H - (E+\mathrm{i}\eta)}\right)\right] \\
&\leqslant 2N\eta \sup_{j \in [\![1,N]\!]} \mathbb{E}[\operatorname{Im}(((H - (E+\mathrm{i}\eta))^{-1})_{jj})].
\end{aligned}$$

Moreover, the Schur complement formula gives

$$((H - (E+\mathrm{i}\eta))^{-1})_{jj} = \frac{1}{H_{jj} - (E + \mathrm{i}\eta + h_j^T (H^{(j)} - (E+\mathrm{i}\eta))^{-1} h_j)},$$

which concludes the proof by the assumption (2.3.4). □

Lemma 2.3.3. *Write $\widetilde{H}_{jj}$ for the law of H_{jj} conditioned on $H_{\widehat{jj}}$. Suppose that there exists a single probability measure μ on $\mathbb{R}$ (independent of N and j) with a bounded density $\mu(\cdot)$, and constants*

$\widetilde{\sigma}_{jj} = \widetilde{\sigma}_{jj}^{(N)}$ *and* $\widetilde{m}_{jj} = \widetilde{m}_{jj}^{(N)}$ *such that*

$$\frac{\widetilde{H}_{jj} - \widetilde{m}_{jj}}{\widetilde{\sigma}_{jj}} \sim \mu$$

for every N *and* $j \in [\![1, N]\!]$. *If there exist* $\alpha, C > 0$ *with*

$$\inf_{j \in [\![1,N]\!]} \widetilde{\sigma}_{jj} \geqslant \frac{1}{C} N^{-\alpha},$$

then (2.3.4) *holds with* $\eta = o(N^{-1-\alpha})$ *for every* $E \in \mathbb{R}$.

Proof. For any *deterministic* $z = E + \mathrm{i}\eta$, and with the notation $S := \|\mu(\cdot)\|_{L^\infty}$, we have

$$\mathbb{E}_{\widetilde{H_{jj}}}\left[\mathrm{Im}\left(\frac{1}{\widetilde{H_{jj}} - z}\right)\right] = \int_{\mathbb{R}} \frac{\eta}{(\widetilde{\sigma}_{jj}x + \widetilde{m}_{jj} - E)^2 + \eta^2} \mu(x)\,\mathrm{d}x \leqslant S\frac{1}{\widetilde{\sigma}_{jj}} \int_{\mathbb{R}} \frac{\eta}{x^2 + \eta^2}\,\mathrm{d}x \leqslant \pi S C N^{\alpha}.$$

Define $z_j = E + \mathrm{i}\eta + h_j^T (H^{(j)} - (E + \mathrm{i}\eta))^{-1} h_j$, and $\widetilde{z_j} = z_j - \mathbb{E}[\widetilde{H}_{jj}]$, and notice that $\widetilde{z_j}$ is measurable with respect to $H_{\widehat{jj}}$ with $\mathrm{Im}(\widetilde{z_j}) \geqslant \eta$ deterministically; thus

$$\sup_{j \in [\![1,N]\!]} \mathbb{E}\left[\mathbb{E}\left[\mathrm{Im}\left(\frac{1}{H_{jj} - z_j}\right)\middle| H_{\widehat{jj}}\right]\right] = \sup_{j \in [\![1,N]\!]} \mathbb{E}_{\widetilde{z_j}}\left[\mathbb{E}_{\widetilde{H_{jj}}}\left[\mathrm{Im}\left(\frac{1}{\widetilde{H_{jj}} - \widetilde{z_j}}\right)\right]\right] \leqslant \pi S C N^{\alpha}$$

which is $o(1/(N\eta))$ for our choice of η. □

2.3.3 Wigner matrices. We will use Theorem 2.1.1 (convexity-preserving functional) and model a Wigner matrix $W_N - E$ as $W_N - E = \Phi(X_1, \ldots, X_M)$, where $M = \frac{N(N+1)}{2}$, the X_i's are independent random variables distributed according to μ, and Φ is $\frac{1}{\sqrt{N}}$ times the identity map which places these entries in the upper triangle of an $N \times N$ matrix, minus $E\,\mathrm{Id}$. This Φ is trivially convex and satisfies $\|\Phi\|_{\mathrm{Lip}} = \frac{1}{\sqrt{N}}$.

Now we check Assumption (E) on expectations, with all μ_N's equal to the semicircle law ρ_{sc}. A. Tikhomirov [147, Theorem 1.1] showed that for every ε in the assumption of $2 + \varepsilon$ moments,

there exists $\eta = \eta(\varepsilon) > 0$ with

$$d_{\mathrm{KS}}(\mathbb{E}[\hat{\mu}_{W_N}], \rho_{\mathrm{sc}}) \leqslant N^{-\eta}. \tag{2.3.5}$$

Now we transfer this inequality from d_{KS} to d_{BL}: If $M > 2$ and $\|f\|_\infty \leqslant 1$, then

$$\left| \int_{-\infty}^{-M} f(x)(\mathbb{E}[\hat{\mu}_{W_N}] - \rho_{\mathrm{sc}})(\mathrm{d}x) \right| = \left| \int_{-\infty}^{-M} f(x)\mathbb{E}[\hat{\mu}_{W_N}](\mathrm{d}x) \right| \leqslant \int_{-\infty}^{-M} \mathbb{E}[\hat{\mu}_{W_N}](\mathrm{d}x) \leqslant N^{-\eta}$$

from (2.3.5), and similarly for $\int_M^\infty$; on $[-M, M]$ we proceed exactly as in the proof of Proposition 2.3.1, to obtain (E)[1].

Now we check the three estimates comprising assumption (C) on coarse bounds.

(2.1.4) Fix $\varepsilon > 0$ and write $W = W_N = A + B = A_N + B_N$, where A is defined entrywise by $A_{ij} = (W_{ij})\mathbb{1}_{|W_{ij}| \leqslant \frac{1}{10N} e^{N^\varepsilon}}$. Notice that all eigenvalues of A have absolute value at most $\frac{1}{10} e^{N^\varepsilon}$. The Weyl inequalities give us

$$\lambda_i(W) = \lambda_i(A + B) \leqslant \lambda_{\max}(A) + \lambda_i(B) \leqslant \frac{1}{10} e^{N^\varepsilon} + \lambda_i(B)$$

and similarly $\lambda_i(W) \geqslant \lambda_i(B) - \frac{1}{10} e^{N^\varepsilon}$, so that for fixed E, for large enough N we have, for any i,

$$\begin{aligned} &1 + |\lambda_i(W - E)|\mathbb{1}_{|\lambda_i(W-E)| > e^{N^\varepsilon}} \leqslant 1 + 2|\lambda_i(W)|\mathbb{1}_{|\lambda_i(W)| > \frac{1}{2} e^{N^\varepsilon}} \leqslant 1 + 2|\lambda_i(W)|\mathbb{1}_{|\lambda_i(B)| > \frac{1}{4} e^{N^\varepsilon}} \\ &\leqslant 1 + (|\lambda_{\max}(A)| + |\lambda_i(B)|)\mathbb{1}_{|\lambda_i(B)| > \frac{1}{4} e^{N^\varepsilon}} \leqslant 1 + 2|\lambda_i(B)|\mathbb{1}_{|\lambda_i(B)| > \frac{1}{4} e^{N^\varepsilon}}. \end{aligned}$$

For $x > 1$ we have $(1 + 2x) < (1 + 100x^2)^{1/2}$, so

$$\prod_{i=1}^{N} (1 + 2|\lambda_i(B)|\mathbb{1}_{|\lambda_i(B)| > \frac{1}{4} e^{N^\varepsilon}}) \leqslant \prod_{i=1}^{N} (1 + 100\lambda_i(B)^2)^{1/2} = \det(\mathrm{Id} + 100B^2)^{1/2}.$$

[1] We also briefly sketch another possible proof of Assumption (E). First, by following the usual Hoffman-Wielandt-based proof that two moments suffice for the Wigner semicircle law (see, e.g., [7, Theorem 2.1.21]), we can assume that the entries W_{ij} are replaced with $W_{ij}\mathbb{1}_{|W_{ij}| \leqslant N^{10\varepsilon}}$, if the $2 + \varepsilon$ moment is finite. Second, for this new matrix, one can apply the usual Stieltjes-transform-based proof of the Wigner semicircle law using Schur complements (see, e.g., [7, Section 2.4.2]); the fourth moments of the new matrix are $\mathrm{O}(N^{40\varepsilon})$, which is more than compensated by $1/N$ prefactors in the error terms.

By Fischer's inequality this can be bounded above by the product of its diagonal entries; that is,

$$\det(\mathrm{Id}+100B^2)^{1/2} \leqslant \prod_{i=1}^{N}\left(1+100\sum_{j=1}^{N}B_{ij}^2\right)^{1/2} \leqslant \prod_{i=1}^{N}\left(1+10\sum_{j=1}^{N}|B_{ij}|\right),$$

where for the last inequality we used $\sum a_i^2 \leqslant (\sum a_i)^2$ for positive numbers a_i. Now, for some constant C we have $\mathbb{E}[|B_{ij}|], \mathbb{E}[|B_{ij}|^2] \leqslant CNe^{-N^\varepsilon} \leqslant e^{-\frac{1}{2}N^\varepsilon}$, and notice that we can calculate $\mathbb{E}\Big[\prod_{i=1}^{N}\Big(1+10\sum_{j=1}^{N}|B_{ij}|\Big)\Big]$ by expansion and factorization again. All matrix elements appear with a power at most two, and for any set I of couples (i,j) which can appear in the expansion, we have $\mathbb{E}[\prod_{\alpha\in I}|B_\alpha|] \leqslant (e^{-\frac{1}{2}N^\varepsilon})^{|I|}$ so that

$$\frac{1}{N}\log\mathbb{E}\left[\prod_{i=1}^{N}\left(1+10\sum_{j=1}^{N}|B_{ij}|\right)\right] \leqslant \frac{1}{N}\log\prod_{i=1}^{N}\left(1+10\sum_{j=1}^{N}e^{-\frac{1}{2}N^\varepsilon}\right) \to 0.$$

(2.1.5) The existence of gaps near zero with high probability (indeed, gaps of polynomial size) was established by Nguyen [125, Theorem 1.4], including the case of general energy levels E.

(2.1.6) Fix δ so small that μ has finite $2+2\delta$ moment. Let S_N be the symmetric group on N letters, and for any permutation $\sigma \in S_N$ define $X_\sigma = \big|(W-E)_{1,\sigma(1)}\cdot\ldots\cdot(W-E)_{N,\sigma(N)}\big|$. Then $|\det(W_N-E)| \leqslant \sum_\sigma X_\sigma$, and by convexity of $x\mapsto x^{1+\delta}$ we have

$$|\det(W_N-E)|^{1+\delta} \leqslant \left(\sum_{\sigma\in S_N}X_\sigma\right)^{1+\delta} \leqslant (N!)^{1+\delta}\frac{\sum_\sigma X_\sigma^{1+\delta}}{N!}.$$

If $\sqrt{N}Y$ is distributed according to μ, then for each $E\in\mathbb{R}$ there exists $c_E = c_E(\mu,\delta)$ such that

$$\max(\mathbb{E}[|Y-E|^{1+\delta}], \mathbb{E}[|Y-E|^{2+2\delta}], \mathbb{E}[|Y|^{1+\delta}], \mathbb{E}[|Y|^{2+2\delta}]) \leqslant c_E < \infty.$$

Thus $\sup_\sigma \mathbb{E}[X_\sigma^{1+\delta}] \leqslant (c_E)^N$. Since $N! \leqslant N^N$, this gives $\mathbb{E}[|\det W_N|^{1+\varepsilon}] \leqslant c_E^N N^{(1+\delta)N}$ up to factors of lower order, which suffices.

To prove assumption (S) on spectral stability, we follow Bordenave, Caputo and Chafaï, see [55,

Lemma C.2] and [54, Lemma 2.2]. Write $W_N^{\mathrm{cut}} = \Phi(X_{\mathrm{cut}})$ for the matrix W_N with entries truncated at level $N^{-\kappa}$, and we will prove (S) for fixed, small enough κ with respect to ε. From interlacing (see, e.g., [18, Theorem A.43]) that

$$d_{\mathrm{KS}}(\hat{\mu}_{W_N}, \hat{\mu}_{W_N^{\mathrm{cut}}}) \leqslant \frac{1}{N} \operatorname{rank}(W_N - W_N^{\mathrm{cut}}) \leqslant \frac{2}{N} \sum_{i \leqslant j} \mathbb{1}_{|W_{ij}| > N^{-\kappa}},$$

where the last inequality follows since the rank of a matrix is at most the number of its nonzero entries. The $\frac{N(N+1)}{2}$ random variables $(\mathbb{1}_{|W_{ij}|>N^{-\kappa}})_{1\leqslant i\leqslant j\leqslant N}$ are i.i.d. Bernoulli variables with parameter

$$p_N = \mathbb{P}(|W_{ij}| > N^{-\kappa}) \leqslant cN^{(-\frac{1}{2}+\kappa)(2+\varepsilon)} \leqslant cN^{-1-\frac{\varepsilon}{4}},$$

if we choose κ small enough, with $c = \int |x|^{2+\varepsilon} \mu(\mathrm{d}x)$. Writing $h(x) = (x+1)\log(x+1) - x$, Bennett's inequality [47] gives

$$\mathbb{P}\left(\sum_{i \leqslant j} \mathbb{1}_{|W_{ij}|>N^{-\kappa}} - \frac{N(N+1)}{2} p_N \geqslant t \right) \leqslant \exp\left(-\sigma^2 h\left(\frac{t}{\sigma^2} \right) \right)$$

with

$$\sigma^2 = \frac{N(N+1)}{2} p_N (1 - p_N) \leqslant \frac{N(N+1)}{2} p_N \leqslant N^{1-\varepsilon/8}$$

for N large enough. With the choice $t = N^{1-\kappa} - \frac{N(N+1)}{2} p_N \geqslant \frac{1}{2} N^{1-\kappa}$ (for κ small enough) we have $\frac{t}{\sigma^2} \to +\infty$, and using $h(x) \sim x \log x$ as $x \to +\infty$ we obtain

$$\log \mathbb{P}(d_{\mathrm{KS}}(\hat{\mu}_{W_N}, \hat{\mu}_{W_N^{\mathrm{cut}}}) > N^{-\kappa}) \leqslant -\sigma^2 h\left(\frac{N^{1-\kappa}}{2\sigma^2} \right) \leqslant -CN^{1-\kappa} \log\left(\frac{N^{1-\kappa}}{2\sigma^2} \right)$$

for some constant C and N large enough, which completes the proof of (2.1.7).

Finally, we prove the remark just before Corollary 2.1.3, namely that $\mathbb{E}[|\det(W_N)|] = +\infty$ when $\mathbb{E}(|W_{12}|^2) = +\infty$. Indeed, we can write $\mathbb{E}[|\det(W_N)|] = \mathbb{E}(|XW_{12}^2 + YW_{12} + Z|)$ where the random vector (X, Y, Z) is independent of W_{12} and $X = \det((W_{ij})_{3\leqslant i,i\leqslant N})$. We have $\mathbb{P}(X = 0) < 1$ so that there exist compact intervals I, J, K in some $[-A, A]$ with $a := \inf_{x\in I} |x| > 0$ and $\mathbb{P}((X, Y, Z) \in$

$I \times J \times K) > 0$. For any such $(X, Y, Z) \in I \times J \times K$ we have $|XW_{12}^2 + YW_{12} + Z| \geqslant aW_{12}^2 - AW_{12} - A$, so that

$$\begin{aligned}\mathbb{E}(|XW_{12}^2 + YW_{12} + Z|) &\geqslant \mathbb{E}(|XW_{12}^2 + YW_{12} + Z|\mathbb{1}_{(X,Y,Z)\in I\times J\times K}) \\ &\geqslant \mathbb{E}((aW_{12}^2 - AW_{12} - A)\mathbb{1}_{(X,Y,Z)\in I\times J\times K}) \\ &= \mathbb{E}(aW_{12}^2 - AW_{12} - A)\mathbb{P}((X, Y, Z) \in I \times J \times K) = +\infty.\end{aligned}$$

2.3.4 Erdős-Rényi matrices. We will use Theorem 2.1.1 (convexity-preserving functional) and model an Erdős-Rényi matrix $H_N - E$ as $H_N - E = \Phi(X_1, \dots, X_M)$, where $M = \frac{N(N+1)}{2}$, the X_i's are independent Bernoulli random variables with parameter p_N, and Φ is $\frac{1}{\sqrt{Np_N(1-p_N)}}$ times the identity map which places these entries in the upper triangle of an $N \times N$ matrix, minus $E\,\mathrm{Id}$. This clearly satisfies assumptions (I) and (M) with $\|\Phi\|_{\mathrm{Lip}} = \frac{1}{\sqrt{Np_N(1-p_N)}}$.

Now we verify assumption (E) with all μ_N's equal to the semicircle law ρ_{sc}. In the proof, we control the extreme eigenvalues (more precisely the smallest and second-largest) with results of Vu [154], improving on earlier results of Füredi-Komlós [83]; and we control the bulk eigenvalues using the local law of Erdős *et al.* [74]. Often we use much weaker consequences of the results, replacing $\log N$ factors by polynomial factors and so on.

More precisely, consider $\widetilde{H_N} = H_N - \mathbb{E}[H_N]$. This matrix has centered entries of variance $\sigma^2 = \frac{1}{N}$, supported in $[-K, K]$ with $K = \frac{1}{\sqrt{\varepsilon N^\varepsilon}}$. Thus the proof of [154, Theorem 1.3, Theorem 1.4] shows that there exist $C, \gamma > 0$ with

$$\mathbb{P}\left(\|\widetilde{H_N}\| > 2 + C\frac{\log N}{(\varepsilon N^\varepsilon)^{1/4}}\right) \leqslant N^{-\gamma} \tag{2.3.6}$$

for N large enough. Recall we order eigenvalues as $\lambda_1 \leqslant \cdots \leqslant \lambda_N$; since $\mathbb{E}[H_N]$ is rank-one and positive semidefinite, interlacing tells us that $\max(|\lambda_1(H_N)|, |\lambda_{N-1}(H_N)|) \leqslant \|\widetilde{H_N}\|$, and thus we

have the very coarse bound

$$\mathbb{P}(\max(|\lambda_1(H_N)|, |\lambda_{N-1}(H_N)|) \geqslant 3) \leqslant N^{-\gamma}$$

for N large enough. In particular, whenever f is a test function with $\|f\|_\infty \leqslant 1$, we have

$$\left| \int_3^\infty f(x)(\mathbb{E}[\hat{\mu}_{H_N}] - \rho_{\mathrm{sc}})(\mathrm{d}x) \right| \leqslant \frac{1}{N} \sum_{i=1}^N \mathbb{P}(\lambda_i(H_N) \geqslant 3) \leqslant \frac{1}{N} + N^{-\gamma},$$

and similarly for the left tail, which is even easier because we do not need to separate out the smallest eigenvalue.

Now we handle the bulk eigenvalues. Let $F_{\rho_{\mathrm{sc}}}$, $F_{\hat{\mu}}$, and $F_{\mathbb{E}[\hat{\mu}]}$ be the distribution functions for ρ_{sc}, $\hat{\mu}_{H_N}$, and $\mathbb{E}[\hat{\mu}_{H_N}]$, respectively. Then [74, Theorem 2.12] shows that there exists $\nu > 0$ such that, for N large enough,

$$\mathbb{P}\left(\sup_{x \in [-3,3]} \left| F_{\rho_{\mathrm{sc}}}(x) - F_{\hat{\mu}(x)} \right| \leqslant N^{-1+\varepsilon} \right) \geqslant 1 - \exp(-\nu(\log N)^{5 \log\log N}).$$

Since $\sup_x |F_{\rho_{\mathrm{sc}}}(x) - F_{\hat{\mu}}(x)| \leqslant 2$ deterministically, this gives

$$\sup_{x \in [-3,3]} \left| F_{\rho_{\mathrm{sc}}}(x) - F_{\mathbb{E}[\hat{\mu}]}(x) \right| \leqslant \frac{N^\varepsilon}{N} + 2\exp(-\nu(\log N)^{5 \log\log N}).$$

The proof of (E) is then easily completed as in the case of Wigner matrices.

Now we check the three estimates comprising assumption (C) on coarse bounds.

(2.1.4) We have

$$\|H_N\|^2 \leqslant \sum_{i,j} |H_{ij}|^2 \leqslant \frac{N}{p_N(1-p_N)} \leqslant \frac{1}{\varepsilon} N^{2-\varepsilon} \tag{2.3.7}$$

almost surely, so (2.1.4) is trivially satisfied.

(2.1.5) For bulk energy levels, meaning $E \in (-2, 2)$, one can show

$$\mathbb{P}\Big(H_N \text{ has no eigenvalues in } \Big(E - \frac{1}{N^2}, E + \frac{1}{N^2}\Big)\Big) = 1 - o(1)$$

using the bulk fixed-energy universality results of Landon-Sosoe-Yau [113, Section 1.1.1]; the argument is given in our discussion below of the free-addition model. For $|E| > 2$, eigenvalues other than λ_N are handled with the result of Vu above (2.3.6). For λ_N (only a concern for positive E values), the Weyl inequalities give

$$\begin{aligned}\lambda_N(H_N) &\geqslant \lambda_N(\mathbb{E}[H_N]) + \lambda_1(H_N - \mathbb{E}[H_N]) = \sqrt{\frac{Np_N}{1-p_N}} + \lambda_1(H_N - \mathbb{E}[H_N]) \\ &\geqslant \frac{1}{\sqrt{\varepsilon}} N^{\varepsilon/2} + \lambda_1(H_N - \mathbb{E}[H_N]).\end{aligned}$$

By (2.3.6), the last term is at least -3 with probability $1 - o(1)$; thus λ_N cannot stick to any fixed $E > 2$.

(2.1.6) This follows from (2.3.7), using $|\det(H_N - E)| \leqslant \|H_N - E\|^N$.

For assumption (S) on spectral stability, we note that the threshold for cutting is

$$\frac{N^{-\kappa}}{\|\Phi\|_{\text{Lip}}} = N^{\frac{1}{2}-\kappa}\sqrt{p_N(1-p_N)} \geqslant \sqrt{\varepsilon} N^{\frac{\varepsilon}{2}-\kappa} \geqslant 1$$

for $\kappa < \frac{\varepsilon}{2}$ and large enough N. Since the $X_i = 0$ or 1, this means that $X = X_{\text{cut}}$, and hence (2.1.7) is trivially satisfied.

2.3.5 Band matrices. We will use Theorem 2.1.1 (convexity-preserving functional) and model a band matrix H_N as $H_N = \Phi(X_1, \ldots, X_M)$, where $M = (W+1)N$, the X_i's are independent random variable distributed according to μ, and Φ is $\frac{1}{\sqrt{2W+1}}$ times the identity map which arranges these entries into a band matrix. This Φ is trivially convex and satisfies $\|\Phi\|_{\text{Lip}} = \frac{1}{\sqrt{2W+1}}$. Throughout this section, the constant ε will be the same as in the assumption $W \geqslant N^{\varepsilon}$.

To check Assumption (E) with $\mu_N \equiv \rho_{\mathrm{sc}}$, we will use Proposition 2.3.1 with $A = 3$. (By translation invariance, it suffices to check (E) at $E = 0$.) The bulk estimate (2.3.1) follows from the stronger local law of Erdős *et al.* [77]; the tail estimate (2.3.2) uses the tail estimates of Benaych-Georges/Péché [46].

Write $s_N(z)$ for the Stieltjes transform of $\hat{\mu}_{H_N}$ and $m_{\mathrm{sc}}(z)$ for the Stieltjes transform of the semicircle law. The local law [77, Theorem 2.1] gives constants C and c such that, if $z = E + \mathrm{i}\eta$ with $E \leqslant 3$, $\kappa := ||E| - 2| \geqslant N^{-\delta}$ for $\delta = 3\varepsilon/20$, and $\eta = N^{6\delta-\varepsilon}$, then

$$\mathbb{P}(|s_N(z) - m_{\mathrm{sc}}(z)| \geqslant N^{-\delta}) \leqslant CN^{-c(\log\log N)}.$$

Together with the trivial bound $|\mathbb{E}[s_N(E + \mathrm{i}\eta)] - m_{\mathrm{sc}}(E + \mathrm{i}\eta)| \leqslant \frac{2}{\eta}$, this gives

$$|\mathbb{E}[s_N(z)] - m_{\mathrm{sc}}(z)| \leqslant N^{-\delta} + 2CN^{\varepsilon-6\delta-c(\log\log N)} \lesssim N^{-\delta}$$

for such z values. Writing $\varepsilon_1 = \varepsilon - 6\delta > 0$ and using again the trivial bound for $\kappa < N^{-\delta}$, we obtain

$$\int_{-3}^{3} |\mathbb{E}[s_N(E + \mathrm{i}N^{-\varepsilon_1})] - m_{\mathrm{sc}}(E + \mathrm{i}N^{-\varepsilon_1})| \lesssim 6N^{-\delta} + 8N^{\varepsilon_1-\delta}.$$

By our choice of δ we have $\varepsilon_1 - \delta < 0$; this suffices to check (2.3.1).

For the tail estimate (2.3.2), we note that $\left|F_{\mathbb{E}[\hat{\mu}]}(x) - F_{\rho_{\mathrm{sc}}}(x)\right| \leqslant \mathbb{P}(\|H_N\| \geqslant x)$ for, say, $x \geqslant 3$. The proof of [46, Theorem 1.4] gives, for any $k \geqslant 1$,

$$\mathbb{P}(\|H_N\| \geqslant x) \leqslant Nx^{-2k}4^k\left(1 - \frac{(12\alpha/e)^6k^{12}}{W}\right)^{-1}.$$

Choosing $k = k_N = N^{\varepsilon/20}$, we verify (2.3.3) and find, for N large enough,

$$\int_3^\infty \mathbb{P}(\|H_N\| \geqslant x)\,\mathrm{d}x \leqslant 6N^{1-\frac{\varepsilon}{20}}\left(\frac{4}{9}\right)^{N^{\varepsilon/20}},$$

which is much faster than we need. The left tail is estimated similarly, and this verifies (2.3.2) and

thus (2.1.3).

Now we check assumption (C) on coarse bounds.

(2.1.4) The proof for Wigner matrices with $2+\varepsilon$ moments works verbatim here.

(2.1.5) Since we assumed our entries have a bounded density, this follows from Proposition 2.3.2 and Lemma 2.3.3.

(2.1.6) The proof for Wigner matrices with $2+\varepsilon$ moments works verbatim here.

The proof of Assumption (S) is similar to the case of Wigner matrices; in particular it holds assuming only that μ has $2+\varepsilon$ finite moments.

2.3.6 Sample covariance matrices. As noted above, this model is not covered by either of our theorems directly. But it can be proved by mimicking the proof of Theorem 2.1.1 (convexity-preserving functional) with the following changes. We let $M = pN$, let $X_1, \dots, X_M$ be independent copies of μ, and consider the map $\Phi = \Phi_E : \mathbb{R}^M \to \mathscr{S}_p$ that places its arguments in the entries of the $p \times N$ matrix $Y = Y_{p,N}$ and returns $\frac{1}{N} YY^T - E$. There are two problems with applying Theorem 2.1.1 as written, but we will implement the following workarounds:

1. Φ is not convex (but we will use the standard Hermitization trick that compares eigenvalues of YY^T with eigenvalues of the $(p+N) \times (p+N)$ block matrix $\left(\begin{smallmatrix} 0 & Y \\ Y^T & 0 \end{smallmatrix}\right)$, which *is* a convex function of the entries of Y).

2. Φ is not Lipschitz, since it grows too quickly at infinity (but the Hermitization is Lipschitz).

Below, we will verify assumption (E) with some value of κ. For now, we redefine X_{cut} (for this model only), using this same κ, as

$$(X_{\text{cut}})_i = X_i \mathbb{1}_{|X_i| \leqslant N^{-\kappa+\frac{1}{2}}}. \tag{2.3.8}$$

We choose this scaling so that $\Phi(X_{\text{cut}})+E$ has entries at most $N^{-2\kappa}$, similar to what happens in the Wigner and Erdős-Rényi cases. Later we will check assumption (S) with this new definition, as well as assumption (C). First we show that all of these assumptions yield determinant concentration.

Much of the proof of Theorem 2.1.1 works verbatim in this new setting, since for example it never uses the old definition of X_{cut} directly, using instead the stability estimate (2.1.7) which will still be true for us under the new definition. The biggest change is in the proof of Lemma 2.2.3, where we applied results of Talagrand using the convexity and Lipschitz properties which no longer hold. The replacement for Lemma 2.2.3 is as follows.

Lemma 2.3.4. *Let* $\widetilde{t_0}(N) = \frac{4\sqrt{2\pi}\cdot 3^{7/4}}{\sqrt{\eta}N^{\frac{1}{2}+\kappa}}$. *If* $t \geqslant \widetilde{t_0}(N)$, *then*

$$\mathbb{P}((\mathcal{E}_{\text{conc}})^c) \leqslant 12\exp\left(-\frac{(t-\widetilde{t_0}(N))^2\eta N^{1+2\kappa}}{171\sqrt{3}}\right).$$

Proof. By translation-invariance, it suffices to check this at $E = 0$. We use the classical trick of considering the $(p+N)\times(p+N)$ matrix

$$\mathfrak{H}_N = \mathfrak{H}_N(X) = \frac{1}{\sqrt{N}}\begin{pmatrix} 0_{p\times p} & Y_{p,N} \\ Y_{p,N}^T & 0_{N\times N}\end{pmatrix}.$$

For any test function f, we have

$$\operatorname{tr}(f(\mathfrak{H}_N^2)) = 2\operatorname{tr}(f(YY^T/N)) + (N-p)f(0). \tag{2.3.9}$$

Thus we need to consider Lipschitz, convex decompositions of the function $x \mapsto \log_\eta^K(x^2)$, which

we do as follows:

$$\widetilde{\log}_1(x) = \begin{cases} -\frac{3^{3/4}}{2\sqrt{\eta}}x - \frac{3}{2} + \log_\eta(\eta\sqrt{3}) & \text{if } x \leqslant -3^{1/4}\sqrt{\eta}, \\ \log_\eta(x^2) & \text{if } -3^{1/4}\sqrt{\eta} \leqslant x \leqslant 3^{1/4}\sqrt{\eta}, \\ \frac{3^{3/4}}{2\sqrt{\eta}}x - \frac{3}{2} + \log_\eta(\eta\sqrt{3}) & \text{if } x \geqslant 3^{1/4}\sqrt{\eta}, \end{cases}$$

$$\widetilde{\log}_2(x) = \begin{cases} \frac{3^{3/4}}{2\sqrt{\eta}}x & \text{if } x \leqslant 3^{1/4}\sqrt{\eta}, \\ \log_\eta^K(x^2) + \frac{3}{2} - \log_\eta(\eta\sqrt{3}) & \text{if } x \geqslant 3^{1/4}\sqrt{\eta}, \end{cases}$$

$$\widetilde{\log}_3(x) = \begin{cases} -\frac{3^{3/4}}{2\sqrt{\eta}}x & \text{if } x \geqslant -3^{1/4}\sqrt{\eta}, \\ \log_\eta^K(x^2) + \frac{3}{2} - \log_\eta(\eta\sqrt{3}) & \text{if } x \leqslant -3^{1/4}\sqrt{\eta}. \end{cases}$$

Notice that $\log_\eta^K(x^2) = \sum_{i=1}^3 \widetilde{\log}_i(x)$, that $\widetilde{\log}_1$ is convex while $\widetilde{\log}_2$ and $\widetilde{\log}_3$ are concave, and that each $\widetilde{\log}_i$ is $\frac{3^{3/4}}{2\sqrt{\eta}}$-Lipschitz. Then we consider the functions $\widetilde{f}_i : [-N^{-\kappa+1/2}, N^{-\kappa+1/2}]^M \to \mathbb{R}$ given by

$$\widetilde{f}_i(X) = (-1)^{\mathbb{1}_{i\neq 1}} \frac{1}{N} \operatorname{tr}(\widetilde{\log}_i(\mathfrak{H}_N(X))).$$

Using (2.3.9) and mimicking the proof of Lemma 2.2.3, we find

$$\mathbb{P}(\mathcal{E}_{\text{conc}}^c) = \mathbb{P}\bigg(\Big|\frac{1}{2N}\operatorname{tr}(\log_\eta^K(\mathfrak{H}_N^2)) - \frac{1}{2N}\mathbb{E}[\operatorname{tr}(\log_\eta^K(\mathfrak{H}_N^2))]\Big| > t\bigg) \leqslant \sum_{i=1}^N \mathbb{P}\bigg(\Big|\widetilde{f}_i(X) - \mathbb{E}[\widetilde{f}_i(X)]\Big| > \frac{2}{3}t\bigg).$$

As in the original proof, each $\widetilde{f}_i$ is $\sqrt{2N}\frac{1}{N}\frac{3^{3/4}}{2\sqrt{\eta}}\frac{1}{\sqrt{N}} = \frac{3^{3/4}}{N\sqrt{2\eta}}$-Lipschitz, and since the map $X \mapsto \mathfrak{H}_N$ is convex (this is the point of the Hermitization) we know that each $\widetilde{f}_i$ is convex as well. Then Talagrand's inequality gives

$$\mathbb{P}\Big(\Big|\widetilde{f}_i - \mathfrak{M}_{f_i}\Big| \geqslant t\Big) \leqslant 4\exp\bigg(-\frac{t^2\eta N^{1+2\kappa}}{96\sqrt{3}}\bigg)$$

and we conclude as before. □

It remains to check assumptions (E), (C), and (S) (the latter under the new definition (2.3.8)). The only assumption that is not translation-invariant (i.e., that depends on the energy level E) is assumption (C).

For assumption (E), [148] proved that if μ has $2+\gamma$ moments then there exists (explicit) $\kappa(\gamma) > 0$ such that

$$d_{\mathrm{KS}}(\mathbb{E}[\hat{\mu}_{\frac{1}{N}YY^T}], \mu_{\mathrm{MP},\frac{p_N}{N}}) \lesssim N^{-\kappa(\gamma)}.$$

From this Kolmogorov-Smirnov distance information we evaluate d_{BL} in the same way as for Wigner matrices with $2+\gamma$ moments. It remains only to understand $d_{\mathrm{BL}}(\mu_{MP,\frac{p_N}{N}}, \mu_{MP,\gamma})$, and this is only necessary in the case $\gamma < 1$ (since when $\gamma = 1$ we assumed $p_N = N$). If $\gamma_1, \gamma_2 \in [\varepsilon, 1-\varepsilon]$, then the difference between the densities gives

$$d_{\mathrm{BL}}(\mu_{\mathrm{MP},\gamma_1}, \mu_{\mathrm{MP},\gamma_2}) = O_\varepsilon\left(\sqrt{|\gamma_1 - \gamma_2|}\right).$$

Since we assumed in (2.1.13) that $\left|\frac{p_N}{N} - \gamma\right|$ is polynomially small, this suffices to prove (2.1.3).

We check the three estimates of assumption (C) as follows:

(2.1.4) This follows the proof of the Wigner case, but using the Weyl inequalities for singular values instead of those for eigenvalues. We write out the beginning of the argument because some of the powers change. For some $\varepsilon > 0$, write $Y/\sqrt{N} = A + B$, where A is defined entrywise by

$$A_{ij} = \frac{1}{\sqrt{N}} Y_{ij} \mathbb{1}_{\frac{1}{\sqrt{N}}|Y_{ij}| \leqslant \frac{1}{10N} e^{\frac{1}{2}N^\varepsilon}}.$$

Then A has singular values at most $\frac{1}{10}e^{\frac{1}{2}N^\varepsilon}$, and the Weyl inequalities give

$$\sigma_i(Y/\sqrt{N}) \leqslant \sigma_{\max}(A) + \sigma_i(B) \leqslant \frac{1}{10}e^{\frac{1}{2}N^\varepsilon} + \sigma_i(B)$$

and similarly $\sigma_i(Y/\sqrt{N}) \geqslant \sigma_i(B) - \frac{1}{10}e^{\frac{1}{2}N^\varepsilon}$, so that for each i we have

$$\begin{aligned} 1 + \left|\lambda_i(YY^T/N - E)\right| \mathbb{1}_{|\lambda_i(YY^T/N-E)|>e^{N^\varepsilon}} &\leqslant 1 + 2\lambda_i(YY^T/N)\mathbb{1}_{\lambda_i(YY^T/N)>\frac{1}{2}e^{N^\varepsilon}} \\ &= 1 + 2\sigma_i^2(Y/\sqrt{N})\mathbb{1}_{\sigma_i(Y/\sqrt{N})>\frac{1}{\sqrt{2}}e^{\frac{1}{2}N^\varepsilon}} \\ &\leqslant 1 + 8\sigma_i^2(B)\mathbb{1}_{\sigma_i(B)>\frac{1}{2}e^{\frac{1}{2}N^\varepsilon}}. \end{aligned}$$

Then from Fischer's inequality we have

$$\prod_{i=1}^{p}(1 + 8\sigma_i^2(B)\mathbb{1}_{\sigma_i(B)>\frac{1}{2}e^{\frac{1}{2}N^\varepsilon}}) \leqslant \det(\mathrm{Id} + 8BB^T) \leqslant \prod_{i=1}^{p}\left(1 + 8\sum_{j=1}^{N} B_{ij}^2\right).$$

Since B is non-Hermitian with independent entries, the same argument as in the Wigner case goes through here: when we expand and factor, each matrix entry appears at a power at most two.

(2.1.5) We mimic the proofs from Section 2.3.2, making the following changes. We closely follow the proof of some Wegner estimates for complex Wigner matrices from [76, Theorem 3.4], as adapted in [58, Proposition B.1] to the symmetric case. Our estimates below will be coarser as we can afford any polynomial error, contrary to the optimal estimates from these references. Let $E, \eta > 0$, $\eta = \varepsilon/N$, $I = [E - \eta, E + \eta]$, $z = E + \mathrm{i}\eta$ and $\mathcal{N}_I = |\{\mu_i \in I\}|$. In the covariance matrix setting, the Schur complement formula gives, for any $1 \leqslant j \leqslant N$ and defining $X = Y/\sqrt{N}$ and $H = YY^*/N$ (see e.g. [53, Equation (3.8)])

$$((H - z)^{-1})_{ii} = \big(- z - zX_i^* R_i(z) X_i\big)^{-1}$$

where we define $X_i = (X_{ij})_j$, $X_{jk}^{(i)} = X_{jk}\mathbb{1}_{j\neq i}$ and $R_i(z) = ((X^{(i)})^* X^{(i)} - z)^{-1}$. This implies,

by the Cauchy-Schwarz inequality,

$$\mathbb{E}[\mathcal{N}_I^2 \mathbb{1}_A] \leqslant C(N\eta)^2 \mathbb{E}\left[\Big(\operatorname{Im} \frac{1}{-z - \frac{z}{N}\sum_{\alpha=1}^N \frac{\xi_\alpha}{\lambda_\alpha - z}}\Big)^2 \mathbb{1}_A\right]$$
$$\leqslant C\varepsilon^2 \mathbb{E}\left[\Big((\sum_{\alpha=1}^N c_\alpha \xi_\alpha)^2 + (E - \sum_{\alpha=1}^N d_\alpha \xi_\alpha)^2\Big)^{-1} \mathbb{1}_A\right]$$

for any event A, where

$$d_\alpha = -\frac{1}{N} - \frac{N\lambda_\alpha(E-\lambda_\alpha)}{N^2(\lambda_\alpha - E)^2 + \varepsilon^2}, \quad c_\alpha = \frac{\lambda_\alpha \varepsilon}{N^2(\lambda_\alpha - E)^2 + \varepsilon^2},$$

with $(\lambda_\alpha)_{1\leqslant\alpha\leqslant N-1}$ the eigenvalues of $(X^{(1)})^* X^{(1)}$, with corresponding L^2-normalized eigenvectors u_α's, and $\xi_\alpha = |u_\alpha \cdot Y_1|^2$.

Let $(\gamma_k)_{1\leqslant k\leqslant N}$ be implicitly defined through $\int_0^{\gamma_k} \mu_{\mathrm{MP},\gamma}(\mathrm{d}x) = k/N$, with $\mu_{\mathrm{MP},\gamma}$ from (2.1.14). If $E < \gamma_{\lfloor N/2\rfloor}$, we define $m = \lfloor 3N/4\rfloor$. If $E > \gamma_{\lfloor N/2\rfloor}$, let $m = \lfloor N/4\rfloor$. Convergence of $N^{-1}\sum_{k=1}^N \delta_{\mu_k}$ ($\mu_1, \dots, \mu_N$ are the eigenvalues of H) to $\mu_{\mathrm{MP},\gamma}$ under the minimal assumption of finite second moment of the entries [155] has the following elementary consequence: For any $c > 0$, $\mathbb{P}(A_N) = 1 - \mathrm{o}(1)$ where $A_N = \cap_{N/7<k<8N/7}\{|\mu_k - \gamma_k| < c\}$. By interlacing, on A_N the $(d_{m+\ell})_{0\leqslant\ell\leqslant 3}$ all have the same sign and absolute value greater than N^{-2}, and $c_m, c_{m+1} > c\varepsilon/N^2$. Hence we can apply [58, Equation (B.4)][2] with $\tau = 0, r = p = 2$ (and either E or $-E$ depending on the sign of the $d_{m+\ell}$'s) to obtain, on A_N,

$$\mathbb{E}_{Y_1}\left[\Big((\sum_{\alpha=1}^N c_\alpha \xi_\alpha)^2 + (E - \sum_{\alpha=1}^{N-1} d_\alpha \xi_\alpha)^2\Big)^{-1}\right] \leqslant \frac{C}{\sqrt{c_m c_{m+1}}\min(d_{m+1}, d_{m+2}, d_{m+3})} \leqslant C\frac{N^{10}}{\varepsilon},$$

so that $\mathbb{E}[\mathcal{N}_I^2 \mathbb{1}_{A_N}] \leqslant N^{10}\varepsilon$ and in particular $\mathbb{P}(\mathcal{N}_I \geqslant 1) \to 0$ for $\varepsilon = e^{-N^\varepsilon}$.

(2.1.6) This proof has the same idea as the one for Wigner matrices; the only difference is that the product of entries associated to one permutation is estimated as follows. Fix δ so small that

[2] The assumption (2.1.12) is exactly the needed input for [58, Lemma B.4]. Note that although this Lemma assumes μ has finite moments of all orders, this is actually not used in its proof.

μ has finite $2+2\delta$ moment. For any permutation $\sigma \in S_p$ define

$$X_\sigma = \Big|(YY^T/N - E)_{1,\sigma(1)} \cdot \ldots \cdot (YY^T/N - E)_{p,\sigma(p)}\Big|.$$

Let

$$c_\delta = \max(\mathbb{E}[|Y_{1,1}|^{1+\delta}], \mathbb{E}[|Y_{1,1}|^{2+2\delta}]) < \infty.$$

Then from convexity of $x \mapsto x^{1+\delta}$ we have

$$\left(\frac{1}{N^p}\sum_{j_1,\ldots,j_p=1}^{N}\prod_{i=1}^{p}\Big|Y_{i,j_i}Y_{\sigma(i),j_i} - E\delta_{i,\sigma(i)}\Big|\right)^{1+\delta} \leqslant \frac{1}{N^p}\sum_{j_1,\ldots,j_p=1}^{N}\left(\prod_{i=1}^{p}\Big|Y_{i,j_i}Y_{\sigma(i),j_i} - E\delta_{i,\sigma(i)}\Big|\right)^{1+\delta},$$

and thus

$$\begin{aligned}
\mathbb{E}[X_\sigma^{1+\delta}] &= \mathbb{E}\left[\left(\prod_{i=1}^{p}\left|\frac{1}{N}\sum_{j=1}^{N}(Y_{i,j}Y_{\sigma(i),j} - E\delta_{i,\sigma(i)})\right|\right)^{1+\delta}\right] \\
&\leqslant \frac{1}{N^p}\sum_{j_1,\ldots,j_p=1}^{N}\mathbb{E}\left[\left(\prod_{i=1}^{p}(\Big|Y_{i,j_i}Y_{\sigma(i),j_i}\Big| + |E|)\right)^{1+\delta}\right] \\
&=: \frac{1}{N^p}\sum_{j_1,\ldots,j_p=1}^{N}\mathbb{E}[(Z_{j_1,\ldots,j_p})^{1+\delta}].
\end{aligned}$$

Now each $Z_{j_1,\ldots,j_p}$ is the sum of 2^p terms, each of the form $|E|^{p-k}\prod_{\ell=1}^{k}|Y_{i_\ell,j_{i_\ell}}Y_{\sigma(i_\ell),j_{i_\ell}}|$ for some $k \in [\![1,p]\!]$ and some collection of *distinct* integers $i_1,\ldots,i_k \in [\![1,p]\!]$. Since they are distinct, each entry of the matrix Y appears with power at most two in such a term; since these entries are independent, we have

$$\mathbb{E}\left[\left(|E|^{p-k}\prod_{\ell=1}^{k}|Y_{i_\ell,j_{i_\ell}}Y_{\sigma(i_\ell),j_{i_\ell}}|\right)^{1+\delta}\right] \leqslant |E|^{(p-k)(1+\delta)}c_\delta^{2k} \leqslant \max(|E|,c_\delta,1)^{2p(1+\delta)} =: c_{\delta,E}^{2p(1+\delta)}$$

Then Minkowski's inequality in $L^{1+\delta}$ gives

$$\mathbb{E}[X_\sigma^{1+\delta}] \leqslant \sup_{j_1,\dots,j_p} \mathbb{E}[(Z_{j_1,\dots,j_p})^{1+\delta}] \leqslant (2c_{\delta,E}^2)^{p(1+\delta)}.$$

The rest of the proof is similar to the Wigner case.

Finally we check assumption (S) with the new definition (2.3.8). Write $Y_{\text{cut}} = \Phi(X_{\text{cut}})$ for the $p \times N$ matrix Y with entries truncated at level $N^{-\kappa+1/2}$; then it is classical that

$$d_{\text{KS}}(\hat{\mu}_{YY^T/N}, \hat{\mu}_{Y_{\text{cut}}Y_{\text{cut}}^T/N}) \leqslant \frac{1}{p} \operatorname{rank}(Y - Y_{\text{cut}})$$

(this follows from interlacing of singular values; see, e.g., [18, Theorem A.44]). The rest of the argument with Bennett's inequality goes through from here; note that $\mathbb{P}(|W_{ij}| > N^{-\kappa})$ and $P(|Y_{ij}| > N^{-\kappa+1/2})$ are of similar order because Y has order-one entries but the Wigner matrix W has order-$\frac{1}{\sqrt{N}}$ entries.

2.3.7 Gaussian matrices with a (co)variance profile. We will use Theorem 2.1.2 (concentrated input) to prove Corollary 2.1.8.B. First we need the following sequence of lemmas establishing consequences of our model assumptions (such as the log-Sobolev inequality and tail decay estimates).

Lemma 2.3.5. *Let $\mathcal{C} = \mathcal{C}_N$ be the covariance matrix of the upper triangle of H_N considered as a Gaussian vector, i.e., $\mathcal{C}$ is an $\frac{N(N+1)}{2} \times \frac{N(N+1)}{2}$ matrix with entries*

$$\mathcal{C}_{(i,j),(k,\ell)} = \operatorname{Cov}(H_{ij}, H_{k\ell}) = \operatorname{Cov}(W_{ij}, W_{k\ell}).$$

Let p be as in the weak-fullness assumption (wF). Then, in the sense of quadratic forms,

$$\mathcal{C} \geqslant N^{-1-p} \operatorname{Id}.$$

Proof. We claim that

$$W \overset{(d)}{=} N^{-\frac{p}{2}} W^{(\mathrm{GOE})} + W' \tag{2.3.10}$$

where W_{GOE} is distributed as a GOE matrix (i.e., independent Gaussian entries up to symmetry with $\mathbb{E}[(W_{ij}^{(\mathrm{GOE})})^2] = \frac{1+\delta_{ij}}{N}$) and where W' is some real symmetric Gaussian matrix independent of $W^{(\mathrm{GOE})}$.

Indeed, consider the $N^2 \times N^2$ covariance matrix $\mathscr{C}_W$ of the *full* matrix W (not just the upper triangle). We will index this by matrix locations, i.e., $\mathscr{C}_W$ has entries $(\mathscr{C}_W)_{(i,j),(k,\ell)}$. Write $\mathscr{C}_{\mathrm{GOE}}$ for the covariance matrix for GOE. We index a vector $B \in \mathbb{R}^{N^2}$ similarly, writing $B_{(i,j)}$, and associate with it the matrix $\tilde{B} \in \mathbb{R}^{N \times N}$ defined by

$$\tilde{B}_{ij} = B_{(i,j)}.$$

Notice that the matrix $\tilde{B}$ need not be symmetric. Whenever B has unit norm, we have

$$\langle B, \mathscr{C}_{\mathrm{GOE}} B \rangle = \frac{1}{N} \sum_{i,j,k,\ell} B_{(i,j)} (\delta_{ik}\delta_{j\ell} + \delta_{i\ell}\delta_{jk}) B_{(k,\ell)} = \frac{1}{N} \operatorname{Tr}(\tilde{B}\tilde{B}^T + \tilde{B}^2) = \frac{1}{N} \operatorname{Tr}\left(\left(\frac{\tilde{B} + \tilde{B}^T}{2} \right)^2 \right).$$

Thus by the weak-fullness assumption (wF) we have

$$\begin{aligned} \langle B, \mathscr{C}_W B \rangle = \mathbb{E}\Big[(\operatorname{Tr}(\tilde{B}W))^2 \Big] &= \mathbb{E}\left[\left(\operatorname{Tr}\left(\left(\frac{\tilde{B} + \tilde{B}^T}{2} \right) W \right) \right)^2 \right] \\ &\geqslant N^{-1-p} \operatorname{Tr}\left(\left(\frac{\tilde{B} + \tilde{B}^T}{2} \right)^2 \right) = \langle B, N^{-p} \mathscr{C}_{\mathrm{GOE}} B \rangle. \end{aligned}$$

To complete the proof of (2.3.10), we write

$$\mathscr{C}_W = N^{-p} \mathscr{C}_{\mathrm{GOE}} + (\mathscr{C}_W - N^{-p} \mathscr{C}_{\mathrm{GOE}})$$

and interpret the matrix in parentheses on the right-hand side, which we just showed is positive semi-definite, as the covariance matrix for W'.

Now we consider the $\frac{N(N+1)}{2} \times \frac{N(N+1)}{2}$ covariance matrix $\mathcal{C} = \mathcal{C}_W$ of the *upper triangle* of W, and define $\mathcal{C}_{\mathrm{GOE}}$ and $\mathcal{C}_{W'}$ similarly. Then whenever $v \in \mathbb{R}^{\frac{N(N+1)}{2}}$ is indexed with upper-triangular entries we have

$$\begin{aligned}\langle v, \mathcal{C}_W v\rangle &= \langle v, N^{-p}\mathcal{C}_{\mathrm{GOE}} v\rangle + \langle v, \mathcal{C}_{W'} v\rangle \geqslant N^{-p}\langle v, \mathcal{C}_{\mathrm{GOE}} v\rangle \\ &= N^{-1-p}\left(\sum_{i\leqslant j} v_{(i,j)}^2 + \sum_i v_{(i,i)}^2\right) \geqslant N^{-1-p}\|v\|_2^2\end{aligned}$$

which concludes the proof. □

Lemma 2.3.6. *For every $\zeta > 0$, there exists $c_\zeta > 0$ such that the law of the upper triangle of H_N, considered as a vector, satisfies the logarithmic-Sobolev inequality with constant $c_\zeta \frac{N^\zeta}{N}$.*

Proof. Since the logarithmic-Sobolev inequality is preserved under translations, it suffices to prove the statement with $H_N = W_N + \mathbb{E}[H_N]$ replaced by W_N. This is essentially an exercise in spelling out our model assumptions, which come from [75].

The upper triangle of W_N is a Gaussian vector with covariance matrix $\mathcal{C}$. Define the matrix $|\mathcal{C}|$ by $|\mathcal{C}|_{(i,j),(k,\ell)} = \left|\mathcal{C}_{(i,j),(k,\ell)}\right|$, and whenever $u \in \mathbb{R}^{\frac{N(N+1)}{2}}$ is a unit vector, define the unit vector $|u|$ by $|u|_{(i,j)} = \left|u_{(i,j)}\right|$. Then

$$\langle u, \mathcal{C}u\rangle \leqslant \langle |u|, |\mathcal{C}||u|\rangle \leqslant |||\mathcal{C}|||.$$

But our assumptions (D) on correlation decay imply that $|||\mathcal{C}||| \leqslant_\zeta \frac{N^\zeta}{N}$; see [75, (6b), Assumption (C)], specifically noting that $\|\kappa\|_2^{\mathrm{av}}$ in their notation is the same as $N|||\mathcal{C}|||$ in ours (the factor N appears since their normalization is $H_N = A_N + \frac{1}{\sqrt{N}} W_N$ to our $H_N = A_N + W_N$).

Since $\mathcal{C}$ is invertible by Lemma 2.3.5, this implies the log-Sobolev inequality via the Bakry-Émery criterion. □

Lemma 2.3.7. *The flatness assumption (F) implies, for each i, j, N,*

$$\frac{1}{pN} \leqslant \mathrm{Var}((W_N)_{ij}) \leqslant \frac{p}{N}.$$

Proof. By writing e_j for the jth canonical basis vector, understood as a column, and writing $(\cdot)^T$ for transposition, we have $\mathbb{E}[W_{ij}^2] = \mathbb{E}[W_{ij}W_{ji}] = \mathbb{E}[We_j(e_j)^TW]_{ii} = (e_i)^T\mathbb{E}[We_j(e_j)^TW]e_i$, but by the flatness assumption (F) we have $\frac{1}{pN} = \frac{1}{pN}\operatorname{Tr}(e_j(e_j)^T) \leqslant (e_i)^T\mathbb{E}[We_j(e_j)^TW]e_i \leqslant \frac{p}{N}\operatorname{Tr}(e_j(e_j)^T) = \frac{p}{N}$. □

Lemma 2.3.8. *We have* $\sup_N \mathbb{E}[\|H_N\|] < \infty$.

Proof. Since we assumed $\sup_N \|A_N\| < \infty$, we need only check $\sup_N \mathbb{E}[\|W_N\|] < \infty$ where $W = W_N = H_N - \mathbb{E}(H_N)$. We apply the relevant local law from [75]. This local law provides a sequence of measures $\widetilde{\mu_N}$ which well-approximate the empirical measure of W. The exact form of $\widetilde{\mu_N}$ does not matter for our purpose; what does matter is [5, Proposition 2.1, Equation (4.2)], which we combine to obtain $\operatorname{supp}(\widetilde{\mu_N}) \subset [-2\sqrt{2p}, 2\sqrt{2p}]$ uniformly in N. Then the local law [75, Corollary 2.3] implies that eigenvalues of W stick to $\operatorname{supp}(\widetilde{\mu_N})$ in the sense that, for some constant C, we have

$$\mathbb{P}(\|W\| \geqslant 2\sqrt{2p} + 1) \leqslant CN^{-100}.$$

Thus

$$\mathbb{E}[\|W\|^2] \leqslant (2\sqrt{2p}+1)^2 + \mathbb{E}[\|W\|^2 \mathbb{1}_{\|W\| \geqslant 2\sqrt{2p}+1}] \leqslant (2\sqrt{2p}+1)^2 + \sqrt{\mathbb{E}[\|W\|^4]\mathbb{P}(\|W\| \geqslant 2\sqrt{2p}+1)}$$

and the last term is $o(1)$ provided $\mathbb{E}[\|W\|^4]$ satisfies some weak bound: Since the entries W_{ij} are centered Gaussian with variance at most $\frac{p}{N}$ by Lemma 2.3.7, Hölder's inequality gives

$$\mathbb{E}[\|W\|^4] \leqslant \mathbb{E}[\operatorname{Tr}(W^4)] \leqslant \sum_{i,j,k,\ell} \mathbb{E}[W_{ij}W_{jk}W_{k\ell}W_{\ell i}] \leqslant \sum_{i,j,k,\ell} (\mathbb{E}[W_{ij}^4]\mathbb{E}[W_{jk}^4]\mathbb{E}[W_{k\ell}^4]\mathbb{E}[W_{\ell i}^4])^{1/4} \leqslant 3p^2N^2,$$

which is sufficient. □

Lemma 2.3.9. *There exists C such that, for every $t > 0$, we have*

$$\mathbb{P}(\|H_N\| \geqslant t) \leqslant e^{-\sqrt{N}\max(t-C,0)}.$$

Proof. For definiteness, we consider the logarithmic Sobolev inequality from Lemma 2.3.6 with constant $cN^{-1/2}$, $c = c_{1/2}$. We apply Herbst's lemma with the map $H_N \mapsto \|H_N\|$, which is Lipschitz with constant $\sqrt{2}$ (by the Hoffman-Wielandt inequality), to obtain for any $\alpha > 0$

$$\mathbb{E}[e^{\alpha\|H_N\|}] \leqslant e^{\alpha \sup_N \mathbb{E}[\|H_N\|] + \frac{c}{2}N^{-1/2}\alpha^2}.$$

To finish, we bound $\mathbb{E}\|H_N\|$ with Lemma 2.3.8, choose $\alpha = \sqrt{N}$, and apply Markov's inequality, so that the result applies for any $C > \sup_N \mathbb{E}[\|H_N\|] + c/2$. □

Proof of Corollary 2.1.8.B. By the Herbst argument, Lemma 2.3.6 implies assumption (L) on Lipschitz concentration.

We now check Assumption (W), with the measures μ_N given as the solutions of the Matrix Dyson Equation. Most of this argument consists of importing results of Ajanki *et al.* and Erdős *et al.* Indeed, combining [5, Proposition 2.1, Equation (4.2)], we find that the supports of the measures μ_N satisfy

$$\operatorname{supp}(\mu_N) \subseteq (-(\|A_N\| + 2\sqrt{2p}), \|A_N\| + 2\sqrt{2p}). \tag{2.3.11}$$

Since the right-hand side is uniformly bounded in N, so is the left-hand side. Furthermore, [5, Proposition 2.2] shows that each μ_N admits a density μ_N with respect to Lebesgue measure (on all of $\mathbb{R}$), and that these densities are c-Hölder continuous for some universal c; hence they are bounded, uniformly in N.

To check (2.1.9), we use Proposition 2.3.1. Write s_N for the (random) Stieltjes transform of $\hat{\mu}_{H_N}$. For the Stieltjes-transform estimate (2.3.1), we use the local law [75, Theorem 2.1(4b)], which implies that there exists a universal constant c such that, for every sufficiently small $\varepsilon > 0$, there exists $C_\varepsilon > 0$ with

$$\mathbb{P}\Big(|s_N(E + \mathrm{i}N^{-c\varepsilon}) - m_N(E + \mathrm{i}N^{-c\varepsilon})| \geqslant N^{\varepsilon(1+2c)-1} \text{ for some } |E| \leqslant N^{100}\Big) \leqslant C_\varepsilon N^{-100}.$$

Using the trivial bound $\frac{1}{\eta}$ for a Stieltjes transform evaluated at $E + \mathrm{i}\eta$, we obtain

$$\left|\mathbb{E}[s_N(E + \mathrm{i}N^{-c\varepsilon})] - m_N(E + \mathrm{i}N^{-c\varepsilon})\right| \leqslant N^{\varepsilon(1+2c)-1} + 2C_\varepsilon N^{c\varepsilon-100}$$

for all $|E| \leqslant N^{100}$, which suffices to check (2.3.1). Moreover, if $x > \max \operatorname{supp}(\mu_N)$ we have

$$\left|F_{\mathbb{E}[\hat{\mu}]}(x) - F_{\mu_N}(x)\right| = 1 - F_{\mathbb{E}[\hat{\mu}]}(x) \leqslant \mathbb{P}(\|H_N\| > x) \leqslant e^{-\sqrt{N}\max(t-C,0)}$$

from Lemma 2.3.9, and similarly for the left edge, which gives (2.3.2) and (2.3.3). This verifies assumption (W).

Finally we check the Wegner estimate, with the general Schur-complement strategy. Recall we wrote $\mathcal{C}$ for the covariance matrix of the upper triangle of $H = H_N$ (we will drop the subscript N for the remainder of this proof). Now we will write $\mathcal{C}_{\widehat{jj}}$ for its minor obtained by erasing the column and row corresponding to H_{jj}. Since $\mathcal{C}$ is invertible by Lemma 2.3.5 (and positive semidefinite), so is its minor $\mathcal{C}_{\widehat{jj}}$ by interlacing. Conditioned on $H_{\widehat{jj}}$, we have that H_{jj} is a Gaussian random variable with (an explicit mean that does not matter now and) variance

$$\begin{aligned}(\widetilde{\sigma_{jj}})^2 :&= \operatorname{Var}(H_{jj}) - \sum_{\substack{k\leqslant\ell,k'\leqslant\ell' \\ (k,\ell)\neq(j,j)\neq(k',\ell')}} \mathcal{C}_{(j,j),(k,\ell)}((\mathcal{C}_{\widehat{jj}})^{-1})_{(k,\ell),(k',\ell')}\mathcal{C}_{(k',\ell'),(j,j)} \\ &= \frac{1}{(\mathcal{C}^{-1})_{jj}} \geqslant \lambda_{\min}(\mathcal{C}) \geqslant N^{-1-p},\end{aligned}$$

where we used Lemma 2.3.5 in the last step. By Lemma 2.3.3 and Proposition 2.3.2, this proves (2.1.5). □

2.3.8 Block-diagonal Gaussian matrices. As in subsection 2.3.7, we will use Theorem 2.1.2 (concentrated input). Considered as a vector, the upper triangle of H_N satisfies log-Sobolev with constant $\frac{p}{N}$, since it consists of independent (possibly degenerate) Gaussians with variance at most $\frac{p}{N}$. This implies the Lipschitz-concentration assumption (L).

Now we check assumption (W). We assumed in (R) that the MDE measures μ_N have a bounded density; they lie in a common compact set by the estimate [6, (3.32a)] and arguments like those around (2.3.11), so it remains only to check (2.1.9), through Proposition 2.3.1. If s_N denotes the Stieltjes transform of H_N, then the local law [6, (B.5)] implies that there exist universal constants $\delta > 0$ and $P \in \mathbb{N}$ such that, for every $0 < \gamma < \delta$, there exists C_γ with

$$\mathbb{P}\left(\left| s_N(E + iN^{-\gamma}) - m_N(E + iN^{-\gamma}) \right| \geqslant \frac{N^{\gamma P}}{N} \text{ for some } E \in \mathbb{R} \right) \leqslant C_\gamma N^{-100}.$$

For the tail estimate (2.3.2), we essentially mimic the proof in the case of Gaussian matrices with a (co)variance profile, with the following differences: Here the estimate $\sup_N \mathbb{E}[\|W_N\|^2] < \infty$ is easier, since (recall that W_N is block-diagonal with blocks $X_1, \dots, X_K$) we have $\mathbb{E}[\|W_N\|^2]^{1/2} \leqslant \sum_{i=1}^K \mathbb{E}[\|X_i\|^2]^{1/2}$, and it is classical that $\sup_N \mathbb{E}[\|X_i\|^2] < \infty$ since X_i is a Gaussian matrix whose entries all have variance order $\frac{1}{N}$, by assumption (MF). Since the log-Sobolev constant is now at most p/N, we obtain $\mathbb{P}(\|H_N\| \geqslant t) \leqslant e^{-cN \max(0, t-C)}$ for some constants $c, C > 0$, which verifies (2.3.2) and (2.3.3). This completes the proof of (2.1.9).

Finally we check the Wegner estimate (2.1.5) with Proposition 2.3.2. Here Lemma 2.3.3 applies immediately, since the conditioning is trivial, and we assumed in (MF) that the variances on the diagonal are all at least of order $\frac{1}{N}$.

2.3.9 Free addition. We will use Theorem 2.1.2 (concentrated input). Write $H_N = E + A_N + O_N B_N O_N^T$. Concentration for Lipschitz test functions follows from classical results of Gromov-Milman: If $S = E + \sup_{N \geqslant 1}(\|A_N\| + \|B_N\|)$ and $f : \mathbb{R} \to \mathbb{R}$ is Lipschitz, then (see, e.g., [7, Corollary 4.4.30])

$$\mathbb{P}\left(\left| \frac{1}{N} \operatorname{Tr}(f(H_N)) - \frac{1}{N} \mathbb{E}[\operatorname{Tr}(f(H_N))] \right| \geqslant \delta \right) \leqslant 2 \exp\left(-\frac{\delta^2 N^2}{128 S^2 \|f\|_{\mathrm{Lip}}^2} \right),$$

which suffices to check (2.1.10) and thus assumption (L).

For assumption (E) with the reference measure $\mu_N \equiv \mu_A \boxplus \mu_B$, we will use the local law of Bao

et al, [22, Corollary 2.8]: for every $\varepsilon > 0$ and all $N \geqslant N_0(\varepsilon)$, we have

$$\mathbb{P}\Big(d_{\mathrm{KS}}(\hat{\mu}_{H_N}, \mu_A \boxplus \mu_B) > N^{-1+\varepsilon}\Big) \leqslant N^{-100}.$$

This implies $d_{\mathrm{KS}}(\mathbb{E}[\hat{\mu}_{H_N}], \mu_A \boxplus \mu_B) \lesssim N^{-1+\varepsilon}$. We obtain the same estimate for W_1 as in the proof of Proposition 2.3.1 (there are no tail estimates because all the measures $\hat{\mu}_{H_N}$ and $\mu_A \boxplus \mu_B$ are supported on a common compact set).

It remains only to check the Wegner estimate (2.1.5). The argument is different depending if E is in the bulk of $\mu_A \boxplus \mu_B$ (meaning in the interior of the single-interval support), or if E is outside the support. In the first case, we prove the Wegner estimate with the much stronger fixed-energy universality results of Che-Landon [66, Theorem 2.1]. This result implies

$$\lim_{N\to\infty} \mathbb{P}\bigg(H_N \text{ has no eigenvalues in } \bigg(E - \frac{\varepsilon}{N(\mu_A \boxplus \mu_B)(E)}, E + \frac{\varepsilon}{N(\mu_A \boxplus \mu_B)(E)}\bigg)\bigg) = 1 - F(\varepsilon),$$

where $F(\varepsilon)$ is a special function found by solving the Painlevé V equation satisfying $\lim_{\varepsilon\downarrow 0} F(\varepsilon) = 0$. Thus

$$\liminf_{N\to\infty} \mathbb{P}\bigg(H_N \text{ has no eigenvalues in } \bigg(E - \frac{1}{N^2}, E + \frac{1}{N^2}\bigg)\bigg) \geqslant 1 - \limsup_{\varepsilon\downarrow 0} F(\varepsilon) = 1.$$

In the second case (if E is outside the support of $\mu_A \boxplus \mu_B$), the Wegner estimate is much easier, since indeed $\mathbb{P}(\text{no eigenvalues in } (E-\delta, E+\delta)) \to 1$ for small enough δ. This follows, e.g., from the large-deviations principle for the extremal eigenvalues of this model established by Guionnet and Maïda [102], or from the edge rigidity of Bao *et al.* [22].

2.3.10 Proofs of examples showing necessity of assumptions. In this subsection we show the importance of two of the tricker assumptions of Theorem 2.1.1. Precisely, for each of (2.1.4) and Assumption (S), we give an explicit example satisfying all the assumptions of that theorem except for the one in question, for which the conclusion fails. All notations refer back to that section.

Our example where (2.1.4) fails and determinant concentration fails is the following: Let $(X_{ij})_{1\leqslant i\leqslant j\leqslant N}$ be centered, i.i.d. with variance 1 and a compactly supported and bounded density; choose some $\theta \in (0,1)$ (e.g. $\theta = 1/8$ works) and let A be deterministic, diagonal, and defined through $A_{ii} = e^{N^\theta}\mathbb{1}_{i<N^{1-\theta}}$ with all other entries zero; and define symmetric $H = \Phi(X)$ as $H_{ij} = \Phi(X)_{ij} = \frac{X_{ij}}{\sqrt{N}} + A_{ij}$ for $i \leqslant j$. In this example, $\mu_N = \rho_{\mathrm{sc}}$.

Our example where Assumption (S) fails and determinant concentration fails is the following: Let $(X_{ij})_{1\leqslant i\leqslant j\leqslant N}$ be as above, include the additional random variable X_0 with $\mathbb{P}(X_0 = N) = N^{-1} = 1 - \mathbb{P}(X_0 = 0)$, and define $A = X_0 \operatorname{Id}_N$; then we let $H = \Phi(X)$ be symmetric defined by $H_{ij} = \Phi(X)_{ij} = \frac{X_{ij}}{\sqrt{N}} + A_{ij}$ for $i \leqslant j$. In this example, $\mu_N = \rho_{\mathrm{sc}}$.

In the remainder of this subsection, we prove that these examples have the claimed properties.

2.3.10.1 Necessity of bounds on large eigenvalues.

Write the compact support of the X_i's as $[-T, T]$. This proof is essentially an application of the Weyl inequalities. Note that $\|\Phi\|_{\mathrm{Lip}} = N^{-1/2}$; since the X_i's are compactly supported, this means $X_{\mathrm{cut}} = X$ for $\kappa < 1/2$ and N large enough, and hence (S) is trivially satisfied. Equation (2.1.5) holds by Lemma 2.3.3. If $\kappa < \theta$, then (E) holds with $\mu_N = \rho_{\mathrm{sc}}$ by interlacing; indeed, defining the matrix G by $G_{ij} = \frac{X_{ij}}{\sqrt{N}}$, we have $d_{\mathrm{KS}}(\hat{\mu}_G, \hat{\mu}_H) \leqslant \frac{1}{N}\operatorname{rank}(A) = N^{-\theta}$. Since G is a Wigner matrix with all moments finite, [19, Theorem 4.1] shows $d_{\mathrm{KS}}(\mathbb{E}[\hat{\mu}_G], \rho_{\mathrm{sc}}) \leqslant N^{-1/4}$, and thus if $\theta < 1/4$ we have

$$d_{\mathrm{KS}}(\mathbb{E}[\hat{\mu}_H], \rho_{\mathrm{sc}}) \lesssim N^{-\theta}.$$

We transfer this from d_{KS} to d_{BL} in the same way as for Wigner matrices, above. For (2.1.6), the Weyl inequalities give deterministically

$$\begin{aligned}|\det(H_N)| &= \prod_{i=1}^{N}|\lambda_i(H_N)| \leqslant \prod_{i=1}^{N}(\lambda_i(A) + T\sqrt{N}) \\ &= (e^{N^\theta} + T\sqrt{N})^{N^{1-\theta}}(T\sqrt{N})^{N-N^{1-\theta}} \leqslant (2e^{N^\theta})^{N^{1-\theta}}(T\sqrt{N})^N\end{aligned}$$

which suffices to check (2.1.6) (with any $\delta > 0$).

On the other hand, (2.1.4) fails. Indeed, by the Weyl inequalities the $N^{1-\theta}$ large eigenvalues of H satisfy

$$\lambda_i \geqslant e^{N^\theta} - T\sqrt{N} \geqslant \frac{1}{2} e^{N^\theta}, \tag{2.3.12}$$

so for $\varepsilon < \theta$ the failure of (2.1.4) follows from the deterministic estimate

$$\prod_{i=1}^{N} (1 + |\lambda_i| \mathbb{1}_{|\lambda_i| > e^{N^\varepsilon}}) \geqslant \prod_{i=N-N^{1-\theta}+1}^{N} |\lambda_i| \mathbb{1}_{|\lambda_i| > e^{N^\varepsilon}} \geqslant \left(\frac{1}{2} e^{N^\theta}\right)^{N^{1-\theta}} = 2^{-N^{1-\theta}} e^N.$$

The proof that determinant concentration fails is somewhat involved, but mimics the proof of the lower bound of Theorem 2.1.1. The idea is that the largest $N^{1-\theta}$ eigenvalues contribute a factor of size e^N, as above, and the rest of the eigenvalues behave as if semicircular (this is the difficulty), so we get a lower bound for the determinant asymptotics that is order-one above what the semicircle would predict. We now sketch how to prove this rigorously. Since $X = X_{\text{cut}}$, we simplify our notation and write $\hat{\mu} = \hat{\mu}_{\Phi(X)}$. Recall our eigenvalues are ordered $\lambda_1 \leqslant \cdots \leqslant \lambda_N$; we decompose this measure as

$$\hat{\mu} = \hat{\mu}^{\text{trunc}} + \hat{\mu}^{\text{r. tail}}, \quad \hat{\mu}^{\text{trunc}} = \frac{1}{N} \sum_{i=1}^{N-N^{1-\theta}} \delta_{\lambda_i}, \quad \hat{\mu}^{\text{r. tail}} = \frac{1}{N} \sum_{i=N-N^{1-\theta}+1}^{N} \delta_{\lambda_i}.$$

Notice that $\hat{\mu}^{\text{trunc}}$ has mass $1 - N^{-\theta}$ and $\hat{\mu}^{\text{r. tail}}$ has mass $N^{-\theta}$. Compared to (2.2.2), the event $\mathcal{E}_{\text{ss}}$ is no longer necessary; the events $\mathcal{E}_{\text{gap}}$ and $\mathcal{E}_b$ remain the same (since they clearly imply the analogues for $\hat{\mu}^{\text{trunc}}$), and each still has probability $1 - o(1)$; the event $\mathcal{E}_{\text{conc}}$ is replaced with

$$\mathcal{E}_{\text{conc}}^{\text{trunc}} = \left\{ \left| \int \log_\eta^K \mathrm{d}(\hat{\mu}^{\text{trunc}} - \mathbb{E}[\hat{\mu}^{\text{trunc}}]) \right| \leqslant t \right\}.$$

This is a likely event, since

$$\left| \int \log_\eta^K \mathrm{d}(\hat{\mu}^{\text{r. tail}} - \mathbb{E}[\hat{\mu}^{\text{r. tail}}]) \right| \leqslant 2 \log_\eta(K) \hat{\mu}^{\text{r. tail}}(\mathbb{R}) \lesssim N^{\varepsilon - \theta} < \frac{t}{2} \tag{2.3.13}$$

(here it matters that θ not be too small), and thus if $\varepsilon < \theta$

$$1 - \mathbb{P}(\mathcal{E}_{\text{conc}}^{\text{trunc}}) \leqslant \mathbb{P}\Big(\Big|\int \log_\eta^K \mathrm{d}(\hat{\mu} - \mathbb{E}[\hat{\mu}])\Big| \geqslant \frac{t}{2}\Big),$$

but the right-hand probability is $o(1)$ by arguments as in the proof of Lemma 2.2.3. By mimicking (2.2.6) but handling the large eigenvalues instead with (2.3.12), $\frac{1}{N}\log \mathbb{E}[|\det(H_N)|]$ is larger than

$$\begin{aligned} &1 - \frac{\log 2}{N^\theta} + \frac{1}{N}\mathbb{E}\Big[e^{N\int(\log|\cdot| - \log_\eta)\mathrm{d}\hat{\mu}^{\text{trunc}}}\mathbb{1}_{\mathcal{E}_{\text{gap}}}\mathbb{1}_{\mathcal{E}_{\text{conc}}^{\text{trunc}}}\Big] - t + \int \log_\eta^K \mathrm{d}\mathbb{E}[\hat{\mu}^{\text{trunc}}] \\ &= 1 + \int \log_\eta^K(\lambda)\mathbb{E}[\hat{\mu}^{\text{trunc}}](\mathrm{d}\lambda) - o(1), \end{aligned}$$

where the last equality follows by mimicking Lemma 2.2.5. Now $\mathbb{E}[\hat{\mu}^{\text{trunc}}] = \mathbb{E}[\hat{\mu}] - \mathbb{E}[\hat{\mu}^{\text{r. tail}}]$, and by (2.2.7) and arguments as in the proof of Lemma 2.2.2, we have $\int \log_\eta^K(\lambda)\mathbb{E}[\hat{\mu}](\mathrm{d}\lambda) \geqslant \int \log|\lambda|\rho_{\text{sc}}(\mathrm{d}\lambda) + o(1)$. The term $\int \log_\eta^K(\lambda)\mathbb{E}[\hat{\mu}^{\text{r. tail}}]$ is handled as in (2.3.13). Overall, this gives $\liminf_{N\to\infty} \frac{1}{N}\log \mathbb{E}[|\det(H_N)|] \geqslant 1 + \int \log|\lambda|\rho_{\text{sc}}(\mathrm{d}\lambda)$ which contradicts (2.1.8).

2.3.10.2 Necessity of spectral stability. With T as above, the eigenvalues of H are at most $N + T\sqrt{N}$ deterministically; this implies (2.1.4) and (2.1.6). For (2.1.5), we note that on the event $\{X_0 = N\}$, the eigenvalues are at least $N - T\sqrt{N} > 0$, so there are clearly no eigenvalues near zero; on the event $\{X_0 = 0\}$, the matrix H is just a Wigner matrix, for which we proved (2.1.5) above. Now we claim that assumption (E) holds with $\mu_N = \rho_{\text{sc}}$. Indeed, for test functions f with $\|f\|_{\text{Lip}} + \|f\|_{L^\infty} \leqslant 1$ we have

$$\mathbb{E}\Big[\Big|\int f(x)(\hat{\mu}_H - \rho_{\text{sc}})(\mathrm{d}x)\Big|\mathbb{1}_{X_0=N}\Big] \leqslant 2\mathbb{P}(X_0 = N) = \frac{2}{N},$$

and on the event $\mathbb{1}_{X_0=0}$ we revert to the Wigner case studied above.

On the other hand, notice that $(X_{\text{cut}})_0$ is always zero, so on the event $\{X_0 = N\}$ the measure $\hat{\mu}_{\Phi(X_{\text{cut}})}$ is supported on $[-T\sqrt{N}, T\sqrt{N}]$ while $\hat{\mu}_{\Phi(X)}$ is supported on $[N - T\sqrt{N}, N + T\sqrt{N}]$. For large enough N these are disjoint, so the measures are one apart in KS distance, and thus

$\mathbb{P}(d_{\mathrm{KS}}(\hat{\mu}_{\Phi(X)}, \hat{\mu}_{\Phi(X_{\mathrm{cut}})}) > N^{-\kappa}) \geqslant \frac{1}{N}$, which shows that (S) fails.

Finally, since

$$\mathbb{E}[|\det(H)|] \geqslant \mathbb{E}[|\det(H)|\mathbb{1}_{X_0=N}] \geqslant (N - T\sqrt{N})^N \mathbb{P}(X_0 = N),$$

we have $\frac{1}{N} \log \mathbb{E}[|\det(H)|] \to +\infty$ and determinant concentration fails.

2.4 Variational principles and long-range correlations

2.4.1 General scheme. In this section, we study expected determinants in the presence of long-range matrix correlations. The prototypical example to keep in mind is

$$H_N = W_N + \xi \,\mathrm{Id},$$

where W_N is drawn from the Gaussian Orthogonal Ensemble (GOE), and $\xi \sim \mathcal{N}(0, 1/N)$ is independent of W_N. Matrices of this style are very common in the landscape-complexity program, but our main theorems do not apply directly because of the presence of long-range correlations (here, along the diagonal of H_N). Nevertheless, there is still a general procedure to understand the determinant asymptotics for such matrices, which we illustrate in the case of this example. We first notice

$$\mathbb{E}[|\det(H_N)|] = \frac{1}{\sqrt{2\pi/N}} \int_{\mathbb{R}} e^{-N\frac{u^2}{2}} \mathbb{E}[|\det(W_N + u)|] \,\mathrm{d}u.$$

Our determinant asymptotics do apply to $W_N + u$, giving $\mathbb{E}[|\det(W_N + u)|] \approx e^{N\Sigma(u)}$ for some constants $\Sigma(u)$; then the Laplace method suggests

$$\lim_{N\to\infty} \frac{1}{N} \log \mathbb{E}[|\det(H_N)|] = \sup_{u\in\mathbb{R}} \left\{ \Sigma(u) - \frac{u^2}{2} \right\}. \tag{2.4.1}$$

This method has appeared before in special cases, for example in [11] and [88]. In Section 2.4.2, we prove results of this type without reference to any particular matrix model. In Section 2.4.3, we prove extensions necessary to understand asymptotics of the form

$$\lim_{N\to\infty} \frac{1}{N} \log \mathbb{E}[|\det(H_N)| \mathbb{1}_{H_N \geqslant 0}].$$

In complexity computations, these "restricted determinants" correspond to counting just the local minima among all critical points. The upshot is that this limit is also a variational problem as in (2.4.1), but restricted to u in some good set instead of all Euclidean space.

2.4.2 Variational principles for unrestricted determinants. For applications to complexity, we will need not just one matrix H_N, but a field of matrices $H_N(u)$ for $u \in \mathbb{R}^m$ (here m is independent of N), with approximating measures $\mu_N(u)$.

Theorem 2.4.1. *Assume the following:*

- ***(Assumptions locally uniform in** u**)** Each $H_N(u)$ satisfies all the assumptions of Theorem 2.1.1, or all the assumptions of Theorem 2.1.2. In addition, all limits, powers, and rates in these assumptions are uniform over compact sets of u.*[3]*
- ***(Limit measures)** There exist probability measures $\mu_\infty(u)$ such that*

$$d_{\mathrm{BL}}(\mu_N(u), \mu_\infty(u)) \leqslant N^{-\kappa} \qquad \text{if we are in the setting of Theorem 2.1.1, or}$$
$$\mathrm{W}_1(\mu_N(u), \mu_\infty(u)) \leqslant N^{-\kappa} \qquad \text{if we are in the setting of Theorem 2.1.2}$$

  for $\kappa = \kappa(u) > 0$ that can, again, be chosen uniformly on compact sets of u. These measures also admit densities $\mu_\infty(u, \cdot)$ on $[-\kappa, \kappa]$ that satisfy $\mu_\infty(u, x) < \kappa^{-1}|x|^{-1+\kappa}$ for all $|x| < \kappa$.
- ***(Continuity and decay in** u**)** For each N, the map $u \mapsto H_N(u)$ is entrywise continuous.*

[3] For example, writing $(\lambda_i(u))_{i=1}^N$ for the eigenvalues of $H_N(u)$, the condition (2.1.4) becomes: for every compact $K \subset \mathbb{R}^m$, $\lim_{N\to\infty} \sup_{u\in K} \frac{1}{N} \log \mathbb{E}\big[\prod_{i=1}^N (1 + |\lambda_i(u)| \mathbb{1}_{|\lambda_i(u)|>e^{N^\varepsilon}})\big] = 0$.

Furthermore, there exists $C > 0$ such that

$$\mathbb{E}[|\det(H_N(u))|] \leqslant (C \max(\|u\|, 1))^N. \tag{2.4.2}$$

Then for any $\alpha > 0$, any fixed $p \in \mathbb{N}$, and any $\mathfrak{D} \subset \mathbb{R}^m$ with positive Lebesgue measure that is the closure of its interior, we have

$$\lim_{N\to\infty} \frac{1}{N} \log \int_{\mathfrak{D}} e^{-(N+p)\alpha\|u\|^2} \mathbb{E}[|\det(H_N(u))|] \, \mathrm{d}u = \sup_{u\in\mathfrak{D}} \left\{ \int_{\mathbb{R}} \log|\lambda| \mu_\infty(u)(\mathrm{d}\lambda) - \alpha\|u\|^2 \right\}.$$

Remark 2.4.2. *A close inspection of the proof shows that the condition "$\mathfrak{D}$ is the closure of its interior" is only necessary for the lower bound in Theorem 2.4.1. For the upper bound, it suffices to assume that $\mathfrak{D}$ is simply closed (and has positive measure). We will use this below.*

The proof of this theorem relies on the following two lemmas, in addition to determinant concentration in the form of Theorem 2.1.1 or 2.1.2. We postpone their proofs until after the proof of the theorem. Recall that B_R is the ball of radius R around zero in $\mathbb{R}^m$.

Lemma 2.4.3.

$$\lim_{R\to\infty} \limsup_{N\to\infty} \frac{1}{N} \log \int_{B_R^c} e^{-N\alpha\|u\|^2} \mathbb{E}[|\det(H_N(u))|] \, \mathrm{d}u = -\infty.$$

Lemma 2.4.4. *The function*

$$\mathcal{S}_\alpha[u] = \int_{\mathbb{R}} \log|\lambda| \mu_\infty(u)(\mathrm{d}\lambda) - \alpha\|u\|^2,$$

is continuous, and $\lim_{\|u\|\to+\infty} \mathcal{S}_\alpha[u] = -\infty$.

Proof of Theorem 2.4.1. First we prove the upper bound. We apply Theorem 2.1.1 or 2.1.2 with the reference measures $\mu_\infty(u)$. Since all inputs are uniform over compact sets of u, so is the conclusion;

that is, for all R, we have

$$\limsup_{N\to\infty}\frac{1}{N}\log\sup_{u\in B_R}\left\{\mathbb{E}[|\det(H_N(u))|]e^{-N\int_{\mathbb{R}}\log|\lambda|\mu_\infty(u)(\mathrm{d}\lambda)}\right\}\leqslant 0$$

and a matching lower bound we will use momentarily. If R is large enough that $|B_R\cap\mathfrak{D}|>0$, then we conclude

$$\begin{aligned}
&\limsup_{N\to\infty}\frac{1}{N}\log\int_{B_R\cap\mathfrak{D}}e^{-(N+p)\alpha\|u\|^2}\mathbb{E}[|\det(H_N(u))|]\,\mathrm{d}u\\
&\leqslant\limsup_{N\to\infty}\frac{1}{N}\log\int_{B_R\cap\mathfrak{D}}e^{-N\alpha\|u\|^2+N\int_{\mathbb{R}}\log|\lambda|\mu_\infty(u,\lambda)\,\mathrm{d}\lambda}\,\mathrm{d}u\\
&\leqslant\sup_{u\in\mathfrak{D}}\mathcal{S}_\alpha[u]+\limsup_{N\to\infty}\left[\frac{\log(|B_R\cap\mathfrak{D})|}{N}\right].
\end{aligned}$$

An application of Lemma 2.4.3 finishes the proof of the upper bound.

Now we prove the lower bound. Lemma 2.4.4 tells us that $\sup_{u\in\mathfrak{D}}\mathcal{S}_\alpha[u]$ is achieved at some (possibly not unique) u_0. Since $\mathcal{S}_\alpha$ is continuous, for every $\varepsilon>0$ there exists a bounded neighborhood $\mathcal{U}_\varepsilon$ of u_0 on which $\mathcal{S}_\alpha[u]\geqslant\mathcal{S}_\alpha[u_0]-\varepsilon$. Since $\mathfrak{D}$ is the closure of its interior, we have $|\mathcal{U}_\varepsilon\cap\mathfrak{D}|>0$.

For each R, applying Theorem 2.1.1 or 2.1.2 with arguments as above yields

$$\liminf_{N\to\infty}\frac{1}{N}\log\inf_{u\in B_R}\left\{\mathbb{E}[|\det(H_N(u))|]e^{-N\int_{\mathbb{R}}\log|\lambda|\mu_\infty(u)(\mathrm{d}\lambda)}\right\}\geqslant 0.$$

If R is so large that $\mathcal{U}_\varepsilon\subset B_R$, then

$$\begin{aligned}
&\liminf_{N\to\infty}\frac{1}{N}\log\int_{\mathfrak{D}}e^{-(N+p)\alpha\|u\|^2}\mathbb{E}[|\det(H_N(u))|]\,\mathrm{d}u\\
&\geqslant\liminf_{N\to\infty}\frac{1}{N}\log\left\{e^{-p\alpha R^2}\int_{\mathcal{U}_\varepsilon\cap\mathfrak{D}}e^{-N\alpha\|u\|^2}\mathbb{E}[|\det(H_N(u))|]\,\mathrm{d}u\right\}\\
&\geqslant\liminf_{N\to\infty}\frac{1}{N}\log\int_{\mathcal{U}_\varepsilon\cap\mathfrak{D}}e^{N\mathcal{S}_\alpha[u]}\,\mathrm{d}u\geqslant\liminf_{N\to\infty}\frac{1}{N}\log\int_{\mathcal{U}_\varepsilon\cap\mathfrak{D}}e^{N(\mathcal{S}_\alpha[u_0]-\varepsilon)}\,\mathrm{d}u\\
&\geqslant\mathcal{S}_\alpha[u_0]-\varepsilon+\liminf_{N\to\infty}\frac{\log(|\mathcal{U}_\varepsilon\cap\mathfrak{D}|)}{N}.
\end{aligned}$$

Letting $\varepsilon \to 0$ completes the proof. □

Proof of Lemma 2.4.3. If ω_m is the surface area of the unit ball in $\mathbb{R}^m$, then from (2.4.2) we have

$$\int_{B_R^c} e^{-N\alpha\|u\|^2}\mathbb{E}[|\det(H_N(u))|]\,\mathrm{d}u \leqslant \int_{B_R^c} e^{N\left(\log(C\|u\|)-\alpha\|u\|^2\right)}\,\mathrm{d}u = \omega_m \int_R^\infty e^{N\left(\log(Cr)-\alpha r^2\right)} r^{m-1}\,\mathrm{d}r$$

which suffices by the Laplace method. □

Proof of Lemma 2.4.4. Fix N. We assumed that $H_N(u)$ is an entrywise continuous function of u. Since the determinant is a continuous function of the matrix entries, dominated convergence (with dominating function given by (2.4.2)) says that $\mathbb{E}[|\det(H_N(u))|]$ is continuous in u, hence so is $\frac{1}{N}\log\mathbb{E}[|\det(H_N(u))|]$. Then Theorem 2.1.1 or 2.1.2 shows

$$\lim_{N\to\infty}\sup_{u\in B_R}\left|\frac{1}{N}\log\mathbb{E}[|\det(H_N(u))|] - \int_{\mathbb{R}}\log|\lambda|\mu_\infty(u)(\mathrm{d}\lambda)\right| = 0, \tag{2.4.3}$$

and $\int_{\mathbb{R}}\log|\lambda|\mu_\infty(u)(\mathrm{d}\lambda)$ is the locally uniform limit of continuous functions. Thus $\mathcal{S}_\alpha[u]$ is continuous.

The decay at infinity follows from

$$\int_{\mathbb{R}}\log|\lambda|\mu_\infty(u)(\mathrm{d}\lambda) \leqslant \liminf_{N\to\infty}\frac{1}{N}\log\mathbb{E}[|\det(H_N(u))|] \leqslant \log(C\max(\|u\|,1)),$$

obtained by (2.4.3) and (2.4.2). □

2.4.3 Variational principles for restricted determinants. Let $\mathcal{G}\subset\mathbb{R}^m$ be the set of "good" u values

$$\mathcal{G} = \{u\in\mathbb{R}^m : \mu_\infty(u)((-\infty,0)) = 0\} = \{u\in\mathbb{R}^m : \mathtt{l}(\mu_\infty(u)) \geqslant 0\}. \tag{2.4.4}$$

For each $\varepsilon > 0$, consider the following inner and outer approximations of $\mathcal{G}$:

$$\begin{aligned} \mathcal{G}_{+\varepsilon} &= \{u \in \mathbb{R}^m : \mathtt{l}(\mu_\infty(u)) \geqslant 2\varepsilon\}, \\ \mathcal{G}_{-\varepsilon} &= \{u \in \mathbb{R}^m : \mu_\infty(u)((-\infty, -\varepsilon)) \leqslant \varepsilon\}. \end{aligned} \tag{2.4.5}$$

Theorem 2.4.5. *Fix some $\mathfrak{D} \subset \mathbb{R}^m$, and suppose that $\mathfrak{D}$ and the matrices $H_N(u)$ satisfy the following.*

- *All the assumptions of Theorem 2.4.1.*
- ***(Superexponential concentration)*** *For every $R > 0$ and every $\varepsilon > 0$, we have*

$$\lim_{N\to\infty} \frac{1}{N \log N} \log \left[\sup_{u \in B_R} \mathbb{P}(d_{\mathrm{BL}}(\hat{\mu}_{H_N(u)}, \mu_\infty(u)) > \varepsilon) \right] = -\infty. \tag{2.4.6}$$

- ***(No outliers)*** *For every $R > 0$ and every $\varepsilon > 0$, we have*

$$\lim_{N\to\infty} \inf_{u \in \mathfrak{D} \cap \mathcal{G}_{+\varepsilon} \cap B_R} \mathbb{P}(\mathrm{Spec}(H_N(u)) \subset [\mathtt{l}(\mu_\infty(u)) - \varepsilon, \mathtt{r}(\mu_\infty(u)) + \varepsilon]) = 1. \tag{2.4.7}$$

- ***(Topology)*** *Each $\mathcal{G}_{+\varepsilon}$ is convex; $\mathfrak{D}$ is convex and closed; the set $\mathfrak{D} \cap \mathcal{G}_{+1}$ has positive Lebesgue measure; and*

$$\overline{\mathfrak{D} \cap \left(\bigcup_{\varepsilon > 0} \mathcal{G}_{+\varepsilon} \right)} = \mathfrak{D} \cap \mathcal{G}. \tag{2.4.8}$$

Then for any $\alpha > 0$ and any fixed $p \in \mathbb{N}$, we have

$$\lim_{N\to\infty} \frac{1}{N} \log \int_{\mathfrak{D}} e^{-(N+p)\alpha\|u\|^2} \mathbb{E}[|\det(H_N(u))| \mathbb{1}_{H_N(u) \geqslant 0}] = \sup_{u \in \mathfrak{D} \cap \mathcal{G}} \left\{ \int_{\mathbb{R}} \log|\lambda| \mu_\infty(\mathrm{d}\lambda) - \alpha\|u\|^2 \right\}.$$

We prove the upper and lower bounds separately in the next two subsubsections.

2.4.3.1 Upper bound. The proof of the upper bound of Theorem 2.4.5 relies on the following three lemmas, which we will prove after.

Lemma 2.4.6. *Each $\mathcal{G}_{-\varepsilon}$ is closed, and $\mathcal{G}$ is closed.*

Lemma 2.4.7. *For every $\varepsilon > 0$, we have*

$$\lim_{N\to\infty} \frac{1}{N} \log \int_{(\mathcal{G}_{-\varepsilon})^c} e^{-N\alpha\|u\|^2} \mathbb{E}[|\det(H_N(u))| \mathbb{1}_{H_N(u)\geqslant 0}] \, \mathrm{d}u = -\infty.$$

Lemma 2.4.8. *We have*

$$\lim_{\varepsilon\downarrow 0} \sup_{u\in\mathfrak{D}\cap\mathcal{G}_{-\varepsilon}} \mathcal{S}_\alpha[u] \leqslant \sup_{u\in\mathfrak{D}\cap\mathcal{G}} \mathcal{S}_\alpha[u].$$

Proof of the upper bound in Theorem 2.4.5. For each $\varepsilon > 0$, Lemma 2.4.7 yields

$$\begin{aligned}
&\limsup_{N\to\infty} \frac{1}{N} \log \int_{\mathfrak{D}} e^{-(N+p)\alpha\|u\|^2} \mathbb{E}[|\det(H_N(u)| \mathbb{1}_{H_N(u)\geqslant 0}] \, \mathrm{d}u \\
&\leqslant \limsup_{N\to\infty} \frac{1}{N} \log \int_{\mathfrak{D}\cap\mathcal{G}_{-\varepsilon}} e^{-N\alpha\|u\|^2} \mathbb{E}[|\det(H_N(u))| \mathbb{1}_{H_N(u)\geqslant 0}] \, \mathrm{d}u \\
&\leqslant \limsup_{N\to\infty} \frac{1}{N} \log \int_{\mathfrak{D}\cap\mathcal{G}_{-\varepsilon}} e^{-N\alpha\|u\|^2} \mathbb{E}[|\det(H_N(u))|] \, \mathrm{d}u \leqslant \sup_{u\in\mathfrak{D}\cap\mathcal{G}_{-\varepsilon}} \mathcal{S}_\alpha[u].
\end{aligned}$$

The last inequality holds by Theorem 2.4.1 applied to $\mathfrak{D}\cap\mathcal{G}_{-\varepsilon}$, which is closed (by Lemma 2.4.6) and has positive measure (as a superset of $\mathfrak{D}\cap\mathcal{G}_{+1}$, which has positive measure by assumption). By Remark 2.4.2, these are the only conditions we need to check. Letting ε tend to zero and applying Lemma 2.4.8 completes the proof. □

Proof of Lemma 2.4.6. Since we assumed that $u \mapsto H_N(u)$ is entrywise continuous and the spectrum is a continuous function of matrix entries, we have that $u \mapsto \hat{\mu}_{H_N(u)}$ is almost surely continuous with respect to the bounded-Lipschitz distance:

$$d_{\mathrm{BL}}(\hat{\mu}_{H_N(u)}, \hat{\mu}_{H_N(u')}) \leqslant \frac{1}{N} \sum_{i=1}^{N} \min(2, |\lambda_i(u) - \lambda_i(u')|).$$

By dominated convergence, this means that $u \mapsto \mathbb{E}[\hat{\mu}_{H_N(u)}]$ is continuous with respect to d_{BL}. But $d_{\mathrm{BL}}(\mathbb{E}[\hat{\mu}_{H_N(u)}], \mu_\infty(u)) \to 0$ uniformly on compact sets of u by assumption (here we use $d_{\mathrm{BL}} \leqslant \mathrm{W}_1$ for the concentrated-input case), so we conclude that $u \mapsto \mu_\infty(u)$ is continuous with respect to d_{BL},

as well. Since d_{BL} metrizes weak convergence, and since the defining properties of $\mathcal{G}$ and $\mathcal{G}_{-\varepsilon}$ can be stated in terms of distribution functions of $\mu_\infty(u)$, which are continuous since each $\mu_\infty(u)$ has a density with respect to Lebesgue, the lemma follows. □

Proof of Lemma 2.4.7. From Lemma 2.4.3, it suffices to show

$$\lim_{N\to\infty} \frac{1}{N} \log \int_{(\mathcal{G}_{-\varepsilon})^c \cap B_R} e^{-N\alpha\|u\|^2} \mathbb{E}[|\det(H_N(u))| \mathbb{1}_{H_N(u)\geqslant 0}] \, \mathrm{d}u = -\infty$$

for each $R > 0$. If $H_N(u) \geqslant 0$ and $u \in (\mathcal{G}_{-\varepsilon})^c$, then by taking some $\frac{1}{2}$-Lipschitz f_ε satisfying $\frac{\varepsilon}{2}\mathbb{1}_{x\leqslant 0} \geqslant f_\varepsilon(x) \geqslant \frac{\varepsilon}{2}\mathbb{1}_{x\leqslant -\varepsilon}$ we obtain $d_{\mathrm{BL}}(\hat{\mu}_{H_N(u)}, \mu_\infty(u)) \geqslant \frac{\varepsilon}{2}\mu_\infty(u)((-\infty, -\varepsilon)) \geqslant \frac{\varepsilon^2}{2}$. For small $\delta > 0$, this gives

$$\begin{aligned}
&\int_{(\mathcal{G}_{-\varepsilon})^c \cap B_R} e^{-N\alpha\|u\|^2} \mathbb{E}[|\det(H_N(u))| \mathbb{1}_{H_N(u)\geqslant 0}] \, \mathrm{d}u \\
&\leqslant |B_R| \left(\sup_{u\in B_R} \mathbb{E}[|\det(H_N(u))|^{1+\delta}]^{\frac{1}{1+\delta}} \right) \left(\sup_{u\in B_R} \mathbb{P}(d_{\mathrm{BL}}(\hat{\mu}_{H_N(u)}, \mu_\infty(u)) \geqslant \frac{\varepsilon^2}{2}) \right)^{\frac{\delta}{1+\delta}}.
\end{aligned}$$

This suffices by (2.1.6) and (2.4.6). □

Proof of Lemma 2.4.8. From their definitions, we have $\bigcap_{\varepsilon>0} \mathcal{G}_{-\varepsilon} = \mathcal{G}$. We take the intersection of both sides with $\mathfrak{D}$. Next, we claim that there exists some $R > 0$ such that

$$\sup_{u\in\mathfrak{D}\cap\mathcal{G}} \mathcal{S}_\alpha[u] = \max_{u\in(\mathfrak{D}\cap\mathcal{G}\cap B_R)} \mathcal{S}_\alpha[u] \quad \text{and} \quad \sup_{u\in\mathfrak{D}\cap\mathcal{G}_{-\varepsilon}} \mathcal{S}_\alpha[u] = \max_{u\in(\mathfrak{D}\cap\mathcal{G}_{-\varepsilon}\cap B_R)} \mathcal{S}_\alpha[u] \tag{2.4.9}$$

for every $\varepsilon > 0$. Indeed, the proof of Lemma 2.4.4 shows that

$$\mathcal{S}_\alpha[u] \leqslant \log(C\|u\|) - \alpha\|u\|^2$$

on $\mathbb{R}^m$. Since $\mathcal{S}_\alpha$ is continuous and $\mathfrak{D} \cap \mathcal{G}$ is closed by Lemma 2.4.6, let $u_* \in \mathfrak{D} \cap \mathcal{G}$ satisfy $\sup_{u\in\mathfrak{D}\cap\mathcal{G}} \mathcal{S}[u] = \mathcal{S}[u_*]$, and let $R > 1$ be so large that $\log(CR) - \alpha R^2 < \mathcal{S}_\alpha[u_*]$.

For each ε, since $\mathfrak{D} \cap \mathcal{G}_{-\varepsilon}$ is closed (again by Lemma 2.4.6), let u_ε be such that $\mathcal{S}_\alpha[u_\varepsilon] =$

$\sup_{u\in\mathfrak{D}\cap\mathcal{G}_{-\varepsilon}}\mathcal{S}_\alpha[u]$. Then $u_\varepsilon\in B_R$; otherwise, we would have

$$\max_{u\in\mathfrak{D}\cap\mathcal{G}_{-\varepsilon}}\mathcal{S}_\alpha[u]=\mathcal{S}_\alpha[u_\varepsilon]\leqslant\log(CR)-\alpha R^2<\mathcal{S}_\alpha[u_*]=\max_{u\in\mathfrak{D}\cap\mathcal{G}}\mathcal{S}_\alpha[u]\leqslant\max_{u\in\mathfrak{D}\cap\mathcal{G}_{-\varepsilon}}\mathcal{S}_\alpha[u].$$

This verifies (2.4.9).

Since the $\{u_\varepsilon\}$ lie in a compact set, they have a limit point u_0 up to extraction. Furthermore, $u_0\in\mathfrak{D}\cap\mathcal{G}=\cap_\varepsilon(\mathfrak{D}\cap\mathcal{G}_{-\varepsilon})$. Indeed, otherwise a neighborhood of u_0 would be contained in $(\mathfrak{D}\cap\mathcal{G}_{-\varepsilon_1})^c$ for some ε_1, hence in $(\mathfrak{D}\cap\mathcal{G}_{-\varepsilon})^c$ for every $\varepsilon<\varepsilon_1$ (since the sets are nested). But then u_0 could not be a limit point of $\{u_\varepsilon\}$.

Thus by continuity of S_α we have

$$\lim_{\varepsilon\downarrow0}\sup_{u\in\mathfrak{D}\cap\mathcal{G}_{-\varepsilon}}\mathcal{S}_\alpha[u]=\lim_{\varepsilon\downarrow0}\mathcal{S}_\alpha[u_\varepsilon]=\mathcal{S}_\alpha[u_0]\leqslant\sup_{u\in\mathfrak{D}\cap\mathcal{G}}\mathcal{S}_\alpha[u].$$

□

2.4.3.2 Lower bound. The proof of the lower bound in Theorem 2.4.5 relies on the following two lemmas, which we will prove after.

Lemma 2.4.9. *For each $u\in\mathbb{R}^m$ and each $\delta,\varepsilon>0$, define the set of probability measures*

$$M(u,\delta,\varepsilon)=\{\mu:d_{\mathrm{BL}}(\mu,\mu_\infty(u))<\delta\ \textit{and}\ \operatorname{supp}(\mu)\subset[\mathtt{l}(\mu_\infty(u))-\varepsilon,\mathtt{r}(\mu_\infty(u))+\varepsilon]\}$$

that are close to $\mu_\infty(u)$ both in d_{BL} and in support. For all R, all δ, and all ε sufficiently small depending on R, we have

$$\inf_{u\in\mathfrak{D}\cap\mathcal{G}_{+\varepsilon}\cap\overline{B_R}}\left(\inf_{\mu\in M(u,\delta,\varepsilon)}\int\log|\lambda|\mu(\mathrm{d}\lambda)-\int\log|\lambda|\mu_\infty(u)(\mathrm{d}\lambda)\right)\geqslant-\frac{2\delta}{\varepsilon}.$$

Lemma 2.4.10. *Each $\mathcal{G}_{+\varepsilon}$ is closed, and for all R large enough we have*

$$\sup_{u\in\mathfrak{D}\cap\mathcal{G}} \mathcal{S}_\alpha[u] = \max_{u\in\mathfrak{D}\cap\mathcal{G}\cap\overline{B_R}} \mathcal{S}_\alpha[u] \quad \textit{and} \quad \sup_{u\in\mathfrak{D}\cap\mathcal{G}_{+\varepsilon}} \mathcal{S}_\alpha[u] = \max_{u\in\mathfrak{D}\cap\mathcal{G}_{+\varepsilon}\cap\overline{B_R}} \mathcal{S}_\alpha[u] \tag{2.4.10}$$

for every $0<\varepsilon<1$. Furthermore,

$$\lim_{\varepsilon\downarrow 0} \sup_{u\in\mathfrak{D}\cap\mathcal{G}_{+\varepsilon}} \mathcal{S}_\alpha[u] = \sup_{u\in\mathfrak{D}\cap\mathcal{G}} \mathcal{S}_\alpha[u]. \tag{2.4.11}$$

Proof of the lower bound in Theorem 2.4.5. Since

$$\begin{aligned}
&\mathbb{P}(\hat{\mu}_{H_N(u)} \notin M(u,\delta,\varepsilon)) \\
&\leqslant \mathbb{P}(\mathrm{Spec}(H_N(u)) \not\subset [\mathtt{l}(\mu_\infty(u)) - \varepsilon, \mathtt{r}(\mu_\infty(u)) + \varepsilon]) + \mathbb{P}(d_{\mathrm{BL}}(\hat{\mu}_{H_N(u)}, \mu_\infty(u)) > \delta),
\end{aligned}$$

(2.4.7) and (2.4.6) tell us that

$$\lim_{N\to\infty} \frac{1}{N} \log\left(\inf_{u\in\mathfrak{D}\cap\mathcal{G}_{+\varepsilon}\cap\overline{B_R}} \mathbb{P}(\hat{\mu}_{H_N(u)} \in M(u,\delta,\varepsilon)) \right) = 0. \tag{2.4.12}$$

Let R satisfy Lemma 2.4.10, and additionally be so large that $\left|\mathfrak{D}\cap\mathcal{G}_{+1}\cap\overline{B_R}\right| > 0$. If $\varepsilon<1$ is sufficiently small depending on R, then by Lemma 2.4.9 we have

$$\begin{aligned}
&\int_{\mathfrak{D}} e^{-(N+p)\alpha\|u\|^2} \mathbb{E}[|\det(H_N(u))| \mathbb{1}_{H_N(u)\geqslant 0}]\,\mathrm{d}u \\
&\geqslant e^{-p\alpha R^2} \int_{\mathfrak{D}\cap\mathcal{G}_{+\varepsilon}\cap\overline{B_R}} e^{-N\alpha\|u\|^2} \exp\left(N \inf_{\mu\in M(u,\delta,\varepsilon)} \int \log|\lambda|\mu(\mathrm{d}\lambda) \right) \mathbb{P}(\hat{\mu}_{H_N(u)} \in M(u,\delta,\varepsilon))\,\mathrm{d}u \\
&\geqslant e^{-p\alpha R^2} \left(\inf_{u\in\mathfrak{D}\cap\mathcal{G}_{+\varepsilon}\cap\overline{B_R}} \mathbb{P}(\hat{\mu}_{H_N(u)} \in M(u,\delta,\varepsilon)) \right) \exp\left(-\frac{2N\delta}{\varepsilon} \right) \int_{\mathfrak{D}\cap\mathcal{G}_{+\varepsilon}\cap\overline{B_R}} e^{N\mathcal{S}_\alpha[u]}\,\mathrm{d}u.
\end{aligned}$$

Now we take the logarithm of both sides, divide by N, let $N\to\infty$, and then let $\delta\downarrow 0$. The set $\mathfrak{D}\cap\mathcal{G}_{+\varepsilon}\cap\overline{B_R}$ is closed and convex, as the finite intersection of such sets. Since closed convex sets in Euclidean space have empty interior if and only if they lie in a lower-dimensional affine space, we conclude that $\mathfrak{D}\cap\mathcal{G}_{+\varepsilon}\cap\overline{B_R}$ has nonempty interior from the fact that it has positive measure. Since

a closed convex set with nonempty interior is the closure of its interior, we can apply Theorem 2.4.1 to this set. From this theorem and from (2.4.12), we have

$$\liminf_{N\to\infty} \int_{\mathfrak{D}} e^{-(N+p)\alpha\|u\|^2} \mathbb{E}[|\det(H_N(u))| \mathbb{1}_{H_N(u)\geqslant 0}]\, \mathrm{d}u \geqslant \sup_{u\in\mathfrak{D}\cap\mathcal{G}_{+\varepsilon}\cap B_R} \mathcal{S}_\alpha[u] = \sup_{u\in\mathfrak{D}\cap\mathcal{G}_{+\varepsilon}} \mathcal{S}_\alpha[u].$$

By (2.4.11), this suffices. □

Proof of Lemma 2.4.9. Consider the function f_u defined on $[\mathtt{l}(\rho_\infty(u))-\varepsilon, \mathtt{r}(\rho_\infty(u))+\varepsilon]$ by $f_u(\lambda) = \log|\lambda|$. If $u \in \mathfrak{D}\cap\mathcal{G}_{+\varepsilon}\cap\overline{B_R}$, then

$$\|f_u\|_{\mathrm{Lip}} + \|f_u\|_{L^\infty} \leqslant \frac{1}{\varepsilon} + \max\{|\log(\varepsilon)|, |\log(\mathtt{r}(\rho_\infty(u))+\varepsilon)|\} \leqslant \frac{2}{\varepsilon}$$

where the last inequality holds for ε sufficiently small, uniformly over $u \in \overline{B_R}$, since $\mathrm{supp}(\rho_\infty(u))$ is compactly supported uniformly over $u \in \overline{B_R}$. This implies that whenever $\mu \in M(u,\delta,\varepsilon)$, we have $|\int \log|\cdot|\mathrm{d}\mu - \int \log|\cdot|\mathrm{d}\mu_\infty(u,\cdot)| \leqslant \frac{2}{\varepsilon} d_{\mathrm{BL}}(\mu, \mu_\infty(u)) \leqslant \frac{2\delta}{\varepsilon}$. □

Proof of Lemma 2.4.10. The proof of Lemma 2.4.6 shows that the map $u \mapsto \rho_\infty(u)$ is continuous with respect to weak convergence; thus each $\mathcal{G}_{+\varepsilon}$ is closed.

The proof of Lemma 2.4.4 shows that $\mathcal{S}_\alpha[u] \leqslant \log(C\|u\|) - \alpha\|u\|^2$ on $\mathbb{R}^m$ and that $\mathcal{S}_\alpha$ is continuous. Since $\mathfrak{D}\cap\mathcal{G}$ is closed by Lemma 2.4.6, and each $\mathfrak{D}\cap\mathcal{G}_{+\varepsilon}$ is closed by the argument above, we can write $\sup_{u\in\mathfrak{D}\cap\mathcal{G}} \mathcal{S}_\alpha[u] = \mathcal{S}_\alpha[u_*]$ for some u_* and $\sup_{u\in\mathfrak{D}\cap\mathcal{G}_{+\varepsilon}} \mathcal{S}_\alpha[u] = \mathcal{S}_\alpha[u_\varepsilon]$ for some u_ε.

Let $R > 1$ be so large that $\log(CR) - \alpha R^2 < \mathcal{S}_\alpha[u_1]$. Then $u_\varepsilon \in B_R$ for each $\varepsilon < 1$; else we would have

$$\max_{u\in\mathfrak{D}\cap\mathcal{G}_{+\varepsilon}} \mathcal{S}_\alpha[u] = \mathcal{S}_\alpha[u_\varepsilon] \leqslant \log(CR) - \alpha R^2 < \mathcal{S}_\alpha[u_1] = \max_{u\in\mathfrak{D}\cap\mathcal{G}_{+1}} \mathcal{S}_\alpha[u] \leqslant \max_{u\in\mathfrak{D}\cap\mathcal{G}_{+\varepsilon}} \mathcal{S}_\alpha[u].$$

This verifies (2.4.10).

For each $\varepsilon > 0$, let

$$f_{+\varepsilon}(u) = \begin{cases} \mathcal{S}_\alpha[u] & \text{if } u \in \mathfrak{D} \cap \mathcal{G}_{+\varepsilon}, \\ -\infty & \text{otherwise,} \end{cases} \qquad f_{+0}(u) = \sup_{\varepsilon>0} f_{+\varepsilon}(u) = \begin{cases} \mathcal{S}_\alpha[u] & \text{if } u \in \mathfrak{D} \cap (\cup_{\varepsilon>0}\mathcal{G}_{+\varepsilon}), \\ -\infty & \text{otherwise.} \end{cases}$$

Since the $\mathcal{G}_{+\varepsilon}$'s are nested and $\mathcal{S}_\alpha$ is continuous, we have

$$\begin{aligned} \lim_{\varepsilon\downarrow 0} \sup_{u\in\mathfrak{D}\cap\mathcal{G}_{+\varepsilon}} \mathcal{S}_\alpha[u] &= \sup_{\varepsilon>0} \sup_{u\in\mathfrak{D}\cap\mathcal{G}_{+\varepsilon}} \mathcal{S}_\alpha[u] = \sup_{\varepsilon>0} \sup_{u\in\mathbb{R}^m} f_{+\varepsilon}(u) = \sup_{u\in\mathbb{R}^m} \sup_{\varepsilon>0} f_{+\varepsilon}(u) = \sup_{u\in\mathbb{R}^m} f_{+0}(u) \\ &= \sup_{u\in\mathfrak{D}\cap(\cup_{\varepsilon>0}\mathcal{G}_{+\varepsilon})} \mathcal{S}_\alpha[u] = \sup_{u\in\overline{\mathfrak{D}\cap(\cup_{\varepsilon>0}\mathcal{G}_{+\varepsilon})}} \mathcal{S}_\alpha[u] = \sup_{u\in\mathfrak{D}\cap\mathcal{G}} \mathcal{S}_\alpha[u], \end{aligned}$$

where the last equality follows from (2.4.8). □

Chapter 3

Landscape complexity beyond invariance and the elastic manifold

This chapter is essentially borrowed from [36], joint with Gérard Ben Arous and Paul Bourgade, which will appear on the arXiv soon.

3.1 Introduction

3.1.1 Complexity of the landscape of disordered elastic systems. The elastic manifold is a paradigmatic representative of the class of disordered elastic systems. These are surfaces with rugged shapes resulting from a competition between random spatial impurities (preferring disordered configurations), on the one hand, and elastic self-interactions (preferring ordered configurations), on the other. The model is defined through its Hamiltonian (3.2.2); for example, a one-dimensional such surface is a polymer; a d-dimensional such surface could describe the interface between ordered phases with opposite signs in a $(d+1)$-dimensional Ising model. Among other motivations, the elastic manifold is interesting because it displays a (de)pinning phase transition, which is a certain nonlinear response to a driving force: if one applies an external force to the surface at zero-temperature equilibrium, then the surface moves if and only if the force is above

the depinning threshold. The elastic manifold also has a long history as a testing ground for new approaches, for example for fixed d by Fisher using functional renormalization group methods [80], and in the high-dimensional limit by Mézard and Parisi using the replica method [122].

In the same diverging dimension regime, we study the energy landscape of this model, through the expected number of configurations that locally minimize the Hamiltonian against small perturbations. We also count the expected number of critical configurations. Our main result, Theorem 3.2.4, gives the phase diagram in the model parameters, and identifies the boundary between simple and glassy phases as a physical parameter known as the Larkin mass, which appears in the (de)pinning theory, confirming recent formulas by Fyodorov and Le Doussal [88].

The proof proceeds by dimension reduction and naturally leads to analyzing a generalization of the zero-dimensional elastic manifold. The original zero-dimensional elastic manifold is

$$\mathcal{H}_N(x) = V_N(x) + \frac{\mu}{2}\|x\|^2, \tag{3.1.1}$$

where $V_N : \mathbb{R}^N \to \mathbb{R}$ is an isotropic Gaussian field and $\mu > 0$. This has been studied by Fyodorov as a toy model of a disordered system; it admits a continuous phase transition between order for large μ and disorder for small μ [84]. We replace the parabolic well confinement $\frac{\mu}{2}\|x\|^2$ with any positive definite quadratic form $\frac{1}{2}\langle x, D_N x\rangle$, to see how different signal strengths in different directions affect the complexity; this defines the model of *soft spins in an anisotropic well*. Theorem 3.2.8 identifies a simple scalar parameter distinguishing between positive and zero complexity in high dimension, namely the negative second moment of the limiting empirical measure of D_N. We also find that the near-critical decay of complexity is described by universal exponents: quadratic for total critical points, and cubic for minima.

Our work is part of the landscape complexity research program, which was initially developed for a variety of functions which are invariant under large classes of isometries (see Section 3.1.3). We address landscapes lacking this property, which we call "non-invariant." The elastic manifold model is a proof of concept for our general approach, which relies on the Kac-Rice formula to reduce

complexity to the calculation of the determinant of random matrices, and on our companion paper [35] for such determinant asymptotics for random matrix ensembles which are not invariant under orthogonal conjugacy. This gives variational formulas for the annealed complexity such as Theorem 3.4.1 for the elastic manifold.

Such variational problems associated to high dimensional Gaussian fields are not solvable in general (see e.g. the companion paper [120] about bipartite spherical spin glasses). However, for the elastic manifold, a key convexity property inherited from the associated Matrix Dyson Equation (see Proposition 3.4.9) reduces the dimension of the relevant variational formula, mapping the problem to the complexity of the soft spins in an anisotropic well model for a specific D_N. We then find integrable dynamics to analyze the variational problems associated to the general soft spins in an anisotropic well model, and obtain the complexity thresholds mentioned above.

3.1.2 Determinants and the Kac-Rice formula. As mentioned in the previous section, the Kac-Rice formula provides a bridge between random geometry and random matrix theory. If f is a Gaussian field with enough regularity on a nice compact manifold $\mathcal{M}$, and if $\mathrm{Crt}_f(t,k)$ denotes the number of critical points of f of index k at which $f \leqslant t$, then this formula reads

$$\mathbb{E}[\mathrm{Crt}_f(t,k)] = \int_{\mathcal{M}} \mathbb{E}\Big[\Big|\det(\nabla^2 f(\sigma)\Big|\mathbf{1}\{f(\sigma)\leqslant t, i(\nabla^2 f(\sigma))=k\}\Big|\nabla f(\sigma)=0\Big]\phi_\sigma(0)\,\mathrm{d}\sigma.$$

Here $i(\cdot)$ is the index and $\phi_\sigma(0)$ is the density of $\nabla f(\sigma)$ at 0. In the models of this paper, we will always take $\mathcal{M}$ to be the whole Euclidean space (with the necessary arguments to account for non-compactness). Thus the Kac-Rice formula transforms questions about critical points into questions about the (conditional) determinant of the random matrix $\nabla^2 f(\sigma)$. For an introduction to the Kac-Rice formula, we direct the reader to [2, 17]. In a digestible special case, if Crt_f is the total number of critical points of f, then

$$\mathbb{E}[\mathrm{Crt}_f] = \int_{\mathcal{M}} \mathbb{E}\Big[\Big|\det(\nabla^2 f(\sigma)\Big|\Big|\nabla f(\sigma)=0\Big]\phi_\sigma(0)\,\mathrm{d}\sigma. \tag{3.1.2}$$

In one dimension, this formula dates back to the 1940s [107, 133]. For many years it was used for small, fixed dimension in applications such as signal processing [134] and oceanography [117]. For more modern results in fixed dimension, we refer the reader to [16].

3.1.3 Rotationally invariant models. In a breakthrough insight, [84] used the Kac-Rice formula in *diverging* dimension, to study *asymptotic* counts of critical points via *asymptotics* of random determinants. For example, if $f = f_N$ in the above discussion is defined on an N-dimensional manifold, one attempts to compute $\lim_{N\to\infty} \frac{1}{N} \log \mathbb{E}[\mathrm{Crt}_{f_N}]$. The papers [84] and [94] studied isotropic Gaussian fields in radially symmetric confining potentials; the centered isotropic case without confining potentials (but in finite volume) was treated in [62]. Work has been done in the mathematics and physics literature on complexity for spherical p-spin models, starting with [10] (for pure models) and [9] (for mixtures). Similar techniques were used to understand the spiked-tensor model in [43]. Intricate questions, such as the number of critical points with fixed index at given overlap from a minimum, are considered for pure p-spin models in [135]. We also mention [78] for an upper bound on the number of critical points of the TAP free energy of the Sherrington-Kirkpatrick model, and the recent works [29, 30] on neural networks, [13] on Gaussian fields with isotropic increments, [38] on stable/unstable equilibria in systems of non-linear differential equations, and [34] on mixed spherical spin glasses with a deterministic external field. In most of these models, the conditioned Hessian is closely related to the Gaussian Orthogonal Ensemble (GOE), a consequence of distributional symmetries of the landscapes.

The above results handle the average number of critical points. It is another question entirely to prove concentration, i.e. to show that the *average* (annealed) number of points is also *typical* (quenched). Proving concentration typically involves intricate second-moment computations, which are also possible via the Kac-Rice formula, but which involve determinant asymptotics for a pair of (usually correlated) random matrices. To our knowledge this has only been carried out for p-spin models, both for pure models [141, 12] and for certain mixtures which are close to pure [44]. The quenched asymptotics are not always expected to match the annealed ones; for more intricate

questions in pure p-spin, physical computations based on the replica trick suggest a qualitative picture of this failure [136, 137].

3.1.4 Non-invariant models. In many models of interest, it happens that the law of the conditioned Hessian in (3.1.2) does not depend on σ, and that it has long-range correlations induced by a fixed (not depending on N) number m of independent Gaussian random variables. For example, this law might match that of $W_N + \xi \,\mathrm{Id}$, where W_N is symmetric with independent Gaussian entries with a variance profile or large zero blocks, and $\xi \sim \mathcal{N}(0, \frac{1}{N})$ is independent of W_N; the resulting matrix has "long-range correlations" because the diagonal entries are all correlated with each other, and $m = 1$ because these correlations are induced by $\xi \in \mathbb{R}^1$. In these models, by integrating over this small number of variables last, the difficult term in the Kac-Rice formula (3.1.2) takes the form

$$\int_{\mathbb{R}^m} e^{-N\frac{\|u\|^2}{2}} \mathbb{E}[|\det(H_N(u))|] \,\mathrm{d}u \tag{3.1.3}$$

for some Gaussian random matrices $H_N(u)$ which may be far from GOE. (In the example above, $H_N(u) = W_N + u \,\mathrm{Id}$.)

The problem then reduces to the exponential asymptotics of (3.1.3). In the companion paper [35], we establish two types of results about (3.1.3). First, we show asymptotics for a single matrix of the form

$$\mathbb{E}[|\det(H_N(u))|] = \exp\Big(N \int_{\mathbb{R}} \log|\lambda| \mu_N(u, \mathrm{d}\lambda) + o(N) \Big). \tag{3.1.4}$$

Here the deterministic probability measures $\mu_N(u) = \mu_N(u, \cdot)$ come from the theory of the *Matrix Dyson Equation* (MDE), developed in the random-matrix literature by Erdős and co-authors in the last several years. Second, after this identification, (3.1.3) looks like a Laplace-type integral (with error terms), but the measures μ_N depend on N, meaning (3.1.3) may take the form $\int_{\mathbb{R}^m} e^{N f_N(u)} \,\mathrm{d}u$ instead of the more-desirable $\int_{\mathbb{R}^m} e^{N f(u)} \,\mathrm{d}u$. In [35], we show that – *assuming* the limits $\mu_N(u) \to \mu_\infty(u)$ exist – the Laplace method can be carried out on (3.1.3).

In this paper we discuss how to identify the limits $\mu_N(u) \to \mu_\infty(u)$ for the elastic manifold

and soft spins in an anisotropic well (a third model is treated in the companion paper [120]). This is model-dependent, although we identify some common techniques. This leads to the following informal statement:

Metatheorem 3.1.1. *Let $\mathcal{M}_N$ be a nice sequence of N-dimensional manifolds, and let $f_N : \mathcal{M}_N \to \mathbb{R}$ be a sequence of Gaussian random landscapes with the properties discussed above (namely, the law of the conditioned Hessian is independent of the basepoint on $\mathcal{M}_N$, and long-range correlations are induced by m independent variables). If the limiting empirical measures $\mu_\infty(u)$ can be identified and some regularity established in u (and we present models where this is possible), then*

$$\lim_{N\to\infty} \frac{1}{N}\mathbb{E}[\mathrm{Crt}_{f_N}] = \sup_{u\in\mathbb{R}^m} \left\{ \int_{\mathbb{R}} \log|\lambda| \mu_\infty(u, \mathrm{d}\lambda) - \frac{\|u\|^2}{2} \right\} + \textit{simpler non-variational term.} \quad (3.1.5)$$

The non-variational term comes from the density of the gradient in the Kac-Rice formula: precisely, it is equal to $\lim_{N\to\infty} \frac{1}{N} \log \int_{\mathcal{M}_N} \phi_\sigma(0)\, \mathrm{d}\sigma$, which is typically easy to calculate.

We also wish to count local minima, for which the analogue of (3.1.3) is

$$\int_{\mathfrak{D}} e^{-N\frac{\|u\|^2}{2}} \mathbb{E}[|\det(H_N(u))| \mathbf{1}_{H_N(u)\geqslant 0}]\, \mathrm{d}u.$$

If we define the set

$$\mathcal{G} = \{u \in \mathbb{R}^m : \mu_\infty(u)((-\infty, 0)) = 0\}$$

of good u values for which $\{H_N(u) \geqslant 0\}$ is a likely event, then the upshot is that at exponential scale we have

$$\mathbb{E}[|H_N(u)| \mathbf{1}_{H_N(u)\geqslant 0}] \approx \begin{cases} \mathbb{E}[|H_N(u)|] & \text{if } u \in \mathcal{G}, \\ 0 & \text{otherwise.} \end{cases} \quad (3.1.6)$$

(All the matrices $H_N(u)$ we encounter have asymptotically no outliers; otherwise, large-deviations estimates for edge eigenvalues would impact the final result.) This gives an analogue of Metatheorem 3.1.1 for the complexity of local minima, where the variational problem is restricted to a supremum over $u \in \mathcal{G}$ instead of $u \in \mathbb{R}^m$. Again, the argument was presented in [35] assuming the existence

of limits $\mu_N(u) \to \mu_\infty(u)$; in this paper we verify this assumption.

The goal of this paper is to carry out this program for the elastic manifold and the anisotropic soft spins model, yielding precise versions of Metatheorem 3.1.1 and its analogue for minima. In fact, for these particular models the variational problem in (3.1.5) turns out to be integrable, as mentioned at the end of Section 3.1.1: By introducing a dynamic version of the optimization (3.1.5), we can distinguish regimes of positive and zero complexity. In addition, we can study near-critical behavior at this phase transition, showing that complexity of total critical points tends to zero quadratically, whereas complexity of local minima tends to zero cubically. These critical exponents were already known for certain models [84, 94]; we show their universality by extending substantially the class of models exhibiting these quadratic and cubic transitions.

We state our main results in Section 3.2. Section 3.3 provides techniques that will be shared across models, showing how the (well-established) stability theory of the MDE allows one to replace $\mu_N(u)$ by $\mu_\infty(u)$ as discussed above, if one has a candidate μ_∞. In the remaining sections, we propose candidates for μ_∞ and carry out this program for each of our models in turn. In Appendix B, we prove a result in free probability necessary to identify near-critical complexity of our models, and possibly of independent interest: The free convolution of any (compactly supported) measure with the semicircle law decays at least as quickly as a square root at its extremal edges.

Notations. We write $\|\cdot\|$ for the operator norm on elements of $\mathbb{C}^{N\times N}$ induced by Euclidean distance on $\mathbb{C}^N$, and if $\mathcal{S} : \mathbb{C}^{N\times N} \to \mathbb{C}^{N\times N}$, we write $\|\mathcal{S}\|$ for the operator norm induced by $\|\cdot\|$. We let

$$\|f\|_{\mathrm{Lip}} = \sup_{x\neq y}\left|\frac{f(x)-f(y)}{x-y}\right|$$

for test functions $f : \mathbb{R} \to \mathbb{R}$, and write d for the bounded-Lipschitz distance on probability measures on $\mathbb{R}$:

$$d_{\mathrm{BL}}(\mu,\nu) = \sup\left\{\left|\int_{\mathbb{R}} f\,\mathrm{d}(\mu-\nu)\right| : \|f\|_{\mathrm{Lip}} + \|f\|_{L^\infty} \leqslant 1\right\}.$$

We will need the semicircle law of variance t, which we write as

$$\rho_{\mathrm{sc},t}(\mathrm{d}x) = \frac{\sqrt{4t - x^2}}{2\pi t} \mathbf{1}_{x \in [-2\sqrt{t}, 2\sqrt{t}]} \, \mathrm{d}x,$$

as well as the abbreviation $\rho_{\mathrm{sc}} = \rho_{\mathrm{sc},1}$ for the usual semicircle law supported in $[-2, 2]$. We write $\mathtt{l}(\mu)$ for the left edge (respectively, $\mathtt{r}(\mu)$ for the right edge) of a compactly supported measure μ. For an $N \times N$ Hermitian matrix M, we write $\lambda_{\min}(M) = \lambda_1(M) \leqslant \cdots \leqslant \lambda_N(M) = \lambda_{\max}(M)$ for its eigenvalues and

$$\hat{\mu}_M = \frac{1}{N} \sum_{i=1}^{N} \delta_{\lambda_i(M)}$$

for its empirical measure. We write $\odot$ for the entrywise (i.e., Hadamard) product of matrices, and $\boxplus$ for the free (additive) convolution of probability measures. Given a matrix T, we write $\mathrm{diag}(T)$ for the diagonal matrix of the same size obtained by setting all off-diagonal entries to zero. In equations, we sometimes identify diagonal matrices with vectors of the same size. We write B_R for the ball of radius R about zero in the relevant Euclidean space. We use $(\cdot)^T$ for the matrix transpose, which should be distinguished both from $(\cdot)^*$ for the matrix conjugate transpose, and from $\mathrm{Tr}(\cdot)$ for the matrix trace.

Unless stated otherwise, z will always be a complex number in the upper half-plane $\mathbb{H} = \{z \in \mathbb{C} : \mathrm{Im}(z) > 0\}$, and we always write its real and imaginary parts as $z = E + \mathrm{i}\eta$.

Acknowledgements. We wish to thank László Erdős and Torben Krüger for many helpful discussions about the MDE for block random matrices, in particular for communicating arguments that became the proofs of Lemmas 3.4.2 and 3.4.10. We are also grateful to Yan Fyodorov and Pierre Le Doussal for bringing the elastic-manifold problem to our attention, and for helpful comments. BM thanks Krishnan Mody for helpful discussions. GBA acknowledges support by the Simons Foundation collaboration Cracking the Glass Problem, PB was supported by NSF grant DMS-1812114 and a Poincaré chair, and BM was supported by NSF grant DMS-1812114.

3.2 Main results

3.2.1 Elastic manifold. Fix positive integers L ("length") and d ("internal dimension"), positive numbers μ_0 ("mass") and t_0 ("interaction strength"), and write Ω for the lattice $[\![1, L]\!]^d \subset \mathbb{Z}^d$, understood periodically. Let V_N be a centered Gaussian field on $\mathbb{R}^N \times \Omega$ with

$$\mathbb{E}[V_N(y_1, x_1)V_N(y_2, x_2)] = NB\left(\frac{\|y_1 - y_2\|^2}{N}\right)\delta_{x_1,x_2},$$

for some function $B : \mathbb{R}_+ \to \mathbb{R}_+$ called the correlator. Schoenberg characterized all possible such correlators [139, Theorem 2] (see also [156]); B must have the representation

$$B(x) = c_0 + \int_0^\infty \exp(-t^2 x)\nu(\mathrm{d}t) \tag{3.2.1}$$

for some $c_0 \geqslant 0$ and some finite non-negative measure ν on $(0, \infty)$. In particular B is infinitely differentiable and non-increasing on $(0, \infty)$. We assume that B is also four times differentiable at zero, which implies via Kolmogorov's criterion that each $V_N(\cdot, x)$ is almost surely twice differentiable. We will also assume

$$0 < |B^{(i)}(0)| \qquad \text{for } i = 0, 1, 2,$$

which should be interpreted as a non-degeneracy condition on the field ($i = 0$), its gradient ($i = 1$), and its Hessian ($i = 2$). This is a very mild assumption; indeed it holds by dominated convergence as soon as the measure ν in (3.2.1) has a finite fourth moment and is not the zero measure.

To each deterministic function $\mathbf{u} : \Omega \to \mathbb{R}^N$ ("point configuration," but sometimes "manifold" after the continuous analogue) associate the random Hamiltonian

$$\mathcal{H}[\mathbf{u}] = \sum_{x,y\in\Omega} (\mu_0 \,\mathrm{Id} - t_0\Delta)_{xy}\langle \mathbf{u}(x), \mathbf{u}(y)\rangle + \sum_{x\in\Omega} V_N(\mathbf{u}(x), x). \tag{3.2.2}$$

Here $\Delta \in \mathbb{R}^{L^d \times L^d}$ is the (periodic) lattice Laplacian on Ω, so the (x, y) entry of $\mu_0 \,\mathrm{Id} - t_0\Delta$ is given

by

$$(\mu_0 \operatorname{Id} - t_0\Delta)_{xy} = \mu_0\delta_{x=y} - t_0(\delta_{x\sim y} - 2d\delta_{x=y}),$$

where $x \sim y$ means that x and y are lattice neighbors. (Following [88], our Laplacian is a negative sign off from the typical mathematical convention.)

Notice that the different energies compete: If the disorder V_N vanished in (3.2.2), then since $\mu_0 \operatorname{Id}$ and $-t_0\Delta$ are both positive semidefinite, the ground-state configuration would be the flat one $\mathbf{u} \equiv 0$. On the other hand, the disorder V_N prefers certain random configurations; the interaction $-t_0\Delta$ prefers to keep these configurations from becoming too jagged; and the confinement μ_0 prefers to keep them close to the origin. See Figure 1.2 for a graphical interpretation.

History. Hamiltonians of this flavor have been used to model a wide variety of problems featuring surfaces with self-interactions in disordered media. For example, when $d = 1$, the model is a polymer, related to the KPZ universality class; when $N = d + 1$, the model is an interface, such as that between regions of opposite magnetization in a ferromagnet. We direct readers to [95] and [96] for a review of disordered elastic media in general and to [89] for a review of this specific Hamiltonian, which we summarize briefly here.

Two phenomena are of primary interest: the *depinning threshold* f_c and the *wandering (or roughness) exponent* ζ. The former refers to the manifold's nonlinear response to an applied force f, a consequence of the impurities in the potential V: at zero temperature, it moves from its preferred position only if the force is above the depinning threshold $f > f_c = f_c(L, d, t_0, N)$, whereas if $f \leqslant f_c$ it does not move at all and is said to be *pinned*. (Depinning is typically discussed in the massless limit $\mu_0 \downarrow 0$, but restricting the manifold points to lie in a finite box. At positive temperature, the manifold can move when $f < f_c$, but the movement is typically slow and is called *creep*; the movement above f_c is faster.) Depinning is related to complexity: Adding a force changes the Hamiltonian, and the landscape is supposed to simplify as f increases; then f_c can be defined as the smallest f for which the resulting (quenched) complexity vanishes. We do not study this

connection further, but refer readers to a discussion in [88].

The wandering exponent ζ, which depends on d and N, is defined by

$$\mathbb{E}[(\mathbf{u}_0(x) - \mathbf{u}_0(y))^2] \sim \|x - y\|^{2\zeta}$$

where $\mathbf{u}_0$ is the ground state. It is generally believed that $\zeta = 0$, i.e. that the manifold is flat, for $d \geqslant 4$. Larkin proposed a simplification of the Hamiltonian (3.2.2), replacing the terms $V_N(\mathbf{u}(x), x)$ with their linearizations $V_N(0, x) + \partial_y V_N(0, x)|_{y=0}\mathbf{u}(x)$. This so-called *Larkin model* is solvable and gives $\zeta = \left(\frac{4-d}{2}\right)_+$; note also that the Larkin model is quadratic in $\mathbf{u}$, hence only has one local minimum, i.e., is necessarily zero-complexity. Physicists believe that the Larkin model is a good approximation for the elastic manifold when L is below the *Larkin length* L_c, with $L_c \sim (B''(0))^{-1/(4-d)}$ for weak disorder. Above the Larkin length the approximation is supposed to break down, and describing the physics of the elastic manifold (in particular finding ζ) is more challenging. This regime inspired early technical developments of Fisher in functional renormalization group methods [80] and of Mézard and Parisi in the replica method [122]; the latter paper suggested that the system exhibits zero-temperature replica symmetry breaking for small μ_0 in the $N \to +\infty$ limit. (This is the same limit we will consider, although of course one is ultimately interested in finite-N results.) Increasing the "mass" μ_0 has the effect of simplifying the landscape, and for μ_0 larger than a *Larkin mass* μ_c (related to the Larkin length L_c), the system is believed to be replica symmetric. In fact the Larkin mass is central to our results; we are making rigorous a result of Fyodorov and Le Doussal suggesting that, for all other parameters fixed, μ_c is precisely the boundary between zero complexity (for $\mu_0 \geqslant 2\sqrt{B''(0)}\mu_c$) and positive complexity (for $\mu_0 \leqslant 2\sqrt{B''(0)}\mu_c$). The same μ_c serves as the boundary both for total critical points and for local minima.

There are some previous complexity results for special cases. When $d = 0$, the system is interpreted by convention as being a single point, i.e., it reduces to the Hamiltonian (3.1.1). Fyodorov computed the complexity of (3.1.1) and found a continuous phase transition in μ: For $\mu \geqslant \mu_c$, the annealed complexity (of the total number of critical points) is zero and the land-

scape is "simple," but for $\mu < \mu_c$ the annealed complexity is positive and the landscape is "complex" or "glassy" [84]. Later, Fyodorov and Williams showed that this phase transition matches that of replica-symmetry/replica-symmetry-breaking at zero temperature [94], interpreting replica-symmetry-breaking as "a replica-symmetric computation of the free energy becomes unstable in the zero-temperature limit." For more discussion of the $d = 0$ case, see Section 3.2.2 below. When $d = 1$, the model is an elastic line, with complexity studied in the case of $N = 1$ and $L \to +\infty$ in [90].

Results. Let $\mathcal{N}_{\mathrm{tot}}$ be the random number of stationary points of the Hamiltonian, i.e., of functions $\mathbf{u} : \Omega \to \mathbb{R}^N$ such that $\partial_{\mathbf{u}_i(x)}\mathcal{H}[\mathbf{u}] = 0$ for every $x \in \Omega$ and every $i = 1, \dots, N$. Let $\mathcal{N}_{\mathrm{st}}$ be the number of local minima.

Definition 3.2.1. *For any $\mu_0, t_0, b > 0$, define*

$$\begin{aligned}
&\Sigma(\mu_0, t_0, b) = \Sigma(\mu_0, t_0, b, L, d) \\
&= -\frac{1}{L^d}\log(\det(\mu_0 \operatorname{Id}_{L^d \times L^d} - t_0\Delta)) + \sup_{u\in\mathbb{R}}\left\{\int_{\mathbb{R}} \log|\lambda - u|(\rho_{sc,b} \boxplus \hat{\mu}_{-t_0\Delta+\mu_0 \operatorname{Id}})(\lambda)\,\mathrm{d}\lambda - \frac{u^2}{2b}\right\}, \\
&\Sigma_{\mathrm{st}}(\mu_0, t_0, b) = \Sigma_{\mathrm{st}}(\mu_0, t_0, b, L, d) \\
&= -\frac{1}{L^d}\log(\det(\mu_0 \operatorname{Id}_{L^d \times L^d} - t_0\Delta)) \\
&\quad + \sup_{u \leqslant 1(\rho_{sc,b}\boxplus\hat{\mu}_{-t_0\Delta+\mu_0 \operatorname{Id}})}\left\{\int_{\mathbb{R}} \log|\lambda - u|(\rho_{sc,b} \boxplus \hat{\mu}_{-t_0\Delta+\mu_0 \operatorname{Id}})(\lambda)\,\mathrm{d}\lambda - \frac{u^2}{2b}\right\}.
\end{aligned} \tag{3.2.3}$$

Theorem 3.2.2. *We have*

$$\begin{aligned}
\lim_{N\to\infty} \frac{1}{NL^d}\log \mathbb{E}[\mathcal{N}_{\mathrm{tot}}] &= \Sigma(\mu_0, t_0, 4B''(0)), \\
\lim_{N\to\infty} \frac{1}{NL^d}\log \mathbb{E}[\mathcal{N}_{\mathrm{st}}] &= \Sigma_{\mathrm{st}}(\mu_0, t_0, 4B''(0)).
\end{aligned} \tag{3.2.4}$$

Definition 3.2.3. *For any $t_0, b > 0$, let the* Larkin mass *$\mu_c = \mu_c(t_0, b, L, d)$ be the unique positive*

solution to

$$\int_{\mathbb{R}} \frac{\hat{\mu}_{-t_0\Delta}(\mathrm{d}\lambda)}{(\mu_c+\lambda)^2} = \frac{1}{b}. \tag{3.2.5}$$

It will also be useful to define, for any $\mu_0, t_0 > 0$, the critical noise parameter

$$b_c = b_c(\mu_0, t_0, L, d) = \left(\int_{\mathbb{R}} \frac{\hat{\mu}_{-t_0\Delta}(\mathrm{d}\lambda)}{(\mu_0+\lambda)^2}\right)^{-1}.$$

For $\mu_0 < \mu_c(t_0, b, L, d)$, we write $c = c(\mu_0, t_0, b, L, d)$ for the unique positive value satisfying

$$\int_{\mathbb{R}} \frac{\hat{\mu}_{-t_0\Delta}(\mathrm{d}\lambda)}{(\mu_0+\lambda)^2 + b^2 c} = \frac{1}{b}$$

and use this to define

$$v = v(\mu_0, t_0, b, L, d) = -b\int_{\mathbb{R}} \frac{\mu_0+\lambda}{(\mu_0+\lambda)^2 + b^2 c}\, \hat{\mu}_{-t_0\Delta}(\mathrm{d}\lambda).$$

Finally, we need the positive numbers

$$c_{\mathrm{tot}}(\mu_0, t_0, L, d) = \frac{\left(\int_{\mathbb{R}} \frac{\hat{\mu}_{-t_0\Delta}(\mathrm{d}\lambda)}{(\mu_0+\lambda)^2}\right)^4}{4\left(\int_{\mathbb{R}} \frac{\hat{\mu}_{-t_0\Delta}(\mathrm{d}\lambda)}{(\mu_0+\lambda)^4}\right)}, \qquad c_{\min}(\mu_0, t_0, L, d) = \frac{\left(\int_{\mathbb{R}} \frac{\hat{\mu}_{-t_0\Delta}(\mathrm{d}\lambda)}{(\mu_0+\lambda)^2}\right)^6}{24\left(\int_{\mathbb{R}} \frac{\hat{\mu}_{-t_0\Delta}(\mathrm{d}\lambda)}{(\mu_0+\lambda)^3}\right)^2}.$$

Theorem 3.2.4. *For each t_0 and $B''(0)$, the Larkin mass μ_c separates the phases of positive and zero complexity, both for total critical points (whose complexity exhibits quadratic near-critical behavior) and for local minima (whose complexity exhibits cubic near-critical behavior).*

More precisely, the complexity functions satisfy the following, where $b = 4B''(0)$:

(i) if $\mu_0 \geqslant \mu_c(t_0, b, L, d)$, then $\Sigma(\mu_0, t_0, b) = \Sigma_{\mathrm{st}}(\mu_0, t_0, b) = 0$;

(ii) if $\mu_0 < \mu_c(t_0, b, L, d)$, then $\Sigma(\mu_0, t_0, b) > \Sigma_{\mathrm{st}}(\mu_0, t_0, b) > 0$, and these are given by

$$\Sigma(\mu_0, t_0, b) = -\frac{1}{L^d}\log(\det(\mu_0 \operatorname{Id} - t_0\Delta)) + \int_{\mathbb{R}} \log|\lambda - v|(\rho_{sc,b} \boxplus \hat{\mu}_{-t_0\Delta+\mu_0 \operatorname{Id}})(\lambda)(\mathrm{d}\lambda) - \frac{v^2}{2b},$$
$$\Sigma_{\mathrm{st}}(\mu_0, t_0, b) = -\frac{1}{L^d}\log(\det(\mu_0 \operatorname{Id} - t_0\Delta)) + \int_{\mathbb{R}} \log|\lambda - \ell|(\rho_{sc,b} \boxplus \hat{\mu}_{-t_0\Delta+\mu_0 \operatorname{Id}})(\lambda)(\mathrm{d}\lambda) - \frac{\ell^2}{2b}$$

where $\ell = \mathtt{l}(\rho_{sc,b} \boxplus \hat{\mu}_{-t_0\Delta+\mu_0 \operatorname{Id}})$ and v is as above; and

(iii) for fixed μ_0 and t_0, and supercritical b, we have

$$\Sigma(\mu_0, t_0, b) = c_{tot}(\mu_0, t_0, L, d) \cdot (b - b_c)^2 + O((b - b_c)^3),$$
$$\Sigma_{\mathrm{st}}(\mu_0, t_0, b) = c_{\min}(\mu_0, t_0, L, d) \cdot (b - b_c)^3 + O((b - b_c)^4).$$

For the proof of this theorem, we use determinant asymptotics from our companion paper [35] to give the complexity as a variational problem over $\mathbb{R}^{L^d}$. Using a remarkable MDE-induced convexity property, we reduce this to a variational problem over $\mathbb{R}$, namely (3.2.3). We analyze this one-dimensional variational problem with a dynamic approach, varying $B''(0)$ for fixed μ_0 and t_0.

We remark that Fyodorov and Le Doussal also exhibited a quadratic/cubic near-critical behavior for this model but in a different scaling, varying μ_0 for fixed $B''(0)$ and t_0 [88].

3.2.2 Soft spins in an anisotropic well. We consider the random Hamiltonian $\mathcal{H}_N : \mathbb{R}^N \to \mathbb{R}$ given by

$$\mathcal{H}_N(x) = \frac{\langle x, D_N x\rangle}{2} + V_N(x),$$

where D_N is a real symmetric matrix satisfying conditions below, and where V_N is an isotropic centered Gaussian field with covariance

$$\mathbb{E}[V_N(x_1)V_N(x_2)] = NB\left(\frac{\|x_1 - x_2\|^2}{2N}\right)$$

with $B : \mathbb{R}_+ \to \mathbb{R}_+$ a correlator function (meaning it has the representation (3.2.1)). As in Section 3.2.1, we assume that B is four times differentiable at zero to ensure twice-differentiability of the field, and we assume

$$0 < |B^{(i)}(0)| \qquad \text{for } i = 0, 1, 2,$$

for nondegeneracy of the field and its first two derivatives.

We suppose that $(D_N)_{N=1}^{\infty}$ is a sequence of real symmetric matrices, $D_N \in \mathbb{R}^{N \times N}$, and that there exists some compactly supported measure μ_D such that, for some $\varepsilon > 0$, we have

$$d_{\mathrm{BL}}(\hat{\mu}_{D_N}, \mu_D) \leqslant N^{-\varepsilon} \tag{3.2.6}$$

and the eigenvalues are uniformly gapped away from zero and from infinity, in that

$$\varepsilon \leqslant \inf_N \lambda_{\min}(D_N) \leqslant \sup_N \lambda_{\max}(D_N) \leqslant \frac{1}{\varepsilon}.$$

Although our results are for the $N \to +\infty$ limit, Figure 1.3 displays how changing D_N can qualitatively change the count of critical points when $N = 2$.

History. Models of the form $V_N(x) + \frac{\mu}{2}\|x\|^2$ (recall (3.1.1)), with various choices of randomness, have been considered in a wide variety of contexts. There are nice overviews of the literature in [84, 94, 13]. In the early 1990s, the model was studied by Mézard-Parisi [123] and by Engel [73] as a zero-dimensional case of the elastic manifold. The complexity was computed by Fyodorov [84] for total critical points and Fyodorov-Williams [94] for minima, finding a phase transition between positive and zero complexity at an explicit μ_c. Fyodorov and Nadal found that the complexity of minima for μ near μ_c, scaled appropriately, tends to a limiting shape related to the Tracy-Widom distribution [91].

There is also a long history of generalizing the model, as we do: Fyodorov and Williams actually studied the complexity after replacing the quadratic confinement $\frac{\mu}{2}\|x\|^2$ with a general radial

confinement $NU(\frac{\|x\|^2}{2N})$ for some function $U : \mathbb{R} \to \mathbb{R}$ which is increasing and convex [94]. In some sense our extension is orthogonal to theirs: they let the confinement be non-quadratic, whereas we let it be non-radial. As another generalization, if $V_N(x)$ is not isotropic but merely has isotropic *increments* (meaning $\mathbb{E}[(V_N(x) - V_N(y))^2]$ depends only $\|x - y\|$), then the model can admit long-range correlations; this was studied in the physics literature by Fyodorov and co-authors [92, 85], and its complexity was recently computed by Auffinger and Zeng [13].

Our generalization is reminiscent of the work of Fan, Mei, and Montanari on an upper bound for the complexity of the TAP free energy of the Sherrington-Kirkpatrick model [78]. Indeed, via the Kac-Rice formula, the random matrix that appears in our problem is a full-rank deformation of GOE (see (3.5.3)). A similar random matrix, in fact with an additional low-rank deformation, appears in [78].

Results. Let $\mathrm{Crt}_N^{\mathrm{tot}}(\mathcal{H}_N)$ be the total number of critical points of $\mathcal{H}_N$ and $\mathrm{Crt}_N^{\mathrm{min}}(\mathcal{H}_N)$ be the total number of local minima.

Definition 3.2.5. *For any $t > 0$ and any μ_D compactly supported in $(0, \infty)$, define*

$$\Sigma^{\mathrm{tot}}(\mu_D, t) = -\int_{\mathbb{R}} \log(\lambda)\mu_D(\mathrm{d}\lambda) + \sup_{u \in \mathbb{R}} \left\{ \int_{\mathbb{R}} \log|\lambda - u|(\rho_{sc,t} \boxplus \mu_D)(\lambda)\,\mathrm{d}\lambda - \frac{u^2}{2t} \right\}, \tag{3.2.7}$$

$$\Sigma^{\mathrm{min}}(\mu_D, t) = -\int_{\mathbb{R}} \log(\lambda)\mu_D(\mathrm{d}\lambda) + \sup_{u \leqslant \mathtt{1}(\rho_{sc,t} \boxplus \mu_D)} \left\{ \int_{\mathbb{R}} \log|\lambda - u|(\rho_{sc,t} \boxplus \mu_D)(\lambda)\,\mathrm{d}\lambda - \frac{u^2}{2t} \right\}. \tag{3.2.8}$$

We will show that these suprema are achieved, possibly not uniquely.

Theorem 3.2.6 below shows the relevance of these functions for complexity, and Theorem 3.2.8 analyzes the variational problems from (3.2.7) and (3.2.8) to describe the phase portrait in μ_D and t. In particular, the regimes of positive complexity for the total number of critical points and local minima coincide for any μ_D, and the exponents describing near-critical behavior are universal in μ_D.

Theorem 3.2.6. *We have*

$$\lim_{N\to\infty} \frac{1}{N} \log \mathbb{E}[\mathrm{Crt}_N^{\mathrm{tot}}(\mathcal{H}_N)] = \Sigma^{\mathrm{tot}}(\mu_D, B''(0)).$$

If in addition D_N has no external outliers, in the sense that

$$\lim_{N\to\infty} \lambda_{\min}(D_N) = \mathtt{l}(\mu_D) \qquad \textit{and} \qquad \lim_{N\to\infty} \lambda_{\max}(D_N) = \mathtt{r}(\mu_D),$$

then

$$\limsup_{N\to\infty} \frac{1}{N} \log \mathbb{E}[\mathrm{Crt}_N^{\min}(\mathcal{H}_N)] = \Sigma^{\min}(\mu_D, B''(0)).$$

Remark 3.2.7. *We emphasize that Theorem 3.2.6 shows that special directions in the environment (meaning outliers in D_N) have no effect on the total number of critical points at exponential scale, as long as there are $o(N)$ many of them. We leave open the effect of special directions on minima.*

We define the important threshold

$$t_c = t_c(\mu_D) = \left(\int_{\mathbb{R}} \frac{\mu_D(\mathrm{d}\lambda)}{\lambda^2} \right)^{-1}. \tag{3.2.9}$$

For $t > t_c$, we write $c = c(t, \mu_D)$ for the unique positive value satisfying

$$\frac{1}{t} = \int_{\mathbb{R}} \frac{1}{\lambda^2 + t^2 c} \, \mu_D(\mathrm{d}\lambda)$$

and use this to define

$$v = v(t, \mu_D) = -t \int_{\mathbb{R}} \frac{\lambda}{\lambda^2 + t^2 c(t, \mu_D)} \, \mu_D(\mathrm{d}\lambda).$$

We also need the positive numbers

$$c_{\mathrm{tot}}(\mu_D) = \frac{\left(\int_{\mathbb{R}} \frac{\mu_D(\mathrm{d}\lambda)}{\lambda^2} \right)^4}{4\left(\int_{\mathbb{R}} \frac{\mu_D(\mathrm{d}\lambda)}{\lambda^4} \right)}, \qquad c_{\min}(\mu_D) = \frac{\left(\int_{\mathbb{R}} \frac{\mu_D(\mathrm{d}\lambda)}{\lambda^2} \right)^6}{24\left(\int_{\mathbb{R}} \frac{\mu_D(\mathrm{d}\lambda)}{\lambda^3} \right)^2}. \tag{3.2.10}$$

Theorem 3.2.8. *For every $t > 0$ and every probability measure μ_D compactly supported in $(0, \infty)$,*

(i) if $t \leqslant t_c$, then $\Sigma^{\mathrm{tot}}(\mu_D, t) = \Sigma^{\mathrm{min}}(\mu_D, t) = 0$;

(ii) if $t > t_c$, then $\Sigma^{\mathrm{tot}}(\mu_D, t) > \Sigma^{\mathrm{min}}(\mu_D, t) > 0$, and these are given by

$$\Sigma^{\mathrm{min}}(\mu_D, t) = -\int_{\mathbb{R}} \log(\lambda)\mu_D(\mathrm{d}\lambda) + \int_{\mathbb{R}} \log|\lambda - \ell|(\rho_{sc,t} \boxplus \mu_D)(\lambda)\,\mathrm{d}\lambda - \frac{\ell^2}{2t}, \tag{3.2.11}$$

$$\Sigma^{\mathrm{tot}}(\mu_D, t) = -\int_{\mathbb{R}} \log(\lambda)\mu_D(\mathrm{d}\lambda) + \int_{\mathbb{R}} \log|\lambda - v|(\rho_{sc,t} \boxplus \mu_D)(\lambda)\,\mathrm{d}\lambda - \frac{v^2}{2t}, \tag{3.2.12}$$

where $\ell = \mathtt{l}(\rho_{sc,t} \boxplus \mu_D)$ and v is as above; and

(iii) for supercritical t, we have

$$\Sigma^{\mathrm{tot}}(\mu_D, t) = c_{\mathrm{tot}}(\mu_D) \cdot (t - t_c)^2 + O((t - t_c)^3),$$
$$\Sigma^{\mathrm{min}}(\mu_D, t) = c_{\mathrm{min}}(\mu_D) \cdot (t - t_c)^3 + O((t - t_c)^4),$$

with $c_{\mathrm{tot}}(\mu_D)$, $c_{\mathrm{min}}(\mu_D)$ as in (3.2.10).

The proof of this theorem relies on a dynamic approach, like for results in Section 3.2.1. We also use two important inputs: (i) the Burgers' equation satisfied by the Stieltjes transform of the semicircle distribution, and (ii) an inequality from free probability, due to Guionnet and Maïda, regarding the subordination function of the free convolution at the edge. We also need a new result in free probability, possibly of independent interest, which we prove in Appendix B: The free convolution of any measure with semicircle decays at least as fast as a square-root at its extremal edges.

We remark that it is not obvious that the same threshold t_c should work both for total critical points and for local minima, and the analogue is false in closely related models. For example, consider the Hamiltonian (3.1.1), i.e. $H_N(x) = \frac{\mu}{2}\|x\|^2 + V_N(x)$, but defined over $\{x \in \mathbb{R}^N : \|x\| \leqslant R\sqrt{N}\}$ for some fixed $R > 0$ rather than over the whole space. Fyodorov, Sommers, and Williams [93] showed that, for some choices of R, the complexity of total critical points is positive but

the complexity of minima vanishes. (See [62] for related independent work.) But [94] proved the analogue of point 1 for their model, discussed above, which is defined on the full space. See [94, Section 2.4] for further discussion of the differences between the full-space models like ours with "smooth confining potentials" and the "hard-wall confining potentials" of [93].

Example 3.2.9. *The model (3.1.1) is a special case when $D_N = \mu \,\mathrm{Id}$ for some scalar $\mu > 0$. In our notation, this corresponds to $\mu_D = \delta_\mu$. Theorem 3.2.6 yields*

$$\Sigma^{\mathrm{tot}}(\delta_\mu, B''(0)) = \begin{cases} \frac{1}{2}\left(\frac{\mu^2}{B''(0)} - 1\right) - \log\left(\frac{\mu}{\sqrt{B''(0)}}\right) & \text{if } \mu \leqslant \mu_c := \sqrt{B''(0)} \\ & \text{(equivalently, if } \int \frac{\mu_D(\mathrm{d}t)}{t^2} \geqslant \frac{1}{B''(0)}\text{)}, \\ 0 & \text{if } \mu \geqslant \mu_c, \end{cases} \tag{3.2.13}$$

$$\Sigma^{\mathrm{min}}(\delta_\mu, B''(0)) = \begin{cases} \frac{1}{2}\left[-3 - \log\left(\frac{\mu^2}{B''(0)}\right) + 4 \cdot \frac{\mu}{\sqrt{B''(0)}} - \frac{\mu^2}{B''(0)}\right] & \text{if } \mu \leqslant \mu_c, \\ 0 & \text{if } \mu \geqslant \mu_c. \end{cases}$$

These recover results of [84, Equations (18-19)] and [94, Equation (81)], respectively. We also recover their results on decay near criticality, as one can check by hand that the behavior predicted by Theorem 3.2.8 (which gives $c_{\mathrm{tot}}(\delta_\mu) = \frac{1}{4\mu^4}$ and $c_{\mathrm{min}}(\delta_\mu) = \frac{1}{24\mu^6}$ here) is correct.

Example 3.2.10. *We give one more explicit example, namely when*

$$\mu_D(\mathrm{d}x) = \rho_{sc}^{(m,\sigma^2)}(\mathrm{d}x) = \frac{\sqrt{(4\sigma^2 - (x-m)^2)_+}}{2\pi\sigma^2}\,\mathrm{d}x$$

is the semicircle law of mean m and variance σ^2. (Notice we need $\mu_D(\mathrm{d}x)$ supported in $(0, \infty)$,

equivalently $m - 2\sigma > 0$, for the model to be well-defined.) In this case we have

$$\Sigma^{\mathrm{tot}}(\mu_D, B''(0)) = \begin{cases} \frac{m}{4\sigma^2}\left(\sqrt{m^2-4\sigma^2} - \frac{m}{1+2\frac{\sigma^2}{B''(0)}}\right) - \log\left(\frac{m+\sqrt{m^2-4\sigma^2}}{2\sqrt{B''(0)+\sigma^2}}\right) \\ \qquad \text{if } \int \frac{\mu_D(\mathrm{d}t)}{t^2} = \frac{-1+\frac{m}{\sqrt{m^2-4\sigma^2}}}{2\sigma^2} \geqslant \frac{1}{B''(0)}, \\ 0 \quad \text{if } \int \frac{\mu_D(\mathrm{d}t)}{t^2} \leqslant \frac{1}{B''(0)}, \end{cases}$$

$$\Sigma^{\mathrm{min}}(\mu_D, B''(0)) = \begin{cases} -1 + \frac{m(-m+\sqrt{m^2-4\sigma^2})}{4\sigma^2} - \frac{m^2+4\sigma^2-4m\sqrt{B''(0)+\sigma^2}}{2B''(0)} - \log\left(\frac{m+\sqrt{m^2-4\sigma^2}}{2\sqrt{B''(0)+\sigma^2}}\right) \\ \qquad \text{if } \int \frac{\mu_D(\mathrm{d}t)}{t^2} \geqslant \frac{1}{B''(0)}, \\ 0 \quad \text{if } \int \frac{\mu_D(\mathrm{d}t)}{t^2} \leqslant \frac{1}{B''(0)}. \end{cases}$$

As a consistency check, in the limit $\sigma \downarrow 0$ we obtain exactly the formulas (3.2.13) with μ replaced by m.

3.3 Stability of the Matrix Dyson Equation

In this section, our goal is to give general results on the stability of the Matrix Dyson Equation. For example, the MDE for GOE matrices is

$$\mathrm{Id} + \left(z\,\mathrm{Id} + \frac{1}{N}\,\mathrm{Tr}(M_N(z)) + \frac{1}{N}M_N(z)^T\right)M_N(z) = 0, \quad \mathrm{Im}\, M_N(z) > 0,$$

but $\frac{1}{N}M_N(z)^T$ should be thought of as an error, and it is more convenient to consider the unique solution $M'_N(z)$ to

$$\mathrm{Id} + \left(z\,\mathrm{Id} + \frac{1}{N}\,\mathrm{Tr}(M'_N(z))\right)M'_N(z) = 0, \quad \mathrm{Im}\, M'_N(z) > 0.$$

In this section we prove stability of such MDEs to conclude $\frac{1}{N}\,\mathrm{Tr}\, M_N(z) \approx \frac{1}{N}\,\mathrm{Tr}\, M'_N(z)$ for their respective unique solutions. Similar arguments have appeared in papers of Erdős and collaborators,

for example [6], but in more involved contexts where an exact deterministic solution of the MDE is compared to a random near-solution with small (random) error. Since we are interested in slightly different perturbations of the MDE, and only in the deterministic case, we adapt their arguments to give a short self-contained proof here.

Fix a sequence $(P_N)_{N=1}^{\infty}$ of positive integers (typically we take $P_N = N$ or P_N independent of N). It is known [106] that, whenever $\mathcal{S} : \mathbb{C}^{P_N \times P_N} \to \mathbb{C}^{P_N \times P_N}$ is a linear operator that is self-adjoint with respect to the inner product $\langle R, T \rangle = \operatorname{Tr}(R^* T)$ and that preserves the cone of positive-semi-definite matrices, and whenever $a(u) \in \mathbb{R}^{P_N \times P_N}$ is symmetric, the problem

$$-M^{-1}(u,z) = z\,\mathrm{Id} - a(u) + \mathcal{S}[M(u,z)] \quad \text{subject to} \quad \operatorname{Im} M(u,z) > 0 \tag{3.3.1}$$

has a unique solution $M(u,z) \in \mathbb{C}^{P_N \times P_N}$ for each $z \in \mathbb{H}$ and $u \in \mathbb{R}^m$, and

$$\|M(u,z)\| \leqslant \frac{1}{\eta}. \tag{3.3.2}$$

Fix two sequences $(\mathcal{S}_N)_{N=1}^{\infty}$ and $(\mathcal{S}'_N)_{N=1}^{\infty}$ of such operators and two sequences $(a_N(u))_{N=1}^{\infty}$ and $(a'_N(u))_{N=1}^{\infty}$ of such matrices (i.e., $\mathcal{S}_N$ and $\mathcal{S}'_N$ act on $\mathbb{C}^{P_N \times P_N}$, and $a_N(u), a'_N(u) \in \mathbb{R}^{P_N \times P_N}$), and consider the associated solutions:

$$\mathcal{S}_N \text{ and } a_N(u) \quad \text{induce} \quad M_N(u,z), \qquad \mathcal{S}'_N \text{ and } a'_N(u) \quad \text{induce} \quad M'_N(u,z).$$

In this section, our goal is to show that M_N and M'_N are close if $\mathcal{S}_N$ and $\mathcal{S}'_N$ are close and $a_N(u)$ and $a'_N(u)$ are close; we will use this to help identify μ_∞ for both of our models.

Lemma 3.3.1. *Suppose that, for some $\kappa > 0$,*

$$\sup_N \max(\|a_N(u)\|, \|a'_N(u)\|) \leqslant \kappa \max(1, \|u\|), \tag{3.3.3}$$

$$\|\mathcal{S}'_N\|_{hs\to\|\cdot\|} \leqslant \kappa, \tag{3.3.4}$$

$$\|\mathcal{S}_N - \mathcal{S}'_N\|_{\|\cdot\|\to\|\cdot\|} \leqslant \frac{\kappa}{N}, \tag{3.3.5}$$

$$\|a_N(u) - a'_N(u)\| \leqslant \frac{\kappa \max(1, \|u\|)}{N}. \tag{3.3.6}$$

If $0 < \gamma < \frac{1}{50}$, then for each R and each A there exists $\delta > 0$ with

$$\sup_{u\in B_R} \frac{1}{N} \int_{-A}^{A} \left|\operatorname{Tr}(M_N(u, E + \mathrm{i}N^{-\gamma})) - \operatorname{Tr}(M'_N(u, E + \mathrm{i}N^{-\gamma}))\right| \mathrm{d}E \leqslant \frac{1}{\delta} N^{-\delta}.$$

Proof. Notice that $M_N(u, z)$ almost solves the MDE (3.3.1) with $\mathcal{S} = \mathcal{S}'_N$ and $a(u) = a'_N(u)$; in fact,

$$-M_N(u,z)^{-1} = z\,\mathrm{Id} - a'_N(u) + \mathcal{S}'_N[M_N(u,z)] + \underbrace{(\mathcal{S}_N - \mathcal{S}'_N)[M_N(u,z)] + a'_N(u) - a_N(u)}_{=:d_N(u,z)},$$

and $d_N(u, z)$ is an error term in the sense that, if $u \in B_R$ (we take $R \geqslant 1$ without loss of generality), from (3.3.2), (3.3.5), and (3.3.6) we have

$$\|d_N(u,z)\| \leqslant \frac{\kappa}{N\eta} + \frac{\kappa \max(1, \|u\|)}{N} \leqslant \frac{\kappa(1 + \eta R)}{N\eta}. \tag{3.3.7}$$

We will apply standard stability theory of the MDE, which lets us conclude from this that M_N is close to M'_N. In fact, our goal is significantly easier than that in the literature, because our approximate solution to the MDE is deterministic. In the generality we need, this theory has been

developed in [6], and manipulations exactly like those preceding (4.25) there let us conclude

$$\begin{aligned}&\|M_N(u,z) - M_N'(u,z)\| \\ &\leqslant \|\mathscr{L}_N^{-1}(u,z)\|\|M_N'(u,z)\|\Big(\|d_N(u,z)\|\|M_N(u,z)\| + \|\mathcal{S}_N'\|\|M_N(u,z) - M_N'(u,z)\|^2\Big).\end{aligned} \tag{3.3.8}$$

Here $\mathscr{L}_N(u,z) : \mathbb{C}^{P_N \times P_N} \to \mathbb{C}^{P_N \times P_N}$ is the "stability operator"

$$\mathscr{L}_N(u,z)[T] = T - M_N'(u,z)\mathcal{S}_N'[T]M_N'(u,z),$$

which is invertible for every u and every z by [6, Lemma 3.7(i)]. Inserting the estimates (3.3.2), (3.3.4), and (3.3.7) into (3.3.8) yields

$$\|M_N(u,z) - M_N'(u,z)\| \leqslant \frac{\kappa}{\eta}\|\mathscr{L}_N^{-1}(u,z)\|\Big(\frac{1+\eta R}{N\eta^2} + \|M_N(u,z) - M_N'(u,z)\|^2\Big). \tag{3.3.9}$$

As usual, this quadratic inequality is fundamental to our strategy: We use it to show that $\|M_N - M_N'\|$ is small for very large η, then fix E and decrease η with a continuity argument. To make this bound useful, we import the following estimate on $\|\mathscr{L}^{-1}(u,z)\|$ from [6, (3.23), (3.22), Convention 3.5] combined with (3.3.2): There exists a constant C such that, for all u and z, we have

$$\|\mathscr{L}^{-1}(u,z)\| \leqslant C\Big(1 + \frac{1}{\eta^2} + \frac{\|M_N'(u,z)^{-1}\|^9}{\eta^{13}}\Big). \tag{3.3.10}$$

We use this estimate differently for $\eta \geqslant 1$ and $\eta \leqslant 1$, which are the two steps in the remainder of our argument.

Step 1 ($\eta \geqslant 1$): If $u \in B_R$ for some R (we take $R \geqslant 1$ without loss of generality), then taking norms directly in the MDE (3.3.1) and applying (3.3.2) and (3.3.3) yields

$$\|M_N'(u,z)^{-1}\| \leqslant |z| + \|a_N(u)\| + \|M_N'(u,z)\| \leqslant |z| + \kappa R + 1.$$

If $\eta \geqslant 1$, then $|z| \leqslant \eta\sqrt{1+E^2}$, so for any choice of $E_{\max}$ there exists a constant $C_{R,E_{\max}} =$

$C_{R,E_{\max},\kappa_1}$ such that

$$\sup_{|E|\leqslant E_{\max}\,,\,\eta\geqslant 1,\,u\in B_R} \frac{\|M_N'(u,z)^{-1}\|}{\eta} \leqslant C_{R,E_{\max}}\,.$$

Inserting this into (3.3.10) gives, for a new constant $\widetilde{C}_{R,E_{\max}} = \widetilde{C}_{R,E_{\max},\kappa_1}$,

$$\sup_{|E|\leqslant E_{\max}\,,\,\eta\geqslant 1,\,u\in B_R} \|\mathscr{L}^{-1}(u,z)\| \leqslant \widetilde{C}_{R,E_{\max}}\,. \tag{3.3.11}$$

Now fix $|E|\leqslant E_{\max}$, and consider the functions $f_N:(0,\infty)\to\mathbb{R}$ and $g_N^{\pm}:[1,\infty)\to\mathbb{R}$ defined by

$$\begin{aligned} f_N(\eta) &= f_{N,u}(\eta) = \|M_N(u,E+\mathrm{i}\eta) - M_N'(u,E+\mathrm{i}\eta)\|,\\ g_N^{\pm}(\eta) &= \frac{\eta}{2\kappa\widetilde{C}_{R,E_{\max}}}\left(1\pm\sqrt{1-\frac{4\kappa^2(\widetilde{C}_{R,E_{\max}})^2(1+\eta R)}{N\eta^4}}\right). \end{aligned}$$

(The functions $g_N^{\pm}(\eta)$ are well-defined if $N\geqslant 4(\widetilde{C}_{R,E_{\max}})^2(1+R)$.) The quadratic inequality (3.3.9) with the estimate (3.3.11) inserted give, for all $N\geqslant 4(\widetilde{C}_{R,E_{\max}})^2$ and all $\eta\geqslant 1$,

$$f_N(\eta)\in[0,g_N^-(\eta)]\cup[g_N^+(\eta),\infty).$$

If $\eta\geqslant\max\left\{1,\sqrt{4\kappa\widetilde{C}_{R,E_{\max}}}\right\}$, then the crude bound (3.3.2) yields

$$f_N(\eta)\leqslant\frac{\eta}{2\widetilde{C}_{R,E_{\max}}}<g_N^+(\eta),$$

so that $f_N(\eta)\leqslant g_N^-(\eta)$. But since $M_N(u,z)$ and $M_N'(u,z)$ are both holomorphic matrix-valued functions of z [106], we know that $f_N(\eta)$ is a continuous function of η. Since $g_N^-(\eta)<g_N^+(\eta)$ for all $\eta>1$, we have $f_N(\eta)\leqslant g_N^-(\eta)$ down to $\eta=1$. Notice that this is uniform in $u\in B_R$.

Step 2 ($\eta\leqslant 1$): Now we estimate

$$\|M_N'(u,z)^{-1}\|\leqslant|z|+\kappa R+\frac{1}{\eta}\leqslant\frac{C'_{R,E_{\max}}}{\eta}$$

for some $C'_{R,E_{\max}} = C'_{R,E_{\max},\kappa_1}$. Inserting this and (3.3.2) shows that, for some $C''_{R,E_{\max}} = C''_{R,E_{\max},\kappa_1}$, we have

$$\sup_{|E|\leqslant E_{\max}\,,\,\eta\leqslant 1,\,u\in B_R} \frac{\|\mathscr{L}^{-1}(u,z)\|}{\eta^{-22}} \leqslant C''_{R,E_{\max}}\,. \tag{3.3.12}$$

Now fix $|E| \leqslant E_{\max}$ and consider the functions $h_N^{\pm} : [N^{-1/50}, 1] \to \mathbb{R}$ defined by

$$h_N^{\pm}(\eta) = \frac{\eta^{23}}{2\kappa C''_{R,E_{\max}}}\left(1 \pm \sqrt{1 - \frac{4\kappa^2 (C''_{R,E_{\max}})^2(1+\eta R)}{N\eta^{48}}}\right).$$

As above, the quadratic inequality (3.3.9) with the estimate (3.3.12) inserted give, for all N and all $\eta \leqslant 1$,

$$f_N(\eta) \in [0, h_N^-(\eta)] \cup [h_N^+(\eta), \infty).$$

But when $\eta = 1$ and $N \geqslant 4\kappa^2 C''_{R,E_{\max}} \widetilde{C}_{R,E_{\max}}(1+R)$ we have (using $1 - \sqrt{1-x} \leqslant x$)

$$f_N(1) \leqslant g_N^-(1) \leqslant \frac{2\kappa \widetilde{C}_{R,E_{\max}}(1+R)}{N} \leqslant \frac{1}{2\kappa C''_{R,E_{\max}}} < h_N^+(1),$$

so the same continuity argument as above gives

$$f_N(\eta) \leqslant h_N^-(\eta) \leqslant \frac{2\kappa C''_{R,E_{\max}}(1+R)}{N\eta^{25}}. \tag{3.3.13}$$

Again, this is uniform over $u \in B_R$ and $|E| \leqslant E_{\max}$.

To show the statement of the lemma, given R, $0 < \gamma < \frac{1}{50}$, and A, we choose $E_{\max} = A$ above; then applying (3.3.13) yields

$$\sup_{u\in B_R} \frac{1}{N}\int_{-A}^{A} \left|\mathrm{Tr}(M_N(u, E + \mathrm{i}N^{-\gamma})) - \mathrm{Tr}(M'_N(u, E+\mathrm{i}N^{-\gamma}))\right| \mathrm{d}E \leqslant (4A\kappa C''_{R,A}(1+R))N^{25\gamma-1}.$$

This holds for N large enough depending on κ, R, and A. □

3.4.1 Establishing the variational formula. In this subsection we establish a variational formula for complexity, given over $\mathbb{R}^{L^d}$. In the next subsection we analyze this variational problem, first by reducing it to a variational problem over $\mathbb{R}$ and then by relating it to the variational problem we analyze in depth for the soft-spins model.

In this subsection, we frequently reference notation and results in the companion paper [35]. Let

$$J = 2\sqrt{B''(0)}$$

which will be an important scaling factor. For each $u \in \mathbb{R}^{L^d}$, define

$$a(u) = (-t_0\Delta + \operatorname{diag}(u) + \mu_0 \operatorname{Id}_{L^d \times L^d}) \in \mathbb{R}^{L^d \times L^d}, \tag{3.4.1}$$

and for each $z \in \mathbb{H}$ let $m_\infty(u,z) = (m_\infty(u,z)_1, \dots, m_\infty(u,z)_{L^d}) \in \mathbb{C}^{L^d}$ be the unique solution to

$$m_\infty(u,z) = \operatorname{diag}[(a(u) - J^2 m_\infty(u,z) - z \operatorname{Id})^{-1}] \quad \text{such that} \quad \operatorname{Im} m_\infty(u,z) > 0 \text{ componentwise.} \tag{3.4.2}$$

(Recall we identify vectors with diagonal matrices; we will prove existence and uniqueness during the proof using the methods of Erdős and co-authors.) Let $\mu_\infty(u)$ (which also depends on L, d, t_0, and μ_0) be the measure whose Stieltjes transform is given by

$$\int \frac{\mu_\infty(u, \mathrm{d}s)}{s - z} = \frac{1}{L^d} \sum_{i=1}^{L^d} m_\infty(u,z)_i.$$

Theorem 3.4.1. *The probability measure $\mu_\infty(u)$ admits a bounded, compactly supported density*

$\mu_\infty(u,\cdot)$ with respect to Lebesgue measure, and

$$\begin{aligned}\Sigma(\mu_0) = \Sigma(\mu_0, L, d, t_0) &:= \lim_{N\to\infty} \frac{1}{NL^d} \log \mathbb{E}[\mathcal{N}_{tot}] \\ &= -\frac{1}{L^d}\log(\det(\mu_0 \,\mathrm{Id} - t_0\Delta)) + \sup_{u\in\mathbb{R}^{L^d}} \left\{ \int \log|\cdot| \mathrm{d}\mu_\infty(u,\cdot) - \frac{\|u\|^2}{2J^2L^d} \right\}.\end{aligned} \tag{3.4.3}$$

Furthermore, if we define the set

$$\mathcal{G} = \{u \in \mathbb{R}^{L^d} : \mu_\infty(u)((-\infty, 0)) = 0\} \tag{3.4.4}$$

of u values whose corresponding measures $\mu_\infty(u)$ are supported in the right half-line, we have

$$\begin{aligned}\Sigma_{st}(\mu_0) = \Sigma(\mu, L, d, t_0) &:= \limsup_{N\to\infty} \frac{1}{NL^d} \log \mathbb{E}[\mathcal{N}_{st}] \\ &= -\frac{1}{L^d}\log(\det(\mu_0 \,\mathrm{Id} - t_0\Delta)) + \sup_{u\in\mathcal{G}} \left\{ \int \log|\cdot| \mathrm{d}\mu_\infty(u,\cdot) - \frac{\|u\|^2}{2J^2L^d} \right\}.\end{aligned} \tag{3.4.5}$$

The suprema in (3.4.3) and (3.4.5) are achieved (possibly not uniquely).

We first build the relevant block matrix. With $a(u)$ as in (3.4.1), let

$$A_N(u) = a(u) \otimes \mathrm{Id}_{N\times N}\,.$$

For each N, let $(X_i)_{i=1}^{L^d}$ be a collection of L^d independent $N \times N$ matrices, each distributed as J times a GOE matrix, with the normalization $\mathbb{E}[((X_i)_{jk})^2] = J^2\frac{1+\delta_{jk}}{N}$. Let

$$\begin{aligned}W_N &= \sum_{i=1}^{L^d} E_{ii} \otimes X_i, \\ H_N(u) &= A_N(u) + W_N.\end{aligned}$$

This matrix is in the class of "block-diagonal Gaussian matrices" studied in [35, Corollary 1.9]. It appears naturally in the Kac-Rice formula, but we also introduce a slight modification that is easier

to work with. Let

$$\widetilde{W_N} = \left(\mathbf{1} - \frac{1}{\sqrt{2}}\,\mathrm{Id}\right) \odot W_N,$$
$$\widetilde{H_N(u)} = A_N(u) + \widetilde{W_N},$$

where $\mathbf{1}$ is the matrix of all ones and $\odot$ is the entrywise product, i.e., $\widetilde{W_N}$ is just W_N rescaled to make the variances J^2/N both on and off the diagonal, coupled appropriately with W_N.

Now we simplify the MDE. It is known [106] that, whenever $\mathcal{S} : \mathbb{C}^{L^d \times L^d} \to \mathbb{C}^{L^d \times L^d}$ is a linear operator that is self-adjoint with respect to the inner product $\langle R, T \rangle = \mathrm{Tr}(R^* T)$ and that preserves the cone of positive-semi-definite matrices, the problem

$$-M^{-1}(u,z) = z\,\mathrm{Id} - a(u) + \mathcal{S}[M(u,z)] \quad \text{subject to} \quad \mathrm{Im}\, M(u,z) > 0 \tag{3.4.6}$$

has a unique solution $M(u,z) \in \mathbb{C}^{L^d \times L^d}$ for each $z \in \mathbb{H}$ and $u \in \mathbb{R}^{L^d}$. We will consider this problem with two choices of operator $\mathcal{S}$:

$$\mathcal{S}_N[T] = J^2 \frac{N+1}{N}\,\mathrm{diag}(T) \quad \text{induces} \quad M_N(u,z),$$
$$\mathcal{S}_\infty[T] = J^2\,\mathrm{diag}(T) \quad \text{induces} \quad M_\infty(u,z).$$

Write $\mathscr{S}_i$ (respectively, $\widetilde{\mathscr{S}_i}$) for the "stability operators" of [35, (1.15)] corresponding to the matrix $H_N(u)$ (respectively, $\widetilde{H_N(u)}$). Then we can compute

$$\mathscr{S}_i[\mathbf{r}] = J^2 \sum_{k=1}^{N} \frac{1+\delta_{ik}}{N}\,\mathrm{diag}(r_k), \qquad \widetilde{\mathscr{S}_i}[\mathbf{r}] = \frac{J^2}{N} \sum_{k=1}^{N} \mathrm{diag}(r_k).$$

Thus the choices $\mathbf{m}(u,z) = (M_N(u,z), \ldots, M_N(u,z))$ and $\widetilde{\mathbf{m}}(u,z) = (M_\infty(u,z), \ldots, M_\infty(u,z))$ exhibit solutions to the block MDE [35, (1.16)] for the matrices $H_N(u)$ and $\widetilde{H_N(u)}$, respectively.

That is, the measure $\mu_N(u)$ that appears in the local laws for $H_N(u)$ has Stieltjes transform

$$\int \frac{\mu_N(u, \mathrm{d}s)}{s - z} = \frac{1}{L^d} \operatorname{Tr}(M_N(u, z)),$$

and the measure that appears in the local laws for $\widetilde{H_N(u)}$ is actually independent of N: We call it $\mu_\infty(u)$, and its Stieltjes transform is given by

$$\int \frac{\mu_\infty(u, \mathrm{d}s)}{s - z} = \frac{1}{L^d} \operatorname{Tr}(M_\infty(u, z)).$$

Lemma 3.4.2. *Both $\mu_N(u)$ and $\mu_\infty(u)$ admit densities $\mu_N(u, \cdot)$ and $\mu_\infty(u, \cdot)$ with respect to Lebesgue measure, and*

$$\sup_{u \in \mathbb{R}^m, z \in \mathbb{H}, N \in \mathbb{N}} \max\{\|M_N(u, z)\|, \|M_\infty(u, z)\|, \|\mu_N(u, \cdot)\|_{L^\infty}, \|\mu_\infty(u, \cdot)\|_{L^\infty}\} \leqslant \sqrt{L^d}/J.$$

Proof. The following arguments are due to László Erdős and Torben Krüger. We prove the result for M_N and μ_N; the proofs for M_∞ and μ_∞ are similar. By taking the imaginary part of (3.4.6) and multiplying on the left by $M_N(u, z)^*$ and on the right by $M_N(u, z)$, then taking the diagonal of both sides, we obtain

$$\begin{aligned} \operatorname{Im} \operatorname{diag}(M_N(u, z)) &= \operatorname{diag}\Big(M_N(u, z)^* \Big(\eta + J^2 \frac{N+1}{N} \operatorname{Im}(\operatorname{diag}(M_N(u, z))) \Big) M_N(u, z) \Big) \\ &= F_N(u, z) \Big(\eta + J^2 \frac{N+1}{N} \operatorname{Im}(\operatorname{diag}(M_N(u, z))) \Big), \end{aligned} \tag{3.4.7}$$

where $F_N(u, z) \in \mathbb{R}^{L^d \times L^d}$ is given elementwise by $F_N(u, z)_{ij} = |(M_N(u, z))_{ij}|^2$. By transposing the MDE (3.4.6) and using the fact that $a(u)$ is symmetric, we see that $M_N(u, z)$ is symmetric (but not Hermitian) as well. Hence $F_N(u, z)$ is a real symmetric matrix with nonnegative entries. The inner product of (3.4.7) with the Perron-Frobenius eigenvector of F_N then gives $\|F_N\| \leqslant \frac{1}{J^2} \frac{N}{N+1}$,

since $\mathrm{Im}(\mathrm{diag}(M_N(u,z))$ has all positive components. Thus

$$\|M_N(u,z)\|^2 \leqslant \mathrm{Tr}(M_N(u,z)^* M_N(u,z)) = \langle 1, F_N(u,z)1\rangle \leqslant \frac{1}{J^2}\frac{N}{N+1}L^d.$$

Now for any interval $[a,b]$ we have

$$\mu_N(u)([a,b]) + \frac{\mu_N(u)(\{a\}) + \mu_N(u)(\{b\})}{2} = \lim_{\eta\downarrow 0}\frac{1}{\pi}\int_a^b \mathrm{Im}\,\frac{1}{L^d}\,\mathrm{Tr}(M_N(u,E+\mathrm{i}\eta))\,\mathrm{d}E \leqslant (b-a)\frac{\sqrt{L^d}}{J\pi}.$$

By standard continuity arguments we extend this to $\mu_N(u)(A) \leqslant |A|\frac{\sqrt{L^d}}{J\pi}$ for any Borel set A; this implies that μ_N is absolutely continuous with respect to Lebesgue measure with a density that is pointwise bounded by $\frac{\sqrt{L^d}}{J\pi}$. □

Lemma 3.4.3. *For every R, there exists $\varepsilon > 0$ such that*

$$\sup_{u\in B_R} \mathrm{W}_1(\mathbb{E}[\hat{\mu}_{H_N(u)}], \mu_\infty(u)) \leqslant N^{-\varepsilon}. \tag{3.4.8}$$

Proof. First, note that

$$\sup_N \|A_N(u)\| = \|a(u)\otimes \mathrm{Id}\,\| = \|a(u)\| \leqslant t\|\Delta\| + \|u\| + \mu < \infty. \tag{3.4.9}$$

Along with Lemma 3.4.2, this verifies the assumptions of [35, Corollary 1.9], the proof of which shows that (3.4.8) holds with $\mu_\infty(u)$ replaced by $\mu_N(u)$ (the result is locally uniform in u since all the assumptions are). To compare $\mu_N(u)$ and $\mu_\infty(u)$, we use the result of Lemma 3.3.1 (with the choices $P_N \equiv L^d$, $\mathcal{S}'_N = \mathcal{S}_\infty$ as above) and follow the proof of [35, Proposition 3.1]. □

Lemma 3.4.4. *There exists $C > 0$ with*

$$\mathbb{E}[|\det(H_N(u))|] \leqslant (C\max(\|u\|,1))^N.$$

Proof. Deterministically,

$$|\det(H_N(u))| \leqslant \|H_N(u)\|^N \leqslant (\|W_N\| + \|A_N(u)\|)^N \leqslant (2\|W_N\|)^N + (2\|A_N(u)\|)^N.$$

In the proof of [35, Corollary 1.9], we showed $\mathbb{P}(\|W_N\| \geqslant t) \leqslant e^{-cN\max(0,t-C)}$ for some constants c, C, which implies $\mathbb{E}[\|W_N\|^N] \leqslant e^{CN}$. With the estimate on $\|A_N(u)\|$ from (3.4.9), this suffices. □

Lemma 3.4.5. *For every R and every $\varepsilon > 0$, we have*

$$\lim_{N\to\infty} \frac{1}{N\log N} \log\left[\sup_{u\in B_R} \mathbb{P}(d_{\mathrm{BL}}(\hat{\mu}_{H_N(u)}, \mu_\infty(u)) > \varepsilon)\right] = -\infty.$$

Proof. The laws of the entries of $\sqrt{N}H_N(u)$ satisfy the log-Sobolev inequality with a uniform constant, since they are Gaussians. (If they are degenerate, we recall that a delta mass satisfies log-Sobolev with any constant.) This is true uniformly over $u \in \mathbb{R}^{L^d}$, since u only affects the mean, and translating measures preserves log-Sobolev with the same constant. Hence [103, Theorem 1.5] give, for some constants C_1 and C_2,

$$\sup_{u\in\mathbb{R}^{L^d}} \mathbb{P}(d_{\mathrm{BL}}(\hat{\mu}_{H_N(u)}, \mathbb{E}[\hat{\mu}_{H_N(u)}]) > \varepsilon) \leqslant \frac{C_1}{\varepsilon^{3/2}} \exp\left(-\frac{C_2}{2L^d} N^2\varepsilon^5\right).$$

To relate $\mathbb{E}[\hat{\mu}_{H_N(u)}]$ to $\mu_\infty(u)$, we use Lemma 3.4.3. □

Lemma 3.4.6. *For every $\varepsilon > 0$ and $R > 0$, we have*

$$\lim_{N\to\infty} \inf_{u\in B_R} \mathbb{P}(\mathrm{Spec}(H_N(u)) \subset [\mathtt{l}(\mu_\infty(u)) - \varepsilon, \mathtt{r}(\mu_\infty(u)) + \varepsilon]) = 1. \tag{3.4.10}$$

and in fact the extreme eigenvalues of $H_N(u)$ converge in probability to the endpoints of $\mu_\infty(u)$.

Proof. The local law [6, Theorem 2.4, Remark 2.5(v)] tells us that, for every ε and R, there exists $C_{\varepsilon,R}$ such that

$$\inf_{u\in B_R} \mathbb{P}\left(\mathrm{Spec}(\widetilde{H_N(u)}) \subset \left[\mathtt{l}(\mu_\infty(u)) - \frac{\varepsilon}{2}, \mathtt{r}(\mu_\infty(u)) + \frac{\varepsilon}{2}\right]\right) \geqslant 1 - \frac{C_{\varepsilon,R}}{N^{100}}. \tag{3.4.11}$$

(We can take the infimum over $u \in B_R$ because the local-law estimates are uniform over all models possessing the same "model parameters," see the remarks just before Theorem 2.4 there. Our model parameters depend on u but can be taken uniformly over $u \in B_R$, for example because $\sup_{u \in B_R} \|A_N(u)\| < \infty$.)

Notice that

$$\Delta_N = H_N(u) - \widetilde{H_N(u)} = W_N - \widetilde{W_N}$$

is a diagonal matrix with independent Gaussian entries of variance J^2/N that does not depend on u. Thus

$$\sup_{u \in \mathbb{R}^{L^d}} \mathbb{P}\left(\left| \lambda_{\min}\left(H_N(u)\right) - \lambda_{\min}\left(\widetilde{H_N(u)}\right) \right| \geqslant \frac{\varepsilon}{2} \right) \leqslant \mathbb{P}\left(\|\Delta_N\| \geqslant \frac{\varepsilon}{2} \right) \leqslant \frac{2J\sqrt{N}L^d}{\varepsilon} e^{-N\varepsilon^2/(8J^2)} \quad (3.4.12)$$

and similarly for $\lambda_{\min}$. Since

$$\begin{aligned} &\mathbb{P}(\lambda_{\max}\left(H_N(u)\right) \geqslant \mathtt{r}(\mu_\infty(u)) + \varepsilon) \\ &\leqslant \mathbb{P}\left(\lambda_{\max}\left(\widetilde{H_N(u)}\right) \geqslant \mathtt{r}(\mu_\infty(u)) + \frac{\varepsilon}{2} \right) + \mathbb{P}\left(\left| \lambda_{\max}\left(H_N(u)\right) - \lambda_{\max}\left(\widetilde{H_N(u)}\right) \right| \geqslant \frac{\varepsilon}{2} \right), \end{aligned}$$

and similarly for $\lambda_{\min}$, (3.4.11) and (3.4.12) imply (3.4.10).

For the other inequality, namely that $\liminf_{N\to\infty} \lambda_{\max}\left(H_N(u)\right) \geqslant \mathtt{r}(\mu_\infty(u))$ in probability, we note that $\hat{\mu}_{H_N(u)}$ concentrates about $\mu_\infty(u)$ in the sense of Lemma 3.4.5. The smallest eigenvalue is handled similarly. □

Lemma 3.4.7. *With $\mathcal{G}_{+\varepsilon}$ as defined in [35, (4.5)] and $\mathcal{G}$ as defined in (3.4.4), we have that each $\mathcal{G}_{+\varepsilon}$ is convex, that $\mathcal{G}_{+1}$ has positive measure, and that*

$$\overline{\bigcup_{\varepsilon > 0} \mathcal{G}_{+\varepsilon}} = \mathcal{G}.$$

Proof. Whenever $u, v \in \mathbb{R}^{L^d}$ and $t \in [0,1]$, one can check $H_N(tu+(1-t)v) = tH_N(u)+(1-t)H_N(v)$;

thus

$$\lambda_{\min}\left(H_N(tu+(1-t)v)\right) \geqslant t\lambda_{\min}\left(H_N(u)\right) + (1-t)\lambda_{\min}\left(H_N(v)\right)$$

almost surely, and by letting $N \to \infty$ and applying the convergence in probability of Lemma 3.4.6 we obtain $\mathtt{l}(\mu_\infty(tu+(1-t)v)) \geqslant t\mathtt{l}(\mu_\infty(u)) + (1-t)\mathtt{l}(\mu_\infty(v))$. Hence each $\mathcal{G}_{+\varepsilon}$ is convex.

Since $-t\Delta$ and $\mu\,\mathrm{Id}$ are positive semidefinite,

$$\lambda_{\min}\left(A_N(u)\right) = \lambda_{\min}\left(a(u)\otimes \mathrm{Id}\right) = \lambda_{\min}\left(a(u)\right) = \lambda_{\min}\left(-t\Delta + \mathrm{diag}(u) + \mu\,\mathrm{Id}\right) \geqslant \min(u_1,\dots,u_{L^d}).$$

On the other hand,

$$\lambda_{\min}\left(W_N\right) = \lambda_{\min}\left(\sum_{i=1}^{L^d} E_{ii}\otimes X_i\right) = \min_{i=1}^{L^d}(\lambda_{\min}\left(X_i\right))$$

which tends almost surely to $-2J$ in our normalization. Thus

$$\liminf_{N\to\infty} \lambda_{\min}\left(H_N(u)\right) \geqslant \min(u_1,\dots,u_{L^d}) - 2J.$$

which, combined with the convergence in probability of Lemma 3.4.6, shows that $\mathcal{G}_{+1}$ has positive measure.

Finally, we note that the inclusion $\cup_{\varepsilon>0}\mathcal{G}_{+\varepsilon} \subset \mathcal{G}$ is clear, and that $\mathcal{G}$ is closed by [35, Lemma 4.6]. To show the reverse inclusion, write $\vec{1} \in \mathbb{R}^{L^d}$ for the vector of all ones; then it is easy to check $H_N(u+\delta\vec{1}) = H_N(u) + \delta\,\mathrm{Id}$, so by the convergence in probability of Lemma 3.4.6 we have $\mathtt{l}(\mu_\infty(u+\delta\vec{1})) = \mathtt{l}(\mu_\infty(u)) + \delta$. This completes the proof. □

Proposition 3.4.8. *We have*

$$\lim_{N\to\infty} \frac{1}{NL^d} \log \int_{\mathbb{R}^{L^d}} e^{-N\frac{\|u\|^2}{2J^2}}\,\mathbb{E}[|\det(H_N(u))|]\,\mathrm{d}u = \sup_{u\in\mathbb{R}^{L^d}} \left\{ \int \log|\lambda|\mu_\infty(u,\lambda)\,\mathrm{d}\lambda - \frac{\|u\|^2}{2J^2L^d} \right\} \quad (3.4.13)$$

and

$$\limsup_{N\to\infty} \frac{1}{NL^d} \log \int_{\mathbb{R}^{L^d}} e^{-N\frac{\|u\|^2}{2J^2}} \mathbb{E}[|\det(H_N(u))|\mathbf{1}_{H_N(u)\geqslant 0}]\,\mathrm{d}u = \sup_{u\in\mathcal{G}}\left\{ \int \log|\lambda|\mu_\infty(u,\lambda)\,\mathrm{d}\lambda - \frac{\|u\|^2}{2J^2L^d}\right\} \tag{3.4.14}$$

where $\mathcal{G}$ is defined in (3.4.4).

Proof. For (3.4.13), we apply [35, Theorem 4.1] with $\alpha = \frac{1}{2J^2L^d}$, $p = 0$, and $\mathfrak{D} = \mathbb{R}^{L^d}$. (Technically, we are applying this theorem with N there replaced by NL^d here, which is the size of H_N; this is why α is $\frac{1}{2J^2L^d}$ and not $\frac{1}{2J^2}$.) We checked the conditions of this theorem in [35, Corollary 1.9] and Lemmas 3.4.3 and 3.4.4. (All the results are locally uniform in u because all the parameters of the random matrices are.) For (3.4.14), we apply [35, Theorem 4.5] with $\alpha = \frac{1}{2J^2L^d}$, $p = 0$, and $\mathfrak{D} = \mathbb{R}^{L^d}$. We checked the conditions for this result in Lemmas 3.4.5, 3.4.6, and 3.4.7. □

Proof of Theorem 3.4.1. Kac-Rice computations in [88] show, exactly for finite N,

$$\begin{aligned}
\frac{1}{NL^d}\log\mathbb{E}[\mathcal{N}_{\mathrm{tot}}] &= -\frac{1}{L^d}\log(\det(u_0 - t_0\Delta)) + \frac{1}{NL^d}\log\int_{\mathbb{R}^{L^d}} \frac{e^{-N\frac{\|u\|^2}{2J^2}}}{(\sqrt{2\pi J^2/N})^{L^d}}\mathbb{E}[|\det(H_N(u))|]\,\mathrm{d}u,\\
\frac{1}{NL^d}\log\mathbb{E}[\mathcal{N}_{\mathrm{st}}] &= -\frac{1}{L^d}\log(\det(u_0 - t_0\Delta))\\
&\quad + \frac{1}{NL^d}\log\int_{\mathbb{R}^{L^d}} \frac{e^{-N\frac{\|u\|^2}{2J^2}}}{(\sqrt{2\pi J^2/N})^{L^d}}\mathbb{E}[|\det(H_N(u))|\mathbf{1}_{H_N(u)\geqslant 0}]\,\mathrm{d}u,
\end{aligned}$$

where $H_N(u)$ is as above. Then we apply the above Proposition. □

3.4.2 Analyzing the variational formula. The following concavity property is the key reason the complexity thresholds can be calculated explicitly, from the variational formulas appearing in the previous section.

Proposition 3.4.9. *The function*

$$\mathcal{S}[u] = \int_{\mathbb{R}} \log|\lambda|\mu_\infty(u,\lambda)\,\mathrm{d}\lambda - \frac{\|u\|^2}{2J^2L^d}$$

is concave.

Proof. We assume $d = 1$ below, the general case requiring only notational change of L into L^d. The MDE for our problem, namely (3.4.6) with the choice $\mathcal{S}[T] = J^2 \operatorname{diag}(T)$, has a matrix solution $M(u,z) = M_\infty(u,z)$ (we now drop the ∞ to save notation). The problem can be reduced to a vector MDE for $\vec{m}(u,z) = \operatorname{diag}(M(u,z)) =: (m_1(u,z), \dots, m_L(u,z))$ by taking the diagonal of both sides of (3.4.6). (In fact, $M(u,z)$ can be reconstructed from knowledge of $\vec{m}(u,z)$ via (3.4.6).) The diagonal MDE takes the form

$$-\operatorname{diag}(m_1, \dots, m_L) = \operatorname{diag}[(z - \mu + t\Delta - \operatorname{diag}(u_1, \dots, u_L) + J^2 \operatorname{diag}(m_1, \dots, m_L))^{-1}]. \quad (3.4.15)$$

We denote $\partial_k = \partial_{u_k}$, and write $m = \frac{1}{L}\sum_1^L m_i$ for the Stieltjes transform of μ_∞. The first essential observation is

$$\frac{\mathrm{d}}{\mathrm{d}u_k}(-Lm) = \frac{\mathrm{d}}{\mathrm{d}z} m_k. \quad (3.4.16)$$

Indeed, for any invertible matrix B, we have

$$(B^{-1})_{kk} = \operatorname{Tr}(B^{-1}|e_k\rangle\langle e_k|) = \partial_{u=0} \log\det(\mathrm{Id} + uB^{-1}|e_k\rangle\langle e_k|) = \partial_{u=0} \log\det(B + u|e_k\rangle\langle e_k|),$$

so that, denoting $B(z,u,\vec{m}) = z - \mu + t\Delta - \operatorname{diag}(u_1, \dots, u_L) + J^2 \operatorname{diag}(m_1, \dots, m_L)$, we have

$$\frac{\mathrm{d}}{\mathrm{d}u_k} \log\det B = \partial_k \log\det B + J^2 \sum_j \partial_{m_j} \log\det B \cdot \partial_k m_j = m_k - J^2 \sum_j m_j \partial_k m_j,$$
$$\frac{\mathrm{d}}{\mathrm{d}z} \log\det B = \partial_z \log\det B + J^2 \sum_j \partial_{m_j} \log\det B \cdot \partial_z m_j = -Lm - J^2 \sum_j m_j \partial_z m_j,$$

i.e.

$$m_k = \frac{\mathrm{d}}{\mathrm{d}u_k}(\log\det B + \frac{J^2}{2}\sum_j m_j^2),$$
$$-Lm = \frac{\mathrm{d}}{\mathrm{d}z}(\log\det B + \frac{J^2}{2}\sum_j m_j^2).$$

We conclude that $\frac{\mathrm{d}}{\mathrm{d}u_k}(-Lm) = \frac{\mathrm{d}}{\mathrm{d}z}m_k$.

With (3.4.16), we obtain

$$\begin{aligned}\partial_k \int \log|\lambda|\mu_\infty(u,\lambda) &= \frac{1}{\pi L}\int \log|\lambda|\partial_\lambda \operatorname{Im} m_k(\lambda+\mathrm{i}0^+) \\ &= -\frac{1}{\pi L}\int \frac{1}{\lambda}\operatorname{Im} m_k(\lambda+\mathrm{i}0^+) = -\frac{1}{L}\operatorname{Re} m_k(\mathrm{i}0^+).\end{aligned} \tag{3.4.17}$$

Now, from (3.4.15), we obtain

$$\partial_k(m_1,\dots,m_L) = M(E_k + J^2\operatorname{diag}(\partial_k m_1,\dots,\partial_k m_L))M = R(J^2\partial_k(m_1,\dots,m_L)+e_k)^T$$

where the matrix E_k (respectively, the vector e_k) is 0's except 1 at position (k,k) (respectively, position k), and where $R = R(u,z)$ is the linear operator defined by $(Rv)_i = \sum(M_{ij})^2 v_j$, with $M = M(u,z)$ the MDE solution matrix. Thus we have

$$(1-J^2R)(\partial_k\vec{m}) = R(e_k)^T,$$

from which

$$\partial_k\vec{m} = \frac{1}{J^2}(1-J^2R)^{-1}(J^2R-1+1)(e_k)^T = \frac{1}{J^2}(-\operatorname{Id}+(1-J^2R)^{-1})(e_k)^T.$$

Taking the jth component of both sides, we get the scalar equation

$$\partial_k m_j = \frac{1}{J^2}(-\operatorname{Id}+(1-J^2R)^{-1})_{jk}.$$

Together with (3.4.17), this gives

$$\nabla_u^2\int \log|\lambda-\mu|\mu_\infty(u,\lambda) = \frac{1}{J^2L}(\operatorname{Id}-\operatorname{Re}[(1-J^2R)^{-1}]). \tag{3.4.18}$$

Lemma 3.4.10 below, due to László Erdős and Torben Krüger, shows that $\operatorname{Re}[(1-J^2R)^{-1}] \geqslant 0$ in the

sense of quadratic forms. Along with (3.4.18), this gives concavity of $\int_{\mathbb{R}} \log|\lambda| \mu_\infty(u,\lambda)\,\mathrm{d}\lambda - \frac{\|u\|^2}{2J^2L^d}$ in u. □

Lemma 3.4.10. *For each $u \in \mathbb{R}^{L^d}$ and each $z \in \mathbb{H}$, let $R(u,z) \in \mathbb{C}^{L^d \times L^d}$ be defined elementwise by*

$$R(u,z)_{jk} = (M(u,z)_{jk})^2 =: e^{\mathrm{i}\theta(u,z)_{jk}} |M(u,z)_{jk}|^2,$$

where $M(u,z) = M_\infty(u,z)$. Then for every $u \in \mathbb{R}^{L^d}$, every $z \in \mathbb{H}$, and every nonzero vector v we have

$$\mathrm{Re}\Big\langle v, (\mathrm{Id} - J^2 R(u,z))v \Big\rangle > 0.$$

In particular, for any w, written as $(1 - J^2R)v$, we have

$$\mathrm{Re}\langle w, (1-J^2R)^{-1} w\rangle = \mathrm{Re}\langle (1-J^2R)v, v\rangle > 0.$$

Proof. Consider the matrix $F(u,z) \in \mathbb{R}^{L^d \times L^d}$ defined elementwise by $F(u,z)_{jk} = |M(u,z)_{jk}|^2$. The proof of Lemma 3.4.2 shows that $\sup_{u \in \mathbb{R}^{L^d}, z \in \mathbb{H}} \|F(u,z)\| \leqslant \frac{1}{J^2}$. Given $v \in \mathbb{C}^{L^d}$, write $|v| = (|v_1|, \dots, |v_{L^d}|)$; then

$$\begin{aligned}
\mathrm{Re}\Big\langle v, (\mathrm{Id} - J^2 R(u,z))v \Big\rangle &\geqslant \mathrm{Re} \sum_{j=1}^{L^d} \Big(1 - J^2 |M(u,z)_{jj}|^2 e^{\mathrm{i}\theta(\xi,z)_{jj}}\Big)|v_j|^2 - J^2 \sum_{j \neq k} |M(u,z)_{jk}|^2 |v_j v_k| \\
&= \sum_{j=1}^{L^d} \Big(1 - J^2 |M(u,z)_{jj}|^2 \cos(\theta(u,z)_{jj})\Big)|v_j|^2 - J^2 \sum_{j \neq k} |M(u,z)_{jk}|^2 |v_j v_k| \\
&= \Big\langle |v|, \Big(\mathrm{Id} - J^2 F(u,z)\Big)|v| \Big\rangle + J^2 \sum_{j=1}^{L^d} (1 - \cos(\theta(u,z)_{jj}))|M(u,z)_{jj}|^2 |v_j|^2 \\
&\geqslant \Big\langle |v|, \Big(\mathrm{Id} - J^2 F(u,z)\Big)|v| \Big\rangle + J^2 \sum_{j=1}^{L^d} 2(\mathrm{Im}(M(u,z)_{jj}))^2 |v_j|^2.
\end{aligned}$$

The first term is nonnegative since $\|F(u,z)\| \leqslant \frac{1}{J^2}$, so we have

$$\mathrm{Re}\Big\langle v, (\mathrm{Id} - J^2 R(u,z))v \Big\rangle \geqslant 2J^2 \left(\min_{j=1}^{L^d} \mathrm{Im}(M(u,z)_{jj}) \right)^2 \|v\|^2.$$

But $\mathrm{Im}(M(u,z)_{jj})$ is the (j,j) entry of the matrix $\mathrm{Im}\, M(u,z)$, which is (strictly) positive definite by the definition of the MDE. □

Proposition 3.4.11. *$\mathcal{S}$ is maximized on the diagonal of $\mathbb{R}^d$, i.e.,*

$$\sup_{u \in \mathbb{R}^{L^d}} \mathcal{S}[u] = \sup_{u \in \mathbb{R}} \mathcal{S}[(u, \dots, u)].$$

Proof. It suffices to show that the set of maximizers

$$\mathcal{M} = \left\{ u \in \mathbb{R}^{L^d} : \mathcal{S}[u] = \sup_{v \in \mathbb{R}^{L^d}} \mathcal{S}[v] \right\}$$

intersects the diagonal. First, $\mathcal{M}$ is nonempty, since (by [35, Lemma 4.4]) $\mathcal{S}$ is continuous with $\lim_{\|u\| \to +\infty} \mathcal{S}[u] = -\infty$. Furthermore, $\mathcal{M}$ is closed under the operation "permute the coordinates (which are indexed by lattice points) with a permutation that is also a translation of the periodic lattice," since such permutations preserve $a(u)$ in (3.4.1) and thus $\mu_\infty(u)$. Finally, $\mathcal{M}$ is convex, since $\mathcal{S}[u]$ is concave.

Given $u \in \mathcal{M}$, its images under all possible lattice translations are thus all in $\mathcal{M}$, so the average of all these points (which is in their convex hull) is in $\mathcal{M}$. Since the lattice is periodic (i.e., translations are in bijection with lattice sites), this average is on the diagonal. □

Proof of Theorem 3.2.2. Using Proposition 3.4.11 to restrict the variational problem from Theorem 3.4.1 to the diagonal, we have

$$\Sigma(\mu_0) = -\frac{1}{L^d} \log(\det(\mu_0 \,\mathrm{Id}_{L^d \times L^d} - t_0 \Delta)) + \sup_{u \in \mathbb{R}} \left\{ \int_{\mathbb{R}} \log|\lambda| \mu_\infty((u, \dots, u), \lambda)\, \mathrm{d}\lambda - \frac{u^2}{2J^2} \right\}$$

and similarly for minima. One can check directly from the MDE that

$$\mu_\infty((u,\dots,u),\lambda) = \mu_\infty((0,\dots,0),\lambda - u),$$

and in fact we have $\mu_\infty((0,\dots,0)) = \rho_{\mathrm{sc},J^2} \boxplus \hat{\mu}_{-t_0\Delta+\mu_0 \mathrm{Id}}$. Indeed, by symmetry all the entries of $m_\infty(0,z)$ must be equal. If we denote by $m(z)$ their shared value (which is also the Stieltjes transform of $\mu_\infty((0,\dots,0))$), then by taking the normalized trace in (3.4.2) we find that $m(z)$ satisfies the self-consistent equation

$$m(z) = \int \frac{\hat{\mu}_{-t_0\Delta+\mu_0 \mathrm{Id}}(\mathrm{d}s)}{s - z - J^2 m(z)}.$$

This *Pastur relation* characterizes [131] the Stieltjes transform $m(z)$ of $\rho_{\mathrm{sc},J^2} \boxplus \hat{\mu}_{-t_0\Delta+\mu_0 \mathrm{Id}}$. Exchanging u and $-u$ gives (3.2.4). □

Proof of Theorem 3.2.4. Since $-L^{-d}\log\det(\mu_0 - t_0\Delta) = -\int \log(\lambda)\hat{\mu}_{-t_0\Delta+\mu_0 \mathrm{Id}}$, the variational problems given in (3.2.3) and (3.2.4) are exactly the variational problems analyzed for the soft spins model in (3.2.7) and (3.2.8), identifying μ_D there with $\hat{\mu}_{-t_0\Delta+\mu_0 \mathrm{Id}}$ here (which is gapped from zero since $\mu_0 > 0$) and $B''(0)$ there with J^2 here. The statement of Theorem 3.2.4 follows from our analysis of that variational problem in the next section, since

$$\int_{\mathbb{R}} \frac{\hat{\mu}_{-t_0\Delta+\mu_0 \mathrm{Id}}(\mathrm{d}\lambda)}{\lambda^2} = \int_{\mathbb{R}} \frac{\hat{\mu}_{-t_0\Delta}(\mathrm{d}\lambda)}{(\lambda+\mu_0)^2}$$

is a strictly decreasing function of μ_0, tending to 0 as $\mu_0 \to +\infty$ and tending to $+\infty$ (since the Laplacian is singular) as $\mu_0 \downarrow 0$. This proves existence and uniqueness of the Larkin mass as claimed. □

Remark 3.4.12. *Here we take Δ to be the lattice Laplacian, which is the classic choice in the elastic manifold, but as suggested in [89] the same methods and proofs allow us to replace Δ everywhere with any symmetric negative semidefinite $L^d \times L^d$ matrix. For example, this allows for interactions*

beyond pairwise.

3.5 Soft spins in an anisotropic well

3.5.1 Establishing the variational formula. In this subsection we prove Theorem 3.2.6, which establishes a variational formula for complexity. In the next subsection we analyze it to prove Theorem 3.2.8.

The Kac-Rice formula [2, Theorem 11.2.1] gives

$$\begin{aligned}\mathbb{E}[\mathrm{Crt}_N^{\mathrm{tot}}(\mathcal{H})] &= \int_{\mathbb{R}^N} \mathbb{E}[\left|\det(\nabla^2\mathcal{H}(\sigma))\right| \mid \nabla\mathcal{H}(\sigma) = 0]\phi_\sigma(0)\,\mathrm{d}\sigma,\\ \mathbb{E}[\mathrm{Crt}_N^{\min}(\mathcal{H})] &= \int_{\mathbb{R}^N} \mathbb{E}[\left|\det(\nabla^2\mathcal{H}(\sigma))\right|\mathbf{1}_{\nabla^2\mathcal{H}(\sigma)\geqslant 0} \mid \nabla\mathcal{H}(\sigma) = 0]\phi_\sigma(0)\,\mathrm{d}\sigma,\end{aligned} \tag{3.5.1}$$

where

$$\phi_\sigma(0) = \frac{1}{(2\pi B'(0))^{N/2}} \exp\left(-\frac{1}{2B'(0)}\|D_N\sigma\|^2\right)$$

is the density of $\nabla\mathcal{H}(\sigma)$ at $0 \in \mathbb{R}^N$. (As stated, the Kac-Rice formula actually counts the mean number of critical points, not in all of $\mathbb{R}^N$, but in a compact subset T of $\mathbb{R}^N$ satisfying some regularity assumptions; then the right-hand integrals in (3.5.1) are over T instead of $\mathbb{R}^N$. To obtain (3.5.1) as written, we use this version of Kac-Rice for some nested sequence $(T_N)_{N=1}^\infty$ of compact sets whose union is $\mathbb{R}^N$ and apply monotone convergence on both sides.)

Since V is isotropic, for each σ we have that $(\nabla^2 V(\sigma), V(\sigma))$ is independent of $\nabla V(\sigma)$; hence for each σ we also have that $(\nabla^2\mathcal{H}(\sigma), \mathcal{H}(\sigma))$ is independent of $\nabla\mathcal{H}(\sigma)$. In fact, since V is isotropic the distribution of $\nabla^2 V(\sigma)$ is independent of σ; and by computation

$$\nabla^2 \frac{1}{2}\langle\sigma, D_N\sigma\rangle = D_N$$

is independent of σ as well. Thus

$$\begin{aligned}\mathbb{E}[\mathrm{Crt}_N^{\mathrm{tot}}(\mathcal{H})] &= \int_{\mathbb{R}^N} \mathbb{E}[\left|\det(\nabla^2\mathcal{H}(\sigma))\right| \mid \nabla\mathcal{H}(\sigma)=0]\phi_\sigma(0)\,\mathrm{d}\sigma = \mathbb{E}\Big[\left|\det(\nabla^2\mathcal{H}(\mathbf{0}))\right|\Big]\int_{\mathbb{R}^N}\phi_\sigma(0)\,\mathrm{d}\sigma \\ &= \frac{1}{\det(D_N)}\mathbb{E}\Big[\left|\det(\nabla^2\mathcal{H}(\mathbf{0}))\right|\Big], \\ \mathbb{E}[\mathrm{Crt}_N^{\mathrm{min}}(\mathcal{H})] &= \frac{1}{\det(D_N)}\mathbb{E}\Big[\left|\det(\nabla^2\mathcal{H}(\mathbf{0}))\right|\mathbf{1}_{\nabla^2\mathcal{H}(\mathbf{0})\geqslant 0}\Big].\end{aligned} \tag{3.5.2}$$

Since the eigenvalues of D_N are gapped away from zero and from infinity, uniformly in N, we have

$$\lim_{N\to\infty}\frac{1}{N}\log\left(\frac{1}{\det(D_N)}\right) = -\int\log(\lambda)\mu_D(\mathrm{d}\lambda).$$

Thus it remains only to study the Hessian.

Classical Gaussian computations (e.g., [2, Section 5.5]) yield

$$\nabla^2\mathcal{H}(\mathbf{0}) \overset{(d)}{=} W_N + \xi\,\mathrm{Id} + D_N,$$

where W_N is distributed according to $\sqrt{B''(0)}$ times the GOE and $\xi\sim\mathcal{N}(0,B''(0)/N)$ is independent of W_N. In fact, since the law of $W_N+\xi\,\mathrm{Id}$ is invariant under conjugation by orthogonal matrices, we can assume without loss of generality that D_N is diagonal. If we define

$$A_N(u) = u\,\mathrm{Id} + D_N$$

and $H_N(u) = W_N + A_N(u)$, then we have

$$\begin{aligned}\mathbb{E}\Big[\left|\det(\nabla^2\mathcal{H}(\mathbf{0}))\right|\Big] &= \frac{1}{\sqrt{2\pi/N}}\int_{\mathbb{R}} e^{-N\frac{u^2}{2B''(0)}}\mathbb{E}[|\det(H_N(u))|]\,\mathrm{d}u, \\ \mathbb{E}\Big[\left|\det(\nabla^2\mathcal{H}(\mathbf{0}))\right|\mathbf{1}_{\nabla^2\mathcal{H}(\mathbf{0})\geqslant 0}\Big] &= \frac{1}{\sqrt{2\pi/N}}\int_{\mathbb{R}} e^{-N\frac{u^2}{2B''(0)}}\mathbb{E}[|\det(H_N(u))|\mathbf{1}_{H_N(u)\geqslant 0}]\,\mathrm{d}u.\end{aligned} \tag{3.5.3}$$

Now we study the relevant MDE. Given a linear operator $\mathcal{S}:\mathbb{C}^{N\times N}\to\mathbb{C}^{N\times N}$ that is self-adjoint

with respect to the inner product $\langle R, T\rangle = \operatorname{Tr}(R^* T)$ and that preserves the cone of positive-semi-definite matrices, the problem

$$-M^{-1}(u,z) = z\,\mathrm{Id} - A_N(u) + \mathcal{S}[M(u,z)] \quad \text{subject to} \quad \operatorname{Im} M(u,z) > 0 \tag{3.5.4}$$

has a unique solution $M(u,z) \in \mathbb{C}^{N\times N}$ for each $z \in \mathbb{H}$ and $u \in \mathbb{R}$. We will consider this problem with two choices of operator $\mathcal{S}$:

$$\begin{aligned} \mathcal{S}_N[T] &= \frac{B''(0)}{N}\operatorname{Tr}(T) + B''(0)\frac{T^{\mathrm{tr}}}{N} \quad \text{induces} \quad M_N(u,z), \\ \mathcal{S}'_N[T] &= \frac{B''(0)}{N}\operatorname{Tr}(T) \quad \text{induces} \quad M'_N(u,z). \end{aligned}$$

Let $\mu_N(u)$ and $\mu'_N(u)$ be the probability measures whose Stieltjes transforms are, respectively, $\frac{1}{N}\operatorname{Tr}(M_N(u,z))$ and $\frac{1}{N}\operatorname{Tr}(M'_N(u,z))$. Recall the notation $\rho_{\mathrm{sc},t}$ for the semicircle law of variance t.

Lemma 3.5.1. *We recognize*

$$\mu'_N(u) = \rho_{sc,B''(0)} \boxplus \hat{\mu}_{A_N(u)}.$$

Proof. Write $m'_N(u,z)$ for the Stieltjes transform of $\rho_{\mathrm{sc},B''(0)} \boxplus \hat{\mu}_{A_N(u)}$. The Pastur relation [131], which characterizes the Stieltjes transform of the free convolution of the semicircle law with another measure, states that $m'_N(u,z)$ satisfies the self-consistent equation

$$m'_N(u,z) = \int \frac{(\hat{\mu}_{D_N + u\,\mathrm{Id}})(\mathrm{d}\lambda)}{\lambda - z - B''(0)m'_N(u,z)} = \frac{1}{N}\sum_{i=1}^{N} \frac{1}{(D_N)_{ii} + u - z - B''(0)m'_N(u,z)}.$$

(Recall we changed variables so that D_N is diagonal.) If we define

$$\mathcal{M}'_N(u,z) = \operatorname{diag}\left(\frac{1}{(D_N)_{11} + u - z - B''(0)m'_N(u,z)}, \ldots, \frac{1}{(D_N)_{NN} + u - z - B''(0)m'_N(u,z)}\right),$$

this Pastur relation then gives $m'_N(u,z) = \frac{1}{N}\operatorname{Tr}(\mathcal{M}'_N(u,z))$, which means that $\mathcal{M}'_N(u,z)$ exhibits a solution to the MDE (3.5.4) with $\mathcal{S} = \mathcal{S}'_N$. (Since $\operatorname{Im} m'_N(u,z) > 0$ when $z \in \mathbb{H}$, one can check that $\operatorname{Im} \mathcal{M}'_N(u,z) > 0$.) Thus $m'_N(u,z)$, which we defined as the Stieltjes transform of $\rho_{\mathrm{sc},B''(0)} \boxplus \hat{\mu}_{A_N(u)}$,

is also the Stieltjes transform of $\mu'_N(u)$. □

Let τ_u be the translation $\tau_u(x) = x + u$, and write $(\tau_u)_*\mu$ for the pushforward of a probability measure μ under τ_u (i.e., the translation of μ by u).

Lemma 3.5.2. *The measures $\mu'_N(u)$ and*

$$\mu_\infty(u) = \rho_{sc,B''(0)} \boxplus ((\tau_u)_*\mu_D)$$

admit bounded and compactly supported densities on $\mathbb{R}$, locally uniformly in u (meaning the bound and the compact set can be taken uniform on compact sets of u).

Proof. These are standard consequences of the regularity of free convolution with the semicircle law, studied in depth by [49]. For a compactly supported measure μ and $t > 0$, we have [49, Corollaries 2, 5] that $\rho_{\mathrm{sc},t} \boxplus \mu$ admits a density $(\rho_{\mathrm{sc},t} \boxplus \mu)(\cdot)$ with

$$(\rho_{\mathrm{sc},t} \boxplus \mu)(x) \leqslant \left(\frac{3}{4\pi^3 t^2}(4 + |\mathtt{r}(\mu) - \mathtt{l}(\mu)|)\right)^{1/3} \mathbf{1}_{\mathtt{l}(\mu)-2\sqrt{t} \leqslant x \leqslant \mathtt{r}(\mu)+2\sqrt{t}}.$$

To study $\mu'_N(u)$, we apply this with $\mu = \hat{\mu}_{A_N(u)}$. Since $\mathtt{r}(\hat{\mu}_{A_N(u)}) \leqslant u + \sup_N \lambda_{\max}(D_N)$ and $\mathtt{l}(\hat{\mu}_{A_N(u)}) \geqslant u$, both of which are uniformly bounded over $u \in B_R$, we obtain the claim for $\mu'_N(u)$. The proof for $\mu_\infty(u)$ is similar. □

Proposition 3.5.3. *We have*

$$\begin{aligned} &\lim_{N\to\infty} \frac{1}{N} \log \int_{\mathbb{R}} e^{-\frac{Nu^2}{2B''(0)}} \mathbb{E}[|\det(H_N(u))|] \, \mathrm{d}u \\ &= \sup_{u\in\mathbb{R}} \left\{ \int_{\mathbb{R}} \log|\lambda - u| \mathrm{d}(\rho_{sc,B''(0)} \boxplus \mu_D)(\lambda) - \frac{u^2}{2B''(0)} \right\}, \end{aligned} \tag{3.5.5}$$

$$\begin{aligned} &\lim_{N\to\infty} \frac{1}{N} \log \int_{\mathbb{R}} e^{-\frac{Nu^2}{2B''(0)}} \mathbb{E}[|\det(H_N(u))| \mathbf{1}_{H_N(u)\geqslant 0}] \, \mathrm{d}u \\ &= \sup_{u \leqslant \mathtt{l}(\rho_{sc,B''(0)} \boxplus \mu_D)} \left\{ \int_{\mathbb{R}} \log|\lambda - u| \mathrm{d}(\rho_{sc,B''(0)} \boxplus \mu_D)(\lambda) - \frac{u^2}{2B''(0)} \right\}. \end{aligned} \tag{3.5.6}$$

Proof. For (3.5.5), we wish to apply [35, Theorem 4.1] with $\alpha = \frac{1}{2B''(0)}$, $p = 0$, $\mathfrak{D} = \mathbb{R}$, and $\mu_\infty(u)$ as above. To do this, we will consider $H_N(u)$ as a sequence of "Gaussian matrices with a (co)variance profile," in the language of [35, Corollary 1.8.A]. So we verify the assumptions of that corollary.

By assumption we have $\sup_N \lambda_{\max}(D_N) < \infty$; thus

$$\sup_N \|A_N(u)\| \leqslant |u| + \sup_N \lambda_{\max}(D_N) < \infty.$$

Since W_N is $\sqrt{B''(0)}$ times a GOE matrix, the proof of Lemma 3.4.4 gives us $\mathbb{E}[|\det(H_N(u))|] \leqslant (C\max(\|u\|, 1))^N$ for some C. For the same reason, we can compute directly

$$\mathbb{E}[W_N T W_N] = \frac{B''(0)}{N}\operatorname{Tr}(T)\operatorname{Id} + \frac{B''(0)}{N}T$$

which verifies the flatness condition. Since everything is locally uniform in u, it remains only to show

$$\mathrm{W}_1(\mu_N(u), \mu_\infty(u)) \leqslant N^{-\kappa} \tag{3.5.7}$$

for some $\kappa > 0$. Since all of these measures are compactly supported, locally uniformly in u, the Wasserstein-1 and bounded-Lipschitz distances are equivalent, so we will work with d_{BL}.

First we relate μ_N to μ'_N, using Lemma 3.3.1 (with $P_N = N$) to estimate the difference between their Stieltjes transforms and then following the proof of [35, Proposition 3.1], using the regularity we established in Lemma 3.5.2. To relate μ'_N and μ_∞, we write d_L for the Lévy distance between probability measures, then combine the translation-invariance of bounded-Lipschitz distance, [71, Corollary 11.6.5, Theorem 11.3.3], and [48, Proposition 4.13] to obtain

$$\begin{aligned}
d_{\mathrm{BL}}(\mu'_N(u), \mu_\infty(u)) &= d_{\mathrm{BL}}(\rho_{\mathrm{sc},B''(0)} \boxplus \hat{\mu}_{A_N(u)}, \rho_{\mathrm{sc},B''(0)} \boxplus ((\tau_u)_*\mu_D)) \\
&= d_{\mathrm{BL}}((\tau_u)_*(\rho_{\mathrm{sc},B''(0)} \boxplus \hat{\mu}_{D_N}), (\tau_u)_*(\rho_{\mathrm{sc},B''(0)} \boxplus \mu_D)) \\
&= d_{\mathrm{BL}}(\rho_{\mathrm{sc},B''(0)} \boxplus \hat{\mu}_{D_N}, \rho_{\mathrm{sc},B''(0)} \boxplus \mu_D) \leqslant 2d_L(\rho_{\mathrm{sc},B''(0)} \boxplus \hat{\mu}_{D_N}, \rho_{\mathrm{sc},B''(0)} \boxplus \mu_D) \\
&\leqslant 2d_L(\hat{\mu}_{D_N}, \mu_D) \leqslant 4\sqrt{d_{\mathrm{BL}}(\hat{\mu}_{D_N}, \mu_D)} \leqslant N^{-\varepsilon},
\end{aligned}$$

uniformly over $u \in \mathbb{R}$, where the last inequality is by assumption (3.2.6). This verifies (3.5.7), and thus [35, Theorem 4.1] yields

$$\lim_{N\to\infty} \frac{1}{N} \log \int_{\mathbb{R}} e^{-N\frac{u^2}{2B''(0)}} \mathbb{E}[|\det(H_N(u))|]\,\mathrm{d}u = \sup_{u\in\mathbb{R}} \left\{ \int_{\mathbb{R}} \log|\lambda| \mu_\infty(u,\lambda)\,\mathrm{d}\lambda - \frac{u^2}{2B''(0)} \right\}.$$

To obtain (3.5.5), we notice that

$$\mu_\infty(u,\lambda) = (\rho_{\mathrm{sc},B''(0)} \boxplus ((\tau_u)_* \mu_D))(\lambda) = ((\tau_u)_* (\rho_{\mathrm{sc},B''(0)} \boxplus \mu_D))(\lambda) = (\rho_{\mathrm{sc},B''(0)} \boxplus \mu_D)(\lambda - u) \quad (3.5.8)$$

and change variables twice (exchanging u and $-u$). This completes the proof of (3.5.5).

For (3.5.6), we wish to apply [35, Theorem 4.5] with $\alpha = \frac{1}{2B''(0)}$, $p = 0$, $\mathfrak{D} = \mathbb{R}$, and $\mu_\infty(u)$ as above. Now we verify its conditions. Arguments as in the elastic-manifold case, specifically Lemma 3.4.5, give [35, (4.6)]. By (3.5.8), $\mathbb{P}(\mathrm{Spec}(H_N(u)) \subset [\mathtt{l}(\mu_\infty(u)) - \varepsilon, \mathtt{r}(\mu_\infty(u)) + \varepsilon])$ is actually independent of u, and when $u = 0$ it takes the form

$$\mathbb{P}(\mathrm{Spec}(W_N + D_N) \subset [\mathtt{l}(\rho_{\mathrm{sc},B''(0)} \boxplus \mu_D) - \varepsilon, \mathtt{r}(\rho_{\mathrm{sc},B''(0)} \boxplus \mu_D) + \varepsilon]).$$

Estimates showing that this tends to one are classical, since W_N is $\sqrt{B''(0)}$ times a GOE matrix and D_N has no outliers by assumption (recall that we made this assumption only for counting local minima, not for counting total critical points). In the generality we need (i.e., with the fewest assumptions on D_N), this estimate follows from the large-deviations result [121, (2.5)] (written for GOE, not $\sqrt{B''(0)}$ times GOE, but clearly goes through in this generality); this verifies [35, (4.7)]. Finally, the topological requirement [35, (4.8)] follows immediately after noticing that (in the notation there)

$$\mathcal{G} = \{u : \mu_\infty(u)((-\infty, 0)) = 0\} = \{u : u \geqslant -\mathtt{l}(\rho_{\mathrm{sc},B''(0)} \boxplus \mu_D)\},$$
$$\mathcal{G}_{+\varepsilon} = \{u : \mathtt{l}(\mu_\infty(u)) \geqslant 2\varepsilon\} = \{u : u \geqslant 2\varepsilon - \mathtt{l}(\rho_{\mathrm{sc},B''(0)} \boxplus \mu_D)\}.$$

Having checked all the conditions, we can apply [35, Theorem 4.5] to complete the proof. □

Proof of Theorem 3.2.6. This follows immediately from (3.5.2), (3.5.3), and Proposition 3.5.3. □

3.5.2 Analyzing the variational formula. The key idea presented here is a dynamical analysis of the variational formulas appearing in the previous section, increasing the noise parameter $B''(0)$. Important ingredients are the Burgers' equation (3.5.10) and the square root edge behavior of the relevant free convolutions, as proved in Appendix B.

Proof of Theorem 3.2.8. In this proof, we state several claims as lemmas, postponing their proofs.

We think of the variational problem as dynamic in the parameter t, which corresponds to the noise parameter $B''(0)$ in the complexity problem, for fixed μ_D. That is, at "time 0" we have a pure signal with zero complexity, and as "time" (meaning noise) increases we find a threshold at which complexity becomes positive. Precisely, write

$$\begin{aligned}\mu_t &= \rho_{\mathrm{sc},t} \boxplus \mu_D,\\ \ell_t &= \mathtt{l}(\mu_t),\\ r_t &= \mathtt{r}(\mu_t),\end{aligned}$$

for the free convolution of μ_D with the semi-circular distribution of variance t (which has density $\mu_t(\cdot)$) and its left and right edges, respectively. Let

$$F(u,t) = -\int_{\mathbb{R}} \log(\lambda)\mu_D(\mathrm{d}\lambda) + \int_{\mathbb{R}} \log|\lambda - u|\mu_t(\lambda)\,\mathrm{d}\lambda - \frac{u^2}{2t}$$

and recall that we are interested in

$$\begin{aligned}\Sigma^{\mathrm{tot}}(\mu_D,t) &= \sup_{u\in\mathbb{R}} F(u,t),\\ \Sigma^{\mathrm{min}}(\mu_D,t) &= \sup_{u\leqslant\ell_t} F(u,t).\end{aligned}$$

Let

$$u_t = -t \int \frac{\mu_D(\mathrm{d}\lambda)}{\lambda}, \tag{3.5.9}$$

and consider the thresholds

$$t_0 = \inf\{t > 0 : u_t = \ell_t\},$$
$$t_c = \left(\int \frac{\mu_D(\mathrm{d}\lambda)}{\lambda^2} \right)^{-1}.$$

Later we will show that $t_0 = t_c$, but for now we distinguish between them. In particular we do not yet assume that t_0 is finite. Since μ_D is supported in $(0, \infty)$, we have $u_0 = 0 < \ell_0$, and by continuity we have $u_t < \ell_t$ for all $t < t_0$.

Let m_t be the Stieltjes transform of μ_t, with the sign convention $m_t(z) = \int_{\mathbb{R}} \frac{\mu_t(\mathrm{d}\lambda)}{\lambda - z}$. It is known (see for example [152, 49], noting their opposite sign convention $m_t(z) = \int_{\mathbb{R}} \frac{\cdots}{z-\lambda}$) that for any z outside the support of μ_t, we have

$$\partial_t m_t(z) - m_t(z) \partial_z m_t(z) = 0. \tag{3.5.10}$$

For $t < t_0$, u_t is not in the support of μ_t, so (3.5.10) gives

$$\frac{\mathrm{d}}{\mathrm{d}t} m_t(u_t) = \partial_u m_t(u) \partial_t u_t + \partial_t m_t(u) = \partial_u m_t(u)(\partial_t u_t + m_t(u_t)) = \partial_u m_t(u)(-m_0(u_0) + m_t(u_t)).$$

The (unique) solution to this differential equation is clearly $m_t(u_t) = m_0(u_0)$, so that

$$-m_t(u_t) - \frac{u_t}{t} = 0, \tag{3.5.11}$$

i.e.

$$\left(\frac{\partial}{\partial u} F(u, t) \right)_{u = u_t} = 0 \tag{3.5.12}$$

for $t < t_0$.

Lemma 3.5.4. *For any $u \in \mathbb{R}$, we have*

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}} \log|\lambda - u|\mu_t(\mathrm{d}\lambda) = \frac{\mathrm{Im}(m_t(u))^2 - \mathrm{Re}(m_t(u))^2}{2}.$$

For $t < t_0$ (when $\mathrm{Im}(m_t(u_t)) = 0$), we can then use (3.5.12), Lemma 3.5.4, and (3.5.11) to obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}F(u_t,t) = \left(\frac{\partial}{\partial t}F(u,t)\right)_{u=u_t} + \left(\frac{\partial}{\partial u}F(u,t)\right)_{u=u_t}\partial_t u_t = \left(\frac{\partial}{\partial t}F(u,t)\right)_{u=u_t} = -\frac{m_t(u_t)^2}{2} + \frac{(u_t)^2}{2t^2} = 0.$$

Together with $F(u_t,t) \to 0$ as $t \to 0$, the above equation gives

$$F(u_t,t) = 0 \tag{3.5.13}$$

for $t < t_0$.

Lemma 3.5.5. *For every measure μ_D and every t, the function $F(u,t)$ is concave in u (possibly not strictly).*

From (3.5.12), (3.5.13), and Lemma 3.5.5 we conclude that

$$\Sigma^{\mathrm{tot}}(\mu_D,t) = \Sigma^{\mathrm{min}}(\mu_D,t) = F(u_t,t) = 0 \quad \text{for all} \quad t < t_0. \tag{3.5.14}$$

Now we study the phase $t > t_0$, showing $t_0 < \infty$ along the way, by considering the evolution of ℓ_t.

Lemma 3.5.6. *For all $t > 0$ we have*

$$\partial_t \ell_t = -\,\mathrm{Re}(m_t(\ell_t)).$$

Since the density of μ_t decays to zero at its edges (in fact at least as quickly as a cube root

[49]), we have $\mathrm{Im}(m_t(\ell_t)) = 0$ for all t. From Lemmas 3.5.6 and 3.5.4 we therefore obtain

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}F(\ell_t,t) &= \Big(\frac{\partial}{\partial u}F(u,t)\Big)_{u=\ell_t}\partial_t\ell_t + \Big(\frac{\partial}{\partial t}F(u,t)\Big)_{u=\ell_t} \\ &= \Big(-\mathrm{Re}(m_t(\ell_t)) - \frac{\ell_t}{t}\Big)(-\mathrm{Re}(m_t(\ell_t))) + \frac{(\mathrm{Im}(m_t(\ell_t)))^2}{2} + \frac{1}{2}\bigg[\Big(\frac{\ell_t}{t}\Big)^2 - (\mathrm{Re}(m_t(\ell_t)))^2\bigg] \\ &= \frac{1}{2}\Big(\frac{\ell_t}{t} + \mathrm{Re}(m_t(\ell_t))\Big)^2 = \frac{1}{2}\bigg[\Big(\frac{\partial}{\partial u}F(u,t)\Big)_{u=\ell_t}\bigg]^2.\end{aligned} \tag{3.5.15}$$

To analyze this, we use the following lemma.

Lemma 3.5.7. *We have*

$$\Big(\frac{\partial}{\partial u}F(u,t)\Big)_{u=\ell_t}\begin{cases} < 0 & \text{if} \quad 0 < t < t_c, \\ = 0 & \text{if} \quad t = t_c, \\ > 0 & \text{if} \quad t > t_c. \end{cases}$$

Thus (3.5.15) is positive for $t \neq t_c$ and vanishes at $t = t_c$. This has two important consequences. First, $F(\ell_t, t)$ is a *strictly* increasing function of t. Second,

$$t_0 = t_c. \tag{3.5.16}$$

Indeed, on the one hand, for $t < t_0$ and small $\varepsilon > 0$, we have

$$F(\ell_t,t) < F(\ell_{t+\varepsilon}, t+\varepsilon) \leqslant F(u_{t+\varepsilon}, t+\varepsilon) = 0 = F(u_t,t),$$

so that $\Big(\frac{\partial}{\partial u}F(u,t)\Big)_{u=\ell_t} \neq 0$ by concavity in u (Lemma 3.5.5) and (3.5.12), and thus $t \neq t_c$. Hence $t_0 \leqslant t_c < \infty$.

On the other hand, if t has the property that $\sup_{u\in\mathbb{R}} F(u,t) = F(\ell_t, t)$, then we have

$$\left(\frac{\partial}{\partial u} F(u,t)\right)_{u=\ell_t} = 0,$$

thus $t = t_c$. But t_0 has this property, now that we know it is finite, since by continuity we have

$$0 = \sup_{u\in\mathbb{R}} F(u,t_0) = F(u_{t_0}, t_0) = F(\ell_{t_0}, t_0).$$

We have shown that $F(\ell_t, t)$ is a strictly increasing function which vanishes at t_c; thus

$$\Sigma^{\mathrm{tot}}(\mu_D, t) \geqslant \Sigma^{\min}(\mu_D, t) \geqslant F(\ell_t, t) > F(\ell_{t_c}, t_c) = 0 \quad \text{for all} \quad t > t_c. \tag{3.5.17}$$

The fact that both complexities vanish if and only if $t \leqslant t_c$ follows immediately from (3.5.14), (3.5.16), and (3.5.17) (the case $t = t_0$ follows from (3.5.14) by continuity).

Lemmas 3.5.5 and 3.5.7 combine to give (3.2.11), as well as strict inequality in $\Sigma^{\mathrm{tot}}(\mu_D, t) > \Sigma^{\min}(\mu_D, t)$ for $t > t_c$. Now we prove (3.2.12). To do this, we will rely on Pastur's relation [131]

$$m_t(z) = \int \frac{\mu_D(\mathrm{d}\lambda)}{\lambda - z - t m_t(z)}. \tag{3.5.18}$$

By taking real and imaginary parts of (3.5.18), we get for any $u \in \mathbb{R}$ the coupled system

$$\mathrm{Re}(m_t(u)) = \int \frac{\lambda - u - t\,\mathrm{Re}(m_t(u))}{(\lambda - u - t\,\mathrm{Re}(m_t(u)))^2 + t^2\,\mathrm{Im}(m_t(u))^2}\,\mu_D(\mathrm{d}\lambda), \tag{3.5.19}$$

$$\mathrm{Im}(m_t(u)) = t\int \frac{\mathrm{Im}(m_t(u))}{(\lambda - u - t\,\mathrm{Re}(m_t(u)))^2 + t^2\,\mathrm{Im}(m_t(u))^2}\,\mu_D(\mathrm{d}\lambda). \tag{3.5.20}$$

If v_t satisfies $F(v_t, t) = \sup_{u\in\mathbb{R}} F(u,t)$, then

$$0 = \left(\frac{\partial}{\partial u} F(u,t)\right)_{u=v_t} = -\frac{v_t}{t} - \mathrm{Re}(m_t(v_t)).$$

We plug this into (3.5.19) and (3.5.20) to obtain, writing $y_t = \mathrm{Im}(m_t(v_t))$,

$$-\frac{v_t}{t} = \int \frac{\lambda}{\lambda^2 + t^2 y_t^2}\, \mu_D(\mathrm{d}\lambda), \tag{3.5.21}$$

$$y_t = t \int \frac{y_t}{\lambda^2 + t^2 y_t^2}\, \mu_D(\mathrm{d}\lambda). \tag{3.5.22}$$

From its definition, $y_t \geqslant 0$. For every $t > 0$, notice that $(u_t, 0)$ is a solution to the coupled system {(3.5.21), (3.5.22)}, where u_t was defined in (3.5.9). We claim that this is the unique solution when $t \leqslant t_c$, but that for $t > t_c$ there is exactly one more solution, with $y_t > 0$, and that for such times this latter solution is the one corresponding to the optimizer (i.e., for $t > t_c$ the point u_t is not an optimizer anymore).

For existence and uniqueness of this second solution exactly when $t > t_c$, we note that the positive solutions y_t to (3.5.22) are exactly the positive solutions to

$$\frac{1}{t} = \int_{\mathbb{R}} \frac{\mu_D(\mathrm{d}\lambda)}{\lambda^2 + t^2 y_t^2}, \tag{3.5.23}$$

but the right-hand side of this equation is a strictly decreasing function of y_t, tending to zero as $y_t \to +\infty$ and tending to $\frac{1}{t_c}$ as $y_t \downarrow 0$ (which is bigger than $\frac{1}{t}$ precisely when $t > t_c$).

To verify the claim that u_t is not an optimizer when $t > t_c$, it suffices to show

$$u_t \leqslant \ell_t \quad \text{for all} \quad t. \tag{3.5.24}$$

Indeed, since $F(u,t)$ is concave and $\left(\frac{\partial}{\partial u} F(u,t)\right)_{u=\ell_t} > 0$ in the regime $t > t_c$, (3.5.24) would imply that u_t is not the optimizer of $F(\cdot, t)$ when $t > t_c$.

To show (3.5.24), we will show that $t \mapsto \ell_t - u_t$ is convex with a vanishing derivative at $t = t_c$, where it takes the value zero. First we claim

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}(\ell_t - u_t) = \frac{\mathrm{d}^2}{\mathrm{d}t^2}\ell_t = \frac{\mathrm{d}}{\mathrm{d}t}(-m_t(\ell_t)) \geqslant 0 \tag{3.5.25}$$

for all t. Indeed, a simple calculation similar to the previous ones gives, for any $\varepsilon > 0$,

$$\frac{\mathrm{d}}{\mathrm{d}t} m_t(\ell_t - \varepsilon) = (m_t(\ell_t - \varepsilon) - m_t(\ell_t))\partial_u m_t(\ell_t - \varepsilon) = \int \frac{-\sqrt{\varepsilon}\mu_t(\mathrm{d}\lambda)}{(\lambda - (\ell_t - \varepsilon))(\lambda - \ell_t)} \int \frac{\sqrt{\varepsilon}\mu_t(\mathrm{d}\lambda)}{(\lambda - (\ell_t - \varepsilon))^2} < 0. \tag{3.5.26}$$

As $\varepsilon \downarrow 0$, each of the integrals on the right-hand side converges, since μ_t decays at its left edge at least as quickly as a square root by Proposition B.1, and since, for example,

$$\lim_{\varepsilon \downarrow 0} \int_0^1 \frac{\sqrt{\varepsilon} x^p}{(x + \varepsilon)x} \, \mathrm{d}x = \begin{cases} \pi & \text{if } p = 1/2, \\ 0 & \text{if } p > 1/2. \end{cases}$$

(When $p = 1/2$, this can be integrated directly at each ε; when $p > 1/2$, we use dominated convergence with dominating function $\frac{1}{2}x^{p-\frac{3}{2}}$.) Thus in the limit $\varepsilon \downarrow 0$ we prove the existence of $\frac{\mathrm{d}}{\mathrm{d}t} m_t(\ell_t) \leqslant 0$, concluding the proof of (3.5.25). Since

$$\left[\frac{\mathrm{d}}{\mathrm{d}t}(\ell_t - u_t)\right]_{t=t_c} = -\operatorname{Re}(m_{t_c}(\ell_{t_c})) - \frac{u_{t_c}}{t_c} = -\operatorname{Re}(m_{t_c}(\ell_{t_c})) - \frac{\ell_{t_c}}{t_c} = \left(\frac{\partial}{\partial u} F(u, t_c)\right)_{u=\ell_{t_c}} = 0$$

with $\ell_{t_c} = u_{t_c}$, we conclude (3.5.24) and thus (3.2.12).

Next we study the degree of vanishing of $\Sigma^{\mathrm{tot}}(\mu_D, t) = F(v_t, t)$ as $t \downarrow t_c$. First, note that v_t and y_t are C^1 functions of $t > t_c$ with the appropriate right-hand limits at criticality (namely $\lim_{t \downarrow t_c} y_t = 0$ and $\lim_{t \downarrow t_c} v_t = \ell_{t_c}$): this is proved, first by studying y_t via (3.5.23) and the implicit function theorem, then studying v_t via (3.5.21) using the knowledge of y_t. For $t > t_c$, Lemma 3.5.4 gives

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} F(v_t, t) &= \underbrace{\left(\frac{\partial}{\partial u} F(u, t)\right)_{u=v_t}}_{=0} \partial_t v_t + \left(\frac{\partial}{\partial t} F(u, t)\right)_{u=v_t} \\ &= \frac{\operatorname{Im}(m_t(v_t))^2 - \operatorname{Re}(m_t(v_t))^2}{2} + \frac{v_t^2}{2t^2} = \frac{\operatorname{Im}(m_t(v_t))^2}{2} = \frac{y_t^2}{2}. \end{aligned}$$

As $t \downarrow t_c$, this tends to zero. Differentiating (3.5.23) in t to find an expression for $y_t y_t'$ and inserting

it, we find

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}F(v_t,t) = y_t y_t' = \frac{1}{2t^4\int_{\mathbb{R}}\frac{\mu_D(\mathrm{d}\lambda)}{(\lambda^2+t^2y_t^2)^2}} - \frac{y_t^2}{t}.$$

As $t \downarrow t_c$, this tends to $\frac{1}{2}t_c^{-4}\Big(\int_{\mathbb{R}}\frac{\mu_D(\mathrm{d}\lambda)}{\lambda^4}\Big)^{-1} = \frac{1}{2}\Big(\int_{\mathbb{R}}\frac{\mu_D(\mathrm{d}\lambda)}{\lambda^2}\Big)^4\Big(\int_{\mathbb{R}}\frac{\mu_D(\mathrm{d}\lambda)}{\lambda^4}\Big)^{-1}$, which is positive. This gives us the quadratic decay and the prefactor.

Finally we study the degree of vanishing of $\Sigma^{\min}(\mu_D,t) = F(\ell_t,t)$ as $t \downarrow t_c$. To do this, we first study regularity of $m_t(\ell_t)$ (we studied regularity of ℓ_t above, around (3.5.25)). Notice that $\mathrm{Im}(m_t(\ell_t)) = 0$ but $\mathrm{Im}(m_t(\ell_t+\varepsilon)) > 0$ for all sufficiently small $\varepsilon > 0$, since μ_t admits a density that vanishes at the endpoints and is analytic where positive [49]; using this in the real and imaginary parts (3.5.19) and (3.5.20) of the Pastur relation, we obtain

$$m_t(\ell_t) = \int \frac{1}{\lambda - \ell_t - tm_t(\ell_t)}\,\mu_D(\mathrm{d}\lambda), \tag{3.5.27}$$

$$1 = t\int \frac{1}{(\lambda - \ell_t - tm_t(\ell_t))^2}\,\mu_D(\mathrm{d}\lambda). \tag{3.5.28}$$

For $t > t_c$, we will show in the proof of Lemma 3.5.7 that $\ell_t + tm_t(\ell_t) < 0$; thus $\int \frac{\mu_D(\mathrm{d}\lambda)}{(\lambda-\ell_t-tm_t(\ell_t))^p} < \infty$ for all $p > 0$. This also implies, using the implicit function theorem, that $\ell_t + tm_t(\ell_t)$ is a C^2 function of $t > t_c$, hence so is $m_t(\ell_t)$. Differentiating (3.5.28) in t and solving for $\frac{\mathrm{d}}{\mathrm{d}t}m_t(\ell_t)$, we find

$$\frac{\mathrm{d}}{\mathrm{d}t}m_t(\ell_t) = -\frac{1}{2t^3\Big(\int_{\mathbb{R}}\frac{\mu_D(\mathrm{d}\lambda)}{(\lambda-\ell_t-tm_t(\ell_t))^3}\Big)}.$$

As $t \downarrow t_c$, this tends to $-\frac{1}{2}\Big(\int_{\mathbb{R}}\frac{\mu_D(\mathrm{d}\lambda)}{\lambda^2}\Big)^3\Big(\int_{\mathbb{R}}\frac{\mu_D(\mathrm{d}\lambda)}{\lambda^3}\Big)^{-1}$. Now we compute derivatives: We have $F(\ell_{t_c},t_c) = 0$, by combining (3.5.15) and Lemma 3.5.7 we find that the first derivative also vanishes at criticality. Next, from (3.5.15) and Lemma 3.5.6 we have

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}F(\ell_t,t) = \Big(\frac{\ell_t}{t} + m_t(\ell_t)\Big)\Big(-\frac{m_t(\ell_t)}{t} - \frac{\ell_t}{t^2} + \frac{\mathrm{d}}{\mathrm{d}t}m_t(\ell_t)\Big).$$

At $t = t_c$, this vanishes by Lemma 3.5.7. The third derivative is

$$\frac{\mathrm{d}^3}{\mathrm{d}t^3}F(\ell_t, t) = \left(\frac{\ell_t}{t} + m_t(\ell_t)\right)\left(-\frac{1}{t}\cdot\frac{\mathrm{d}}{\mathrm{d}t}m_t(\ell_t) + 2\frac{m_t(\ell_t)}{t^2} + 2\frac{\ell_t}{t^3} + \frac{\mathrm{d}^2}{\mathrm{d}t^2}m_t(\ell_t)\right) + \left(-\frac{m_t(\ell_t)}{t} - \frac{\ell_t}{t^2} + \frac{\mathrm{d}}{\mathrm{d}t}m_t(\ell_t)\right)^2.$$

Since $\frac{\ell_{t_c}}{t_c} + m_{t_c}(\ell_{t_c}) = 0$, at $t = t_c$ this reduces to $\left[\left(\frac{\mathrm{d}}{\mathrm{d}t}m_t(\ell_t)\right)_{t=t_c}\right]^2$, which we computed above (and which is clearly nonzero). This gives the cubic decay and the prefactor, and completes the proof. □

Proof of Lemma 3.5.4. For large (in absolute value) negative A and small $\eta > 0$, by (3.5.10) we have

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}}(\log|\lambda - (u+\mathrm{i}\eta)| - \log|\lambda - (A+\mathrm{i}\eta)|)\mu_t(\lambda)\,\mathrm{d}\lambda \\
&= -\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}}\int_A^u \frac{\lambda - x}{(\lambda-x)^2+\eta^2}\,\mathrm{d}x\,\mu_t(\mathrm{d}\lambda) = -\frac{\mathrm{d}}{\mathrm{d}t}\int_A^u\left[\int_{\mathbb{R}}\mathrm{Re}\,\frac{\mu_t(\mathrm{d}\lambda)}{\lambda-(x+\mathrm{i}\eta)}\right]\mathrm{d}x \\
&= -\,\mathrm{Re}\left[\int_A^u \frac{\mathrm{d}}{\mathrm{d}t}m_t(x+\mathrm{i}\eta)\,\mathrm{d}x\right] = -\frac{1}{2}\,\mathrm{Re}\left[\int_A^u \partial_z(m_t(x+\mathrm{i}\eta)^2)\,\mathrm{d}x\right] \\
&= -\frac{\mathrm{Re}(m_t(u+\mathrm{i}\eta)^2) - \mathrm{Re}(m_t(A+\mathrm{i}\eta)^2)}{2}.
\end{aligned} \tag{3.5.29}$$

We will take $A \to -\infty$ and $\eta \downarrow 0$ in that order. After these two limits, the right-hand side of (3.5.29) reads

$$\frac{\mathrm{Im}(m_t(u))^2 - \mathrm{Re}(m_t(u))^2}{2}.$$

Now we claim that

$$\lim_{A\to-\infty}\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}}\log|\lambda-(A+\mathrm{i}\eta)|\mu_t(\lambda)\,\mathrm{d}\lambda = 0 \tag{3.5.30}$$

for every $\eta > 0$. Indeed, since μ_t has mass one for all t and $\mu_t(r_t) = \mu_t(\ell_t) = 0$, we have

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}} \log|\lambda - (A + \mathrm{i}\eta)|\mu_t(\lambda)\,\mathrm{d}\lambda &= \frac{\mathrm{d}}{\mathrm{d}t}\int_{\ell_t}^{r_t} \log\left|\frac{-\lambda + \mathrm{i}\eta}{A} + 1\right|\mu_t(\lambda)\,\mathrm{d}\lambda \\ &= \int_{\ell_t}^{r_t} \log\left|\frac{-\lambda + \mathrm{i}\eta}{A} + 1\right|\partial_t\mu_t(\lambda)\,\mathrm{d}\lambda.\end{aligned}$$

As $A \to -\infty$, the integrand on the right-hand side tends pointwise to zero, and it is bounded in absolute value by

$$|\partial_t\mu_t(\lambda)| \max\left\{\left|\log\left|\frac{-\ell_t + \mathrm{i}\eta}{A_0} + 1\right|\right|, \left|\log\left|\frac{-r_t + \mathrm{i}\eta}{A_0} + 1\right|\right|\right\}$$

for all $A \geqslant A_0 = A_0(t)$. This is integrable by Lemma 3.5.8 below and Hölder's inequality, so we can conclude the proof of (3.5.30) by dominated convergence.

Thus as $A \to -\infty$ the left-hand side of (3.5.29) tends to

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}} \log|\lambda - (u + \mathrm{i}\eta)|\mu_t(\lambda)\,\mathrm{d}\lambda = \int_{\ell_t}^{r_t} \log|\lambda - (u + \mathrm{i}\eta)|\partial_t\mu_t(\lambda)\,\mathrm{d}\lambda.$$

As $\eta \downarrow 0$, this tends to $\frac{\mathrm{d}}{\mathrm{d}t}\int \log|\lambda - u|\mu_t(\lambda)\,\mathrm{d}\lambda$ by dominated convergence, using for example the dominating function

$$\max\{-\log|\lambda - u|, -\log(1/2), \log\sqrt{(\lambda - u)^2 + 1/2}\}|\partial_t\mu_t(\lambda)|\mathbf{1}_{\lambda\in[\ell_t, r_t]}$$

for $\eta^2 < 1/2$, which is integrable by Lemma 3.5.8 and Hölder's inequality. □

Lemma 3.5.8. *The derivative $\partial_t\mu_t(\lambda)$ is in $L^p(\mathbb{R})$, as a function of λ, for $1 < p < 3/2$.*

Proof. For $\eta > 0$, the Burgers equation (3.5.10) gives

$$\begin{aligned}\partial_t \operatorname{Im}(m_t(\lambda + \mathrm{i}\eta)) &= \operatorname{Im}\left(\partial_z\left(\frac{m_t(\lambda + \mathrm{i}\eta)^2}{2}\right)\right) = \partial_z\left[\operatorname{Re}(m_t(\lambda + \mathrm{i}\eta))\operatorname{Im}(m_t(\lambda + \mathrm{i}\eta))\right] \\ &= \left[\partial_\lambda \operatorname{Re}(m_t(\lambda + \mathrm{i}\eta))\right]\operatorname{Im}(m_t(\lambda + \mathrm{i}\eta)) + \left[\partial_\lambda \operatorname{Im}(m_t(\lambda + \mathrm{i}\eta))\right]\operatorname{Re}(m_t(\lambda + \mathrm{i}\eta)).\end{aligned} \tag{3.5.31}$$

We now consider $\eta \downarrow 0$. As μ_t is analytic on $\{\lambda : \mu_t(\lambda) > 0\}$ [49, Corollary 4], if λ is not an edge or cusp of μ_t,

$$\lim_{\eta \downarrow 0} \partial_\lambda \operatorname{Im}(m_t(\lambda + \mathrm{i}\eta)) = \lim_{\eta \downarrow 0} \left(\pi \int_{\mathbb{R}} \frac{\eta}{(s-\lambda)^2 + \eta^2} \mu_t'(s)\, \mathrm{d}s \right) = \pi \mu_t'(\lambda).$$

As μ_t' is compactly supported and analytic on the set where it does not vanish, this limit is locally uniform in λ. By the same argument, this local uniformity also holds for $\lim_{\eta \downarrow 0} \operatorname{Im}(m_t(\lambda + \mathrm{i}\eta)) = \pi \mu_t(\lambda)$. We argue similarly for the real part (noting that the interchange $\lim_{\eta \downarrow 0} \partial_\lambda \operatorname{Re}(m_t(\lambda + \mathrm{i}\eta)) = \partial_\lambda \operatorname{Re}(m_t(\lambda + \mathrm{i}0^+))$ is simply a rephrasing of the fact that the Hilbert transform commutes with differentiation). Furthermore, [49, Proposition 2, Lemma 3] gives

$$\sup_{z \in \mathbb{H}} |m_t(z)| \leqslant \frac{1}{\sqrt{t}}. \tag{3.5.32}$$

Thus the right-hand side of (3.5.31) tends to $\pi \partial_\lambda [\operatorname{Re}(m_t(\lambda + \mathrm{i}0^+)) \mu_t(\lambda)]$ as $\eta \downarrow 0$, and this limit is locally uniform in λ. This justifies swapping the limit and derivative on the left-hand side of (3.5.31), and dividing through by π we obtain

$$\partial_t \mu_t(\lambda) = \partial_\lambda \Big[\operatorname{Re}(m_t(\lambda + \mathrm{i}0^+)) \mu_t(\lambda) \Big] \tag{3.5.33}$$

for λ not an edge or cusp of μ_t.

Now we prove the regularity claim. Since μ_t decays at most like a cube root near its edges and possible cusps [49, Corollary 5], we have $\partial_\lambda \mu_t \in L^p(\mathrm{d}\lambda)$, for any $1 \leqslant p < 3/2$. Since the Hilbert transform commutes with differentiation and is bounded on L^p for $1 < p < \infty$, we also have $\partial_\lambda(\operatorname{Re}(m_t(\lambda + \mathrm{i}0^+))) \in L^p(\mathrm{d}\lambda)$, for the same range of p values. Expanding the derivative in (3.5.33) and using (3.5.32), we conclude that $\partial_t \mu_t$ is in L^p for $1 < p < 3/2$. □

Proof of Lemma 3.5.5. Assume first that $\operatorname{supp}(\mu_D)$ is connected. By [49, Proposition 3] $\operatorname{supp}(\mu_t)$ is connected for any $t > 0$. By [49, Corollary 4] μ_t has a density that is analytic on $\{x : \mu_t(x) > 0\}$ (although it can have cusps).

Outside of $\operatorname{supp}(\mu_t)$, the function $F(\cdot, t)$ is concave as the sum of concave functions. For $\mu_t(u) >$

0, we compute $\partial_{uu}F(u,t)$ below. For any $\eta = \operatorname{Im} z > 0$ by taking the imaginary part of (3.5.18) we have on the one hand

$$\operatorname{Im} m_t(z) = \int \frac{\mu_D(\mathrm{d}\lambda)(\eta + t \operatorname{Im} m_t(z))}{|\lambda - z - t m_t(z)|^2},$$

i.e.

$$1 = t \int \frac{\mu_D(\mathrm{d}\lambda)}{|\lambda - z - t m_t(z)|^2}, \tag{3.5.34}$$

for $z = u + \mathrm{i}\eta$ and $\eta = 0^+$. On the other hand, differentiation of (3.5.18) gives

$$\operatorname{Re} \partial_z m_t(z) = \operatorname{Re} \frac{X}{1 - tX} = \frac{1}{t} \operatorname{Re} \frac{1}{1 - tX} - \frac{1}{t} \quad \text{with } X := \int \frac{\mu_D(\mathrm{d}\lambda)}{(\lambda - z - t m_t(z))^2}.$$

From (3.5.34), for $z = u + \mathrm{i}0^+$ we have $|tX| \leqslant 1$ so that $\operatorname{Re} \frac{1}{1-tX} \geqslant 0$. Note that by analyticity, $\partial_z m = \partial_u \operatorname{Re} m + \mathrm{i}\partial_u \operatorname{Im} m$, so we have proved

$$\partial_u m_t(u) \geqslant -\frac{1}{t},$$

so that

$$\frac{\partial^2}{\partial u^2} F(u,t) \leqslant \frac{1}{t} - \frac{1}{t} \leqslant 0.$$

Since $F(\cdot, t)$ is differentiable at ℓ_t (with derivative $-\operatorname{Re}(m_t(\ell_t)) - \ell_t/t$) and similarly for r_t, this completes the proof if $\operatorname{supp}(\mu_D)$ is connected. In the general case, write I for the convex hull of $\operatorname{supp}(\mu_D)$, which is necessarily an interval gapped away from zero, write ν_I for uniform measure on I, and consider the probability measures $\mu_D^{(\varepsilon)} = (1-\varepsilon)\mu_D + \varepsilon\nu_I$. We temporarily add the measure to the notation for $F(u,t)$, writing $F(u,t,\mu_D)$. We have $\mu_D^{(\varepsilon)} \to \mu_D$ weakly as $\varepsilon \to 0$; in particular, since $\lambda \mapsto \log(\lambda)$ is bounded and continuous on I, we have

$$\lim_{\varepsilon \to 0} \int_{\mathbb{R}} \log(\lambda) \mu_D^{(\varepsilon)}(\mathrm{d}\lambda) = \int_{\mathbb{R}} \log(\lambda) \mu_D(\mathrm{d}\lambda).$$

Combined with Lemma 3.5.9 below, this lets us conclude that $F(\cdot, t, \mu_D^{(\varepsilon)}) \to F(\cdot, t, \mu_D)$ pointwise as $\varepsilon \to 0$. Since each $\operatorname{supp}(\mu_D^{(\varepsilon)}) = I$ is connected, $F(\cdot, t, \mu_D)$ is thus concave as the pointwise limit

of concave functions. □

Proof of Lemma 3.5.6. Differentiating both sides of (3.5.27) in t and using (3.5.28), we obtain

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\operatorname{Re}(m_t(\ell_t)) &= \int \frac{\partial_t \ell_t + t\frac{\mathrm{d}}{\mathrm{d}t}\operatorname{Re}(m_t(\ell_t)) + \operatorname{Re}(m_t(\ell_t))}{(\lambda - \ell_t - t\operatorname{Re}(m_t(\ell_t)))^2}\mu_D(\mathrm{d}\lambda) \\ &= \frac{\partial_t \ell_t + \operatorname{Re}(m_t(\ell_t))}{t} + \frac{\mathrm{d}}{\mathrm{d}t}\operatorname{Re}(m_t(\ell_t)).\end{aligned}$$

(Differentiability of $m_t(\ell_t)$ was established using (3.5.26).) We note that the idea to study the evolution of the edge by differentiating a self-consistent equation that it satisfies also appears in the proof of [1, Proposition 3.4]. □

Proof of Lemma 3.5.7. Notice that

$$\left(\frac{\partial}{\partial u}F(u,t)\right)_{u=\ell_t} = -\frac{\ell_t}{t} - \operatorname{Re}(m_t(\ell_t)).$$

We work with the right-hand side. We claim that

$$\ell_t + t\operatorname{Re}(m_t(\ell_t)) \leqslant \mathtt{l}(\mu_D). \tag{3.5.35}$$

This is in fact a special case of an inequality established by Guionnet-Maïda in the proof of [102, Lemma 6.1], which says that if μ and ν are compactly supported probability measures and ω is the so-called *subordination function* defined implicitly by

$$\int \frac{(\mu \boxplus \nu)(\mathrm{d}\lambda)}{\lambda - z} = \int \frac{\mu(\mathrm{d}\lambda)}{\lambda - \omega(z)},$$

then

$$\omega(\mathtt{r}(\mu \boxplus \nu)) \geqslant \mathtt{r}(\mu).$$

In our case, $\nu = \rho_{\mathrm{sc},t}$ and $\mu = \mu_D$, so that $\mu \boxplus \nu = \mu_t$, and the Pastur relation (3.5.18) shows that the subordination function is $\omega(z) = z + tm_t(z)$. (In fact, these choices give us results about the

right edge; to get (3.5.35), one should choose $\mu = -\mu_D$, the measure defined by $-\mu_D(A) = \mu_D(-A)$ for Borel A, then track the negative signs.)

Combined with (3.5.28), the result (3.5.35) shows that $w_t = \ell_t + t\,\mathrm{Re}(m_t(\ell_t))$ is a solution to the following constrained problem:

$$\begin{cases} \dfrac{1}{t} = \displaystyle\int \frac{\mu_D(\mathrm{d}\lambda)}{(\lambda - w_t)^2}, \\ w_t \leqslant \mathtt{l}(\mu_D). \end{cases} \tag{3.5.36}$$

A short differential calculation shows that $f(w) = \int \frac{\mu_D(\mathrm{d}\lambda)}{(\lambda-w)^2}$ is *strictly* increasing for $w \leqslant \ell(\mu_D)$, so (3.5.36) has at most one solution. Furthermore, $f(0) = \frac{1}{t_c}$; this means that the unique solution (which we showed is $\ell_t + t\,\mathrm{Re}(m_t(\ell_t))$) must be positive if $0 < t < t_c$, must be zero if $t = t_c$, and must be negative if $t > t_c$. □

Lemma 3.5.9. *Suppose that μ_N is a sequence of probability measures, all supported on some $[a, b]$, tending weakly to some μ_∞ which is also supported on $[a, b]$. Then for every $t > 0$ and every $u \in \mathbb{R}$ we have*

$$\lim_{N\to\infty} \int_{\mathbb{R}} \log|\lambda - u|(\rho_{sc,t} \boxplus \mu_N)(\lambda)\,\mathrm{d}\lambda = \int_{\mathbb{R}} \log|\lambda - u|(\rho_{sc,t} \boxplus \mu_\infty)(\lambda)\,\mathrm{d}\lambda.$$

Proof. For small positive $\eta = \eta_N$ to be chosen, define $\log_\eta : \mathbb{R} \to \mathbb{R}$ by $\log_\eta(x) = \log|x + \mathrm{i}\eta|$. For any probability measure μ supported on $[a, b]$, [49, Corollaries 2, 5] yields

$$(\rho_{\mathrm{sc},t} \boxplus \mu)(\lambda) \leqslant \left(\frac{3}{4\pi^3 t^2}(4 + b - a)\right)^{1/3} \mathbf{1}_{a-2\sqrt{t}\leqslant\lambda\leqslant b+2\sqrt{t}}.$$

Since $\int_{\mathbb{R}} \frac{\log_\eta(\lambda)-\log|\lambda|}{\eta}\,\mathrm{d}\lambda = \pi$, we have

$$\left|\int_{\mathbb{R}} \log|\lambda - u|(\rho_{\mathrm{sc},t} \boxplus \mu)(\lambda)\,\mathrm{d}\lambda - \int_{\mathbb{R}} \log_\eta(\lambda - u)(\rho_{\mathrm{sc},t} \boxplus \mu)(\lambda)\,\mathrm{d}\lambda\right| \leqslant \left(\frac{3}{4\pi^3 t^2}(4 + b - a)\right)^{1/3} \pi\eta$$

which depends on μ only through $[a, b]$. On the other hand, the function $f_{u,\eta}(\lambda) = \log_\eta(\lambda - u)$ is

$\frac{1}{2\eta}$-Lipschitz and bounded on $[a-2\sqrt{t}, b+2\sqrt{t}]$ by

$$\max\{|\log(\eta)|, \left|\log_\eta(b-u+2\sqrt{t})\right|, \left|\log_\eta(a-u-2\sqrt{t})\right|\} = |\log(\eta)|$$

where the equality holds for η sufficiently small depending on a, b, and u. Since combining [71, Corollary 11.65, Theorem 11.3.3] and [48, Proposition 4.13] gives

$$d_{\mathrm{BL}}(\rho_{\mathrm{sc},t} \boxplus \mu_N, \rho_{\mathrm{sc},t} \boxplus \mu_\infty) \leqslant 4\sqrt{d_{\mathrm{BL}}(\mu_N, \mu_\infty)},$$

we bound $|\int_{\mathbb{R}} \log|\cdot - u| \mathrm{d}(\rho_{\mathrm{sc},t} \boxplus \mu_N) - \int_{\mathbb{R}} \log|\cdot - u| \mathrm{d}(\rho_{\mathrm{sc},t} \boxplus \mu_\infty)|$ with

$$\begin{aligned}
&\sum_{\nu=\mu_N,\mu_\infty} \left|\int_{\mathbb{R}} (\log|\cdot - u| - \log_\eta(\cdot - u)) \mathrm{d}(\rho_{\mathrm{sc},t} \boxplus \nu)\right| + \left|\int_{\mathbb{R}} \log_\eta(\cdot - u) \mathrm{d}(\rho_{\mathrm{sc},t} \boxplus \mu_N - \rho_{\mathrm{sc},t} \boxplus \mu_\infty)\right| \\
&\leqslant \left(\frac{3}{4\pi^3 t^2}(4+b-a)\right)^{1/3} \pi\eta + \left(\frac{1}{2\eta} + |\log(\eta)|\right) d_{\mathrm{BL}}(\rho_{\mathrm{sc},t} \boxplus \mu_N, \rho_{\mathrm{sc},t} \boxplus \mu_\infty) \\
&\leqslant \left(\frac{3}{4\pi^3 t^2}(4+b-a)\right)^{1/3} \pi\eta + \left(\frac{1}{2\eta} + |\log(\eta)|\right) 4\sqrt{d_{\mathrm{BL}}(\mu_N, \mu_\infty)},
\end{aligned}$$

for η sufficiently small depending on u. If we choose $\eta = \eta_N = \left(d_{\mathrm{BL}}(\mu_N, \mu_\infty)\right)^{1/4}$, which tends to zero as $N \to \infty$, this upper bound also tends to zero as $N \to \infty$. □

Chapter 4

Complexity of bipartite spherical spin glasses

This chapter is essentially borrowed from [120], which will appear on the arXiv soon.

4.1 Introduction

4.1.1 History and motivations. Multi-species spin systems were first introduced in the 1970s in the physics of metamagnets [109], and in the last fifteen years, their development has been accelerated by applications of two kinds. First, in many social and biological networks it is natural to group individuals into two populations, and the result can be modelled with bipartite spin glasses, for example in immunology with two types of immune cells [3]. Second, certain types of neural networks, such as Hopfield networks and restricted Boltzman machines, can be mapped to bipartite spin systems [26, 4, 28].

Partially motivated by these applications, physical properties like the free energy of bipartite spin glasses have been developed, mostly for what we will later call $(1, 1)$ models with Ising spins or variations thereof. These were treated both in the physics literature, first by Korenblit-Shender

and Fyodorov-Korenblit-Shender [112, 86, 87] and later by Guerra and co-authors [27, 25], both under the assumption of replica symmetry, and then by Hartnett *et al.* assuming replica symmetry breaking [105]; and in the mathematical literature, first as an upper bound due to Barra *et al.* [24] and then a matching lower bound due to Panchenko [129]. The free energy for *spherical* bipartite models was established by Auffinger and Chen at high temperature [11], allowing for mixtures and small external fields, and by Baik and Lee at all temperatures other than some critical one [20], restricted to what we will call pure spherical $(1,1)$ models. Recently, bipartite spin glasses appeared as a model example in Mourrat's program to relate the free energy of disordered systems to infinite-dimensional Hamilton-Jacobi equations [124].

Beyond applications, bipartite spin systems also serve as a toy model for spin glasses beyond the purely mean-field regime. Spins interact with each other in two groups, a waystation between the best-understood mean-field spin glasses (where all spins interact with each other on equal footing) and the eventual goal of spin glasses with nearest-neighbor interactions.

4.1.2 Results. In this paper, we study the *complexity* of high-dimensional bipartite spherical models. That is, write $\mathcal{H}_N$ for an N-dimensional bipartite spin glass, which is a real-valued random function defined on a product of two high-dimensional spheres (see precise definitions in Section 4.2). Write $\mathrm{Crt}_N^{\mathrm{tot}}(t)$ for the (random) number of critical points of $\mathcal{H}_N$ at which $\mathcal{H}_N \leqslant Nt$, and $\mathrm{Crt}_N^{\mathrm{min}}(t)$ for the number of such local minima. We wish to understand the large-N asymptotics of $\frac{1}{N}\log \mathbb{E}[\mathrm{Crt}_N^{\mathrm{tot}}(t)]$ and $\frac{1}{N}\log \mathbb{E}[\mathrm{Crt}_N^{\mathrm{min}}(t)]$.

This landscape-complexity program – counting critical points of high-dimensional random functions to understand their geometry – was initiated by Fyodorov [84] for a certain toy model of disordered systems, and re-discovered by Auffinger-Ben Arous-Černý for spherical spin glasses [10, 9]. Complexity of spherical bipartite models was first studied by Auffinger and Chen [11], who found continuous functions $J, K : \mathbb{R} \to \mathbb{R}$ such that

$$J(t) \leqslant \lim_{N\to\infty} \frac{1}{N}\log \mathbb{E}[\mathrm{Crt}_N^{\mathrm{min}}(t)] \leqslant K(t).$$

Their strategy was to compare bipartite spin glasses with a coupled pair of usual (single-species) spin glasses. They also established that $J(t) > 0$ for some t, so that the system has positive complexity, and that $\lim_{t\to-\infty} K(t) = -\infty$, so that it makes sense to define the "smallest zero of K" which is thus a lower bound for the ground state.

In Theorem 4.2.1 below, we give exact formulas for

$$\lim_{N\to\infty} \frac{1}{N} \log \mathbb{E}[\mathrm{Crt}_N^{\mathrm{tot}}(t)], \qquad \lim_{N\to\infty} \frac{1}{N} \log \mathbb{E}[\mathrm{Crt}_N^{\mathrm{min}}(t)]$$

that are of the form

$$\sup_{u\in\mathfrak{D}} \left\{ \int_{\mathbb{R}} \log|\lambda| \mu_\infty(u,\lambda)\, \mathrm{d}\lambda - \frac{\|u\|^2}{2} \right\}.$$

Here $\mathfrak{D}$ is some subset of $\mathbb{R}$ (for pure models) or $\mathbb{R}^3$ (for mixtures), and the deterministic probability measures $\mu_\infty(u,\cdot)$ are found by solving a system of two coupled quadratic equations in two scalar unknowns. This system arises from the *Matrix Dyson Equation (MDE)*, developed to describe the local eigenvalue behavior of large random matrices in [5, 75, 6].

For the special case of pure (p,q) models with ratio $\gamma = \frac{p}{p+q}$ (see definitions below), the measures $\mu_\infty(u,\cdot)$ are rescalings the semicircle law, so these variational problems can be solved explicitly. The resulting complexity functions turn out to be the same as those describing the pure $p+q$ usual (single-species) spherical spin glass, as established by Auffinger-Ben Arous-Černý [10]; see Corollary 4.2.5 below. This is surprising, since the models look quite different. It remains to be seen if this analogy holds for other types of critical points, such as saddle points, for bipartite models with ratios γ other than $\frac{p}{p+q}$, or for more than two communities.

We also show that pure (p,q) models, with any ratio γ, exhibit a band-of-minima phenomenon similar to pure spherical spin glasses. More precisely, there exists a threshold $-E_\infty(p,q,\gamma) < 0$ such that, with high probability and for any $\varepsilon > 0$, all local minima have energy values below $N(-E_\infty(p,q,\gamma) + \varepsilon)$; see Corollary 4.2.4 below. It would be interesting to understand the role of this threshold in, say, Langevin dynamics.

The paper is organized as follows. In Section 4.2 we state our main results, both variational

formulas for general models and closed-form formulas for the special case stated above. In Section 4.3 we give the proofs, which rely on determinant asymptotics for large random matrices as established in the companion paper [35], and strategies for applying these to complexity as established in the companion paper [36].

Notations. We write $\|\cdot\|$ for the operator norm on elements of $\mathbb{C}^{N\times N}$ induced by Euclidean distance on $\mathbb{C}^N$, and if $\mathcal{S} : \mathbb{C}^{N\times N} \to \mathbb{C}^{N\times N}$, we write $\|\mathcal{S}\|$ for the operator norm induced by $\|\cdot\|$. We let

$$\|f\|_{\mathrm{Lip}} = \sup_{x\neq y}\left|\frac{f(x)-f(y)}{x-y}\right|$$

for test functions $f : \mathbb{R} \to \mathbb{R}$, and consider the following two distances on probability measures on the real line (called bounded-Lipschitz and Wasserstein-1, respectively):

$$d_{\mathrm{BL}}(\mu,\nu) = \sup\left\{\left|\int_{\mathbb{R}} f\,\mathrm{d}(\mu-\nu)\right| : \|f\|_{\mathrm{Lip}} + \|f\|_{L^\infty} \leqslant 1\right\},$$
$$\mathrm{W}_1(\mu,\nu) = \sup\left\{\left|\int_{\mathbb{R}} f\,\mathrm{d}(\mu-\nu)\right| : \|f\|_{\mathrm{Lip}} \leqslant 1\right\}.$$

We write $\mathtt{l}(\mu)$ for the left edge (respectively, $\mathtt{r}(\mu)$ for the right edge) of a compactly supported measure μ. For an $N \times N$ Hermitian matrix M, we write $\lambda_{\min}(M) = \lambda_1(M) \leqslant \cdots \leqslant \lambda_N(M) = \lambda_{\max}(M)$ for its eigenvalues and

$$\hat{\mu}_M = \frac{1}{N}\sum_{i=1}^{N} \delta_{\lambda_i(M)}$$

for its empirical measure. We write $\odot$ for the entrywise (i.e., Hadamard) product of matrices. Given a matrix T, we write $\mathrm{diag}(T)$ for the diagonal matrix of the same size obtained by setting all off-diagonal entries to zero. In equations, we sometimes identify diagonal matrices with vectors of the same size. We write $B_R(0)$ for the ball of radius R about zero in the relevant Euclidean space. We use $(\cdot)^T$ for the matrix transpose, which should be distinguished both from $(\cdot)^*$ for the matrix conjugate transpose, and from $\mathrm{Tr}(\cdot)$ for the matrix trace.

Unless stated otherwise, z will always be a complex number in the upper half-plane $\mathbb{H} = \{z \in$

$\mathbb{C} : \mathrm{Im}(z) > 0\}$, and we always write its real and imaginary parts as $z = E + \mathrm{i}\eta$.

Acknowledgements. We wish to thank Tuca Auffinger for bringing the bipartite spin glass model to our attention, and Gérard Ben Arous and Paul Bourgade for many helpful discussions. BM was supported by NSF grant DMS-1812114.

4.2 Main results

We follow the notation of [11]. If $M \in \mathbb{N}$, write S^M for the $(M-1)$ sphere in $\mathbb{R}^M$ with radius $\sqrt{M}$. Fix some $\gamma \in (0,1)$, suppose that we decompose each positive integer $N \geqslant 2$ as $N = N_1 + N_2$, where N_1 and N_2 are positive integers satisfying $N_1 \approx \gamma N$ in the precise sense

$$\frac{N_1 - 1}{N - 2} = \gamma. \tag{4.2.1}$$

(Notice the abuse of notation: N_1 is actually a sequence of positive integers.) For any $p, q \geqslant 1$, define the pure bipartite Hamiltonian for $u = (u_1, \ldots, u_{N_1}) \in S^{N_1}$ and $v = (v_1, \ldots, v_{N_2}) \in S^{N_2}$ as

$$\mathcal{H}_{N,p,q}(u,v) = \sum_{1 \leqslant i_1, \ldots, i_p \leqslant N_1} \sum_{1 \leqslant j_1, \ldots, j_q \leqslant N_2} g_{i_1,\ldots,i_p,j_1,\ldots,j_q} u_{i_1} \ldots u_{i_p} v_{j_1} \ldots v_{j_q}$$

where the g variables are i.i.d. centered Gaussians with variance $N/(N_1^p N_2^q)$. Equivalently, $\mathcal{H}_{N,p,q}$ is the centered Gaussian process on $S^{N_1} \times S^{N_2}$ with covariance

$$\mathbb{E}[\mathcal{H}_{N,p,q}(u,v)\mathcal{H}_{N,p,q}(u',v')] = N \left(\frac{1}{N_1} \sum_{i=1}^{N_1} u_i u_i' \right)^p \left(\frac{1}{N_2} \sum_{i=1}^{N_2} v_i v_i' \right)^q .$$

Notice that this interaction is genuinely bipartite, i.e. it is not a pure spin glass of the concatenated vector $(u_1, \dots, u_{N_1}, v_1, \dots, v_{N_2})$. Define the "mixed" Hamiltonian

$$\mathcal{H}_N(u, v) = \sum_{p,q \geqslant 1} \beta_{p,q} \mathcal{H}_{N,p,q}(u, v)$$

where the nonnegative double sequence $(\beta_{p,q})_{p,q \geqslant 1}$ is not identically zero and decays fast enough; for example, $\sum_{p,q \geqslant 1} 2^{p+q} \beta_{p,q}^2 < \infty$ suffices. Define $\xi : [0,1]^2 \to \mathbb{R}$ by

$$\xi(x, y) = \sum_{p,q \geqslant 1} \beta_{p,q}^2 x^p y^q$$

assumed to be normalized as

$$\xi(1, 1) = 1.$$

We will say the model is "pure (p_0, q_0)" if $\beta_{p,q} = \delta_{pp_0} \delta_{qq_0}$, and "pure" if it is pure (p_0, q_0) for some p_0, q_0. Define

$$\xi_1' = \partial_x \xi(1,1) = \sum_{p,q \geqslant 1} p \beta_{p,q}^2, \qquad \xi_1'' = \partial_{xx} \xi(1,1) = \sum_{p,q \geqslant 1} p(p-1) \beta_{p,q}^2,$$
$$\xi_2' = \partial_y \xi(1,1) = \sum_{p,q \geqslant 1} q \beta_{p,q}^2, \qquad \xi_2'' = \partial_{yy} \xi(1,1) = \sum_{p,q \geqslant 1} q(q-1) \beta_{p,q}^2.$$

Since $\xi(1,1) = 1$, one can check with Cauchy-Schwarz that $\xi_i'' + \xi_i' - (\xi_i')^2 \geqslant 0$ for each $i = 1, 2$, so that we may define

$$\alpha_i = \sqrt{\xi_i'' + \xi_i' - (\xi_i')^2}.$$

Notice that $\alpha_1 = \alpha_2 = 0$ if and only if the model is pure.

Results. Write $\mathrm{Crt}_N^{\mathrm{tot}}(t)$ for the number of critical points of $\mathcal{H}_N$ at which $\mathcal{H}_N \leqslant Nt$, and $\mathrm{Crt}_N^{\mathrm{tot}}$ for the total number of critical points of $\mathcal{H}_N$. Write also $\mathrm{Crt}_N^{\mathrm{min}}(t)$ for the number of local minima of $\mathcal{H}_N$ at which $\mathcal{H}_N \leqslant Nt$, and $\mathrm{Crt}_N^{\mathrm{min}}$ for the total number of local minima of $\mathcal{H}_N$.

In the statement, we will need the half-space

$$H_t = \{(u_0, u_1, u_2) : u_0 \leqslant t\} \subset \mathbb{R}^3.$$

Theorem 4.2.1. *Suppose that*

$$\xi_1'' > 0 \quad \textit{and} \quad \xi_2'' > 0. \tag{4.2.2}$$

This condition is satisfied if and only if the model is neither a pure $(1, q)$ spin for some q, nor a pure $(p, 1)$ spin for some p.

For each $u \in \mathbb{R}^3$, there exists a compactly supported probability measure $\mu_\infty(u)$ with a bounded density $\mu_\infty(u, \cdot)$ (see Remark 4.2.2 below for its definition) such that, if we define

$$\mathcal{S}_{\mathrm{bsg}}[u] = \int_{\mathbb{R}} \log|\lambda| \mu_\infty(u, \lambda)\, \mathrm{d}\lambda - \frac{\|u\|_2^2}{2},$$

then

$$\begin{aligned} \Sigma^{\mathrm{tot}}(t) &:= \lim_{N\to\infty} \frac{1}{N} \log \mathbb{E}[\mathrm{Crt}_N^{\mathrm{tot}}(t)] = \frac{1 + \gamma \log\left(\frac{\gamma}{\xi_1'}\right) + (1-\gamma)\log\left(\frac{1-\gamma}{\xi_2'}\right)}{2} + \sup_{u \in H_t} \mathcal{S}_{\mathrm{bsg}}[u], \\ \Sigma^{\mathrm{tot}} &:= \lim_{N\to\infty} \frac{1}{N} \log \mathbb{E}[\mathrm{Crt}_N^{\mathrm{tot}}] = \frac{1 + \gamma \log\left(\frac{\gamma}{\xi_1'}\right) + (1-\gamma)\log\left(\frac{1-\gamma}{\xi_2'}\right)}{2} + \sup_{u \in \mathbb{R}^3} \mathcal{S}_{\mathrm{bsg}}[u], \end{aligned} \tag{4.2.3}$$

and these suprema are achieved, possibly not uniquely.

Furthermore, define the set

$$\mathcal{G} = \{u \in \mathbb{R}^3 : \mu_\infty(u)((-\infty, 0)) = 0\} \tag{4.2.4}$$

of u values whose corresponding measures $\mu_\infty(u)$ are supported in the right half-line. Then $\mathcal{G}$ is

convex and closed, and we have

$$\Sigma^{\min}(t) := \limsup_{N\to\infty} \frac{1}{N} \log \mathbb{E}[\mathrm{Crt}_N^{\min}(t)] = \frac{1 + \gamma \log\left(\frac{\gamma}{\xi_1'}\right) + (1-\gamma)\log\left(\frac{1-\gamma}{\xi_2'}\right)}{2} + \sup_{u \in H_t \cap \mathcal{G}} \mathcal{S}_{\mathrm{bsg}}[u],$$
$$\Sigma^{\min} := \limsup_{N\to\infty} \frac{1}{N} \log \mathbb{E}[\mathrm{Crt}_N^{\min}] = \frac{1 + \gamma \log\left(\frac{\gamma}{\xi_1'}\right) + (1-\gamma)\log\left(\frac{1-\gamma}{\xi_2'}\right)}{2} + \sup_{u \in \mathcal{G}} \mathcal{S}_{\mathrm{bsg}}[u], \tag{4.2.5}$$

and these suprema are achieved, possibly not uniquely ($\mathcal{H}_t \cap \mathcal{G}$ is nonempty for every t).

Remark 4.2.2. *The measures $\mu_\infty(u)$ are found as follows: For each $u = (u_0, u_1, u_2) \in \mathbb{R}^3$ and each $z \in \mathbb{H}$, let $\{m_1(u,z), m_2(u,z)\} \in \mathbb{C}^2$ be the unique pair satisfying*

$$\begin{cases} 1 + \Big(z - \frac{1}{\gamma}(\alpha_1 u_1 - \xi_1' u_0) + \frac{\xi_1''}{\gamma} m_1(u,z) + \frac{\xi_1'\xi_2'}{\gamma} m_2(u,z)\Big) m_1(u,z) = 0, \\ 1 + \Big(z - \frac{1}{1-\gamma}(\alpha_2 u_2 - \xi_2' u_0) + \frac{\xi_2''}{1-\gamma} m_2(u,z) + \frac{\xi_1'\xi_2'}{1-\gamma} m_1(u,z)\Big) m_2(u,z) = 0, \\ \mathrm{Im}(m_1(u,z)) > 0, \\ \mathrm{Im}(m_2(u,z)) > 0. \end{cases} \tag{4.2.6}$$

Then $\mu_\infty(u)$ is the measure whose Stieltjes transform at z is $\gamma m_1(u,z) + (1-\gamma) m_2(u,z)$. (In Lemma 4.3.5 below, we will prove all of the implicit claims here about existence, uniqueness, and regularity using the methods of Erdős and co-authors).

Remark 4.2.3. *In the special case when the model is pure (p,q), the result simplifies somewhat: The terms $\alpha_1 u_1$ and $\alpha_2 u_2$ vanish in (4.2.6), so $\mu_\infty((u_0,u_1,u_2))$ is a function of u_0 only. Thus $\mathcal{G}$ takes the form*

$$\mathcal{G} = \{u_0 \times \mathbb{R}^2 : u_0 \in \mathcal{G}_{\mathrm{pure}}\}$$

for some set $\mathcal{G}_{\mathrm{pure}} = \mathcal{G}_{\mathrm{pure}}(p,q,\gamma) \subset \mathbb{R}$. Since $\mathcal{G}$ is convex and closed, and the proof shows that it contains points whose first coordinates are arbitrarily large and negative, in fact $\mathcal{G}_{\mathrm{pure}}$ must be an interval of the form

$$\mathcal{G}_{\mathrm{pure}} = (-\infty, -E_\infty(p,q,\gamma)] \tag{4.2.7}$$

for some $E_\infty(p,q,\gamma)$, which will turn out to be an important threshold. (This notation and sign convention is intended to invoke [10]; see discussion below.)

One consequence of this simplification is that the variational problems for pure (p,q) models are one-dimensional:

$$\lim_{N\to\infty}\frac{1}{N}\log\mathbb{E}[\mathrm{Crt}_N^{\mathrm{tot}}(t)] = \frac{1+\gamma\log\left(\frac{\gamma}{p}\right)+(1-\gamma)\log\left(\frac{1-\gamma}{q}\right)}{2}+\max_{u_0\leqslant t}\mathcal{S}_{\mathrm{bsg}}[(u_0,0,0)],$$
$$\lim_{N\to\infty}\frac{1}{N}\log\mathbb{E}[\mathrm{Crt}_N^{\mathrm{tot}}] = \frac{1+\gamma\log\left(\frac{\gamma}{p}\right)+(1-\gamma)\log\left(\frac{1-\gamma}{q}\right)}{2}+\max_{u_0\in\mathbb{R}}\mathcal{S}_{\mathrm{bsg}}[(u_0,0,0)],$$

and similarly for minima.

Corollary 4.2.4. *For every pure (p,q) model satisfying (4.2.2), the quantity $-E_\infty(p,q,\gamma)$ defined in (4.2.7) is strictly negative, and almost all local minima have energy below $-NE_\infty(p,q,\gamma)$ in the following senses:*

- *For all $t\geqslant -E_\infty(p,q,\gamma)$, we have $\Sigma^{\min}(t)=\Sigma^{\min}(-E_\infty(p,q,\gamma))$.*
- *For any Borel set B, write $\mathrm{Crt}_N^{\min}(B)$ for the number of local minima of $\mathcal{H}_N$ at which $\mathcal{H}_N\in NB$. This corresponds to our previous notation as $\mathrm{Crt}_N^{\min}(t)=\mathrm{Crt}_N^{\min}((-\infty,t))$. For any $\varepsilon>0$, we have*

$$\lim_{N\to\infty}\frac{1}{N}\log\mathbb{P}(\mathrm{Crt}_N^{\min}((-E_\infty(p,q,\gamma)+\varepsilon,\infty))\geqslant 1)=-\infty.$$

In the extra-special case of a pure (p,q) model with $\gamma=\frac{p}{p+q}$, we can solve the variational problems explicitly, because then the relevant Hessian is (almost, up to small error) a generalized Wigner matrix and $\mu_\infty(u)$ is (exactly) a rescaled semicircle law. In the following we write the log-potential of semicircle as

$$\Omega(x)=\int_{-2}^{2}\log|\lambda-x|\frac{\sqrt{4-\lambda^2}}{2\pi}\,\mathrm{d}\lambda=\begin{cases}\frac{x^2}{4}-\frac{1}{2} & \text{if } |x|\leqslant 2,\\ \frac{x^2}{4}-\frac{1}{2}-\left(\frac{|x|}{4}\sqrt{x^2-4}-\log\left(\frac{|x|+\sqrt{x^2-4}}{2}\right)\right) & \text{if } |x|\geqslant 2.\end{cases}$$

Corollary 4.2.5. *For a pure (p,q) model with $\gamma = \frac{p}{p+q}$, we have*

$$E_\infty\Big(p,q,\frac{p}{p+q}\Big) = 2\sqrt{\frac{p+q-1}{p+q}},$$

and

$$\Sigma_{p+q}(t) := \lim_{N\to\infty}\frac{1}{N}\log\mathbb{E}[\mathrm{Crt}_N^{\mathrm{tot}}(t)] = \begin{cases} \frac{1+\log(p+q-1)}{2} + \Omega\Big(t\sqrt{\frac{p+q}{p+q-1}}\Big) - \frac{t^2}{2} & \text{if } t \leqslant 0, \\ \frac{\log(p+q-1)}{2} & \text{if } t \geqslant 0, \end{cases} \tag{4.2.8}$$

$$\Sigma_{p+q} := \lim_{N\to\infty}\frac{1}{N}\log\mathbb{E}[\mathrm{Crt}_N^{\mathrm{tot}}] = \frac{\log(p+q-1)}{2},$$

$$\Sigma_{p+q,\min}(t) := \lim_{N\to\infty}\frac{1}{N}\log\mathbb{E}[\mathrm{Crt}_N^{\min}(t)] = \begin{cases} \Sigma_{p,q}(t) & \text{if } t \leqslant -E_\infty(p,q,\frac{p}{p+q}), \\ \Sigma_{p,q}(-E_\infty(p,q,\frac{p}{p+q})) & \text{if } t \geqslant -E_\infty(p,q,\frac{p}{p+q}), \end{cases}$$

$$\Sigma_{p+q,\min} := \lim_{N\to\infty}\frac{1}{N}\log\mathbb{E}[\mathrm{Crt}_N^{\min}] = \Sigma_{p,q}\Big(-E_\infty\Big(p,q,\frac{p}{p+q}\Big)\Big) = \frac{\log(p+q-1)}{2} + \frac{2}{p+q} - 1.$$

Notice the surprising fact that, as implicit in the notation, these functions depend only on $p+q$ rather than on p and q individually.

Proof. Since u_1 and u_2 play no role, we drop them from the notation. The scalar problem (4.2.6) is solved by $m_1(u_0,z) = m_2(u_0,z) = m(u_0,z)$, which satisfies a quadratic equation given by (4.2.6); this yields

$$\mu_\infty(u_0,\mathrm{d}\lambda) = \frac{\sqrt{(4(p+q-1)(p+q) - (\lambda+(p+q)u_0)^2)_+}}{2\pi(p+q-1)(p+q)}\,\mathrm{d}\lambda.$$

Since the left edge of this measure is explicit, $E_\infty(p,q,\frac{p}{p+q})$ can be computed directly. After changing variables we obtain

$$\mathcal{S}_{\mathrm{bsg}}[(u_0,0,0)] = \log\Big(\sqrt{(p+q-1)(p+q)}\Big) + \Omega\Big(u_0\sqrt{\frac{p+q}{p+q-1}}\Big) - \frac{(u_0)^2}{2}$$

which is even, strictly concave, and uniquely maximized at zero; this allows us to solve the variational problem. □

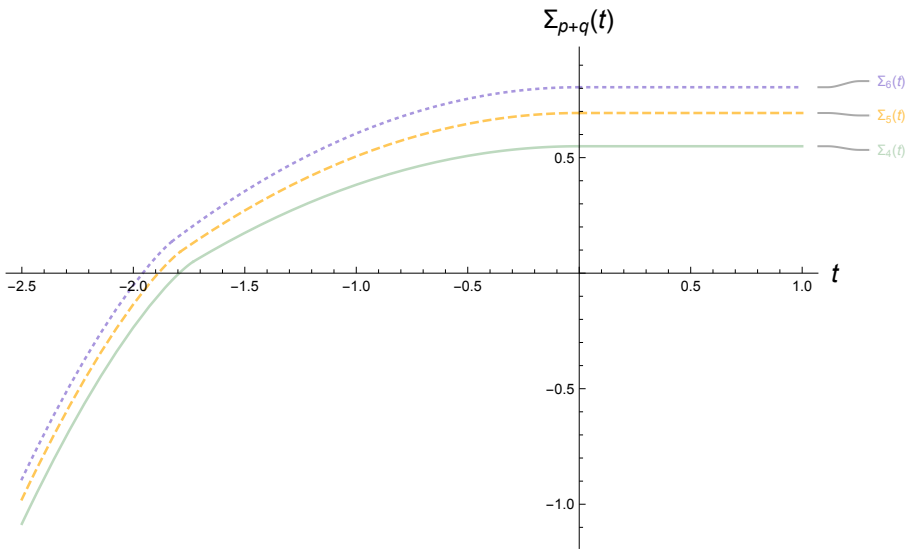


Figure 4.1: Plots of $\Sigma_{p+q}(t)$, which captures the asymptotic complexity of total critical points with field values in $(-\infty, Nt)$ of the pure (p,q) model at $\gamma = \frac{p}{p+q}$, for $p+q = 4,5,6$ (solid green, dashed yellow, dotted purple, respectively). Negative values of $\Sigma_{p+q}(t)$ are irrelevant for us, since we can prove that the zero of Σ_{p+q} is a lower bound for the ground state (and we believe it is equal to the ground state). The functions stabilize at $t = 0$: this is consistent with distributional symmetry $\mathcal{H}_{N,p,q} \stackrel{d}{=} -\mathcal{H}_{N,p,q}$, since we would expect the total number of critical points to be twice the number of critical points with values in $(-\infty, 0)$ on average.

The functions $\Sigma_{p+q}(t)$ are strictly increasing on $(-\infty, 0)$, and $\Sigma_{p+q}(0) > 0$, so they each have a unique zero. They are plotted for $p+q = 4,5,6$ in Figure 4.1. Notice that $p+q = 4$ (corresponding to a pure $(2,2)$ bipartite spin glass) is the smallest value to which Theorem 4.2.1 applies. As a corollary, we obtain a lower bound on the ground state of $\mathcal{H}_{N,p,q}$ in the classical way.

Corollary 4.2.6. *Let $-E_0(p+q)$ be the unique zero of the function Σ_{p+q} defined in (4.2.8), and consider the Hamiltonian $\mathcal{H}_{N,p,q}$ of a pure (p,q) model with $\gamma = \frac{p}{p+q}$. For any $\varepsilon > 0$ there exist $C_1, C_2 > 0$ such that*

$$\mathbb{P}\left(\min_{u,v} \mathcal{H}_{N,p,q}(u,v) \leqslant N(-E_0(p+q) - \varepsilon)\right) \leqslant C_1 \exp(-C_2 N).$$

We compute numerically $-E_0(4) \approx -1.794$, $-E_0(5) \approx -1.888$, *and* $-E_0(6) \approx -1.959$. *Finally,*

$$\lim_{p+q\to\infty} \frac{E_0(p+q)}{\sqrt{\log(p+q)}} = 1.$$

Proof. To locate the ground state, note that if $\min_{u,v} \mathcal{H}_{N,p,q}(u,v) \leqslant N(-E_0(p+q)-\varepsilon)$, then $\mathrm{Crt}_N^{\mathrm{tot}}(-E_0(p+q)-\varepsilon) \geqslant 1$; then apply Markov's. To estimate $-E_0(p+q)$, use the crude bounds $0 \leqslant \Omega(-t\sqrt{\frac{p+q}{p+q-1}}) \leqslant -t\sqrt{\frac{p+q}{p+q-1}}$, valid for all $t \leqslant -2$, to get upper and lower bounds for $\Sigma_{p+q}(t)$ and hence for $-E_0(p+q)$. □

In fact, the functions $\Sigma_{p+q}(t)$ and $\Sigma_{p+q,\min}(t)$ have already appeared in the literature, in [10]: They give exactly the complexities of the numbers of critical points and of local minima, respectively, of a spherical pure $(p+q)$-spin glass below level Nt. That is, define the spherical pure $(p+q)$-spin Hamiltonian $\mathcal{H}_{N,p+q}$ over $\sigma = (\sigma_1, \dots, \sigma_N) \in S^{N-1}$ by

$$\mathcal{H}_{N,p+q}(\sigma) = \frac{1}{N^{(p+q-1)/2}} \sum_{i_1,\dots,i_{p+q}=1}^{N} J_{i_1,\dots,i_{p+q}} \sigma_{i_1} \cdots \sigma_{i_{p+q}},$$

where the J variables are i.i.d. standard Gaussians, and let $\mathrm{Crt}_N^{\text{pure } p+q}(t)$ be the number of critical points (and $\mathrm{Crt}_N^{\text{pure } p+q,\min}(t)$ be the number of local minima) of $\mathcal{H}_{N,p+q}$ at which $\mathcal{H}_{N,p+q} \leqslant Nt$. Then [10, Theorems 2.5, 2.8] show that

$$\lim_{N\to\infty} \frac{1}{N} \log \mathbb{E}[\mathrm{Crt}_N^{\text{pure } p+q}(t)] = \Sigma_{p+q}(t), \qquad \lim_{N\to\infty} \frac{1}{N} \log \mathbb{E}[\mathrm{Crt}_N^{\text{pure } p+q,\min}(t)] = \Sigma_{p+q,\min}(t).$$

(A computation shows that our Σ_{p+q} and $\Sigma_{p+q,\min}$ are their Θ_{p+q} and $\Theta_{0,p+q}$, respectively. We have used their notation for $-E_0(p+q)$ in the same normalization.)

But we emphasize that, despite the superficial similarity between the pure $p+q$-spin Hamiltonian $\mathcal{H}_{N,p+q}$ and the pure bipartite (p,q)-spin Hamiltonian $\mathcal{H}_{N,p,q}$ with $\gamma = \frac{p}{p+q}$, they are different

processes: Their covariance structures are (assuming $N_1 = \gamma N$ for clarity)

$$\mathbb{E}[\mathcal{H}_{N,p+q}(\sigma)\mathcal{H}_{N,p+q}(\sigma')] = N^{1-(p+q)}\left(\sum_{i=1}^{N}\sigma_i\sigma_i'\right)^{p+q},$$
$$\mathbb{E}[\mathcal{H}_{N,p,q}(u,v)\mathcal{H}_{N,p,q}(u',v')] = N^{1-(p+q)}\frac{(p+q)^{p+q}}{p^p q^q}\left(\sum_{i=1}^{\gamma N}u_i u_i'\right)^p\left(\sum_{i=1}^{(1-\gamma)N}v_i v_i'\right)^q.$$

Remark 4.2.7. *We believe that the restriction "neither a pure $(1,q)$ spin nor a pure $(p,1)$ spin" can be relaxed to "not a pure $(1,1)$ spin," which is the restriction in [11]. See Remark 4.3.8 for a discussion of the obstacles.*

4.3 Proofs

Notation. *We write*

$$I_1 = [\![1, N_1 - 1]\!], \qquad I_2 = [\![N_1, N-2]\!].$$

For each $u \in \mathbb{R}^3$, define $A_N(u), A_N'(u) \in \mathbb{R}^{((N_1-1)+(N_2-1))\times((N_1-1)+(N_2-1))}$ by

$$A_N(u) = A_N(u_0,u_1,u_2) = \begin{pmatrix} \frac{N}{N_1}(\alpha_1 u_1 - \xi_1' u_0)\,\mathrm{Id}_{N_1-1} & 0 \\ 0 & \frac{N}{N_2}(\alpha_2 u_2 - \xi_2' u_0)\,\mathrm{Id}_{N_2-1}\end{pmatrix},$$
$$A_N'(u) = A_N'(u_0,u_1,u_2) = \begin{pmatrix} \frac{1}{\gamma}(\alpha_1 u_1 - \xi_1' u_0)\,\mathrm{Id}_{N_1-1} & 0 \\ 0 & \frac{1}{1-\gamma}(\alpha_2 u_2 - \xi_2' u_0)\,\mathrm{Id}_{N_2-1}\end{pmatrix}.$$

Next, we define random matrices $W_N, W_N' \in \mathbb{R}^{((N_1-1)+(N_2-1))\times((N_1-1)+(N_2-1))}$ one block at a time. Write G for an $(N_1-1)\times(N_2-1)$ matrix with i.i.d. centered Gaussian entries, each with variance $\frac{N\xi_1'\xi_2'}{N_1N_2}$. For each $i = 1, 2$, let $G_i = \sqrt{\frac{N(N_i-1)\xi_i''}{N_i^2}}M^{N_i}$, where each M^{N_i} is an $(N_i-1)\times(N_i-1)$ GOE matrix with normalization $\mathbb{E}[(M^{N_i})_{ij}^2] = \frac{1+\delta_{ij}}{N_i-1}$, and where the M^{N_i}'s are independent of each other

and of G. Then we define W_N by

$$W_N = \begin{pmatrix} G_1 & G \\ G^T & G_2 \end{pmatrix}.$$

Let $T_N \in \mathbb{R}^{((N_1-1)+(N_2-1))\times((N_1-1)+(N_2-1))}$ be given entrywise by

$$(T_N)_{jk} = \begin{cases} \sqrt{\frac{N_1^2}{(1+\delta_{ij})\gamma^2 N(N-2)}} & \text{if } j,k \in I_1, \\ \sqrt{\frac{N_1 N_2}{\gamma N(N_2-1)}} = \sqrt{\frac{N_1 N_2}{(1-\gamma)N(N_1-1)}} & \text{if } j \in I_1, k \in I_2 \text{ or } j \in I_2, k \in I_1, \\ \sqrt{\frac{N_2^2}{(1+\delta_{ij})(1-\gamma)^2 N(N-2)}} & \text{if } j,k \in I_2, \end{cases}$$

and let

$$W'_N = T_N \odot W_N.$$

(That is, W'_N is like W_N, but all the variances are multiplied by a carefully chosen factor close to one.) Finally, let

$$H_N(u) = A_N(u) + W_N, \qquad H'_N(u) = A'_N(u) + W'_N.$$

The matrix $H_N(u)$ is the one naturally appearing in the Kac-Rice formula, as we shall see, but it is well approximated by the matrix $H'_N(u)$, which is easier to work with.

While the definitions are fresh, we store the following lemma for later use:

Lemma 4.3.1. *For every $R > 0$ and every $\varepsilon > 0$, we have*

$$\sup_{u \in B_R(0)} \mathbb{P}(\|H_N(u) - H'_N(u)\| \geqslant \varepsilon) = \mathrm{O}_{R,\varepsilon}(e^{-N^{0.49}}).$$

Proof. Write $E_N = W_N - W'_N$. From the definitions, we check that E_N is a matrix of independent Gaussian entries up to symmetry, and that there exists some constant C such that the off-diagonal entries of E_N have variance at most C/N^3 and the diagonal entries have variance at most C/N. If

$\|\cdot\|_{\max}$ is the maximum norm for matrices, we thus have

$$\mathbb{P}\Big(\|E_N - \operatorname{diag}(E_N)\|_{\max} \geqslant \frac{C}{N^{5/4}}\Big) \leqslant \frac{N(N-1)}{2}\mathbb{P}\Big(|\mathcal{N}(0,1)| \geqslant N^{1/4}\Big) \leqslant N^2 e^{-\frac{\sqrt{N}}{2}}.$$

Then

$$\begin{aligned}\mathbb{P}(\|E_N\| \geqslant \varepsilon) &\leqslant \mathbb{P}\Big(\|E_N\| \geqslant \varepsilon, \|E_N - \operatorname{diag}(E_N)\| \leqslant \frac{C}{N^{1/4}}\Big) + N^2 e^{-\frac{\sqrt{N}}{2}} \\ &\leqslant \mathbb{P}\Big(\|\operatorname{diag}(E_N)\| \leqslant \frac{\varepsilon}{2}\Big) + N^2 e^{-\frac{\sqrt{N}}{2}},\end{aligned}$$

where the last inequality holds for N large enough. But now $\operatorname{diag}(E_N)$ has independent Gaussian entries with variance order $1/N$, so $\mathbb{P}(\|\operatorname{diag}(E_N)\| \leqslant \varepsilon/2)$ is order e^{-N} up to polynomial factors in N; thus

$$\mathbb{P}(\|E_N\| \geqslant \varepsilon) = \mathrm{O}(e^{-N^{0.49}}),$$

say. On the other hand, we have

$$\|A_N(u) - A'_N(u)\| = \max\Big\{\Big|\frac{N}{N_1} - \frac{1}{\gamma}\Big||\alpha_1 u_1 - \xi'_1 u_0|, \Big|\frac{N}{N_2} - \frac{1}{1-\gamma}\Big||\alpha_1 u_2 - \xi'_2 u_0|\Big\} = \mathrm{O}\Big(\frac{\|u\|}{N}\Big).$$

Since

$$\mathbb{P}(\|H_N(u) - H'_N(u)\| \geqslant \varepsilon) \leqslant \mathbb{P}\Big(\|W_N - W'_N\| \geqslant \frac{\varepsilon}{2}\Big) + \mathbf{1}_{\|A_N(u) - A'_N(u)\| \geqslant \frac{\varepsilon}{2}},$$

this completes the proof. □

An easy variation on the Kac-Rice arguments found in [11, Equation (27), Lemma 2, Lemma 3, Equation (37)] yields the following lemma.

Lemma 4.3.2. *With the prefactor*

$$f(N_1, N_2) = \frac{2(\pi N_1)^{N_1/2}}{\Gamma(N_1/2)} \cdot \frac{2(\pi N_2)^{N_2/2}}{\Gamma(N_2/2)} \cdot \left(\sqrt{\frac{N}{2\pi}}\right)^3 \left((2\pi N)^{N-2} \cdot \left(\frac{\xi'_1}{N_1}\right)^{N_1-1}\left(\frac{\xi'_2}{N_2}\right)^{N_2-1}\right)^{-1/2},$$

we have

$$\begin{aligned}
\mathbb{E}[\mathrm{Crt}_N^{\mathrm{tot}}(t)] &= f(N_1,N_2)\int_{H_t} e^{-N\frac{\|u\|^2}{2}}\mathbb{E}[|\det(H_N(u))|]\,\mathrm{d}u,\\
\mathbb{E}[\mathrm{Crt}_N^{\mathrm{tot}}] &= f(N_1,N_2)\int_{\mathbb{R}^3} e^{-N\frac{\|u\|^2}{2}}\mathbb{E}[|\det(H_N(u))|]\,\mathrm{d}u,\\
\mathbb{E}[\mathrm{Crt}_N^{\mathrm{min}}(t)] &= f(N_1,N_2)\int_{H_t} e^{-N\frac{\|u\|^2}{2}}\mathbb{E}[|\det(H_N(u))|\mathbf{1}_{H_N(u)}\geqslant 0]\,\mathrm{d}u,\\
\mathbb{E}[\mathrm{Crt}_N^{\mathrm{min}}] &= f(N_1,N_2)\int_{\mathbb{R}^3} e^{-N\frac{\|u\|^2}{2}}\mathbb{E}[|\det(H_N(u))|\mathbf{1}_{H_N(u)}\geqslant 0]\,\mathrm{d}u.
\end{aligned}$$

Notice

$$\lim_{N\to\infty}\frac{1}{N}\log f(N_1,N_2) = \frac{1+\gamma\log\left(\frac{\gamma}{\xi_1'}\right)+(1-\gamma)\log\left(\frac{1-\gamma}{\xi_2'}\right)}{2}.$$

Thus it remains only to understand the integrals appearing in Lemma 4.3.2. We will do this with [35, Theorems 4.1, 4.5] with the choices $\alpha = 1/2$, $p = 2$, and $\mathfrak{D} = \mathbb{R}^3$ or $\mathfrak{D} = H_t$. In the following lemmas, we check the conditions of these theorems.

The matrix $H_N(u)$ belongs both to the class of "Gaussian matrices with a (co)variance profile" and the class of "block-diagonal Gaussian matrices" (with one block) considered in [35, Corollaries 1.8.A, 1.9]. The latter turns out to be more convenient, so we check the regularity assumptions of [35, Corollary 1.9] as well.

Lemma 4.3.3. *Define the matrix*

$$\sigma = \mathrm{diag}\left(\underbrace{\frac{N\xi_1''}{N_1^2},\frac{N\xi_1''}{N_1^2},\ldots}_{N_1-1\ \text{times}},\underbrace{\frac{N\xi_2''}{N_2^2},\frac{N\xi_2''}{N_2^2},\ldots}_{N_2-1\ \text{times}}\right)$$

and consider the linear operators $\mathcal{S}_N,\mathcal{S}_N' : \mathbb{C}^{(N_1-1)\times(N_2-1)}\to\mathbb{C}^{(N_1-1)\times(N_2-1)}$ *defined on block*

matrices $T = \left(\begin{smallmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{smallmatrix}\right)$ *by*

$$\begin{aligned}
&\mathcal{S}_N\left[\begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix}\right] \\
&= \begin{pmatrix} \left(\frac{N\xi_1''}{N_1^2}\operatorname{Tr}(T_{11}) + \frac{N\xi_1'\xi_2'}{N_1N_2}\operatorname{Tr}(T_{22})\right)\operatorname{Id} & 0 \\ 0 & \left(\frac{N\xi_1'\xi_2'}{N_1N_2}\operatorname{Tr}(T_{11}) + \frac{N\xi_2''}{N_2^2}\operatorname{Tr}(T_{22})\right)\operatorname{Id} \end{pmatrix} + \sigma \odot \operatorname{diag}(T), \\
&\mathcal{S}_N'\left[\begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix}\right] \\
&= \begin{pmatrix} \left(\frac{\xi_1''}{\gamma(N_1-1)}\operatorname{Tr}(T_{11}) + \frac{\xi_1'\xi_2'}{\gamma(N_2-1)}\operatorname{Tr}(T_{22})\right)\operatorname{Id} & 0 \\ 0 & \left(\frac{\xi_1'\xi_2'}{(1-\gamma)(N_1-1)}\operatorname{Tr}(T_{11}) + \frac{\xi_2''}{(1-\gamma)(N_2-1)}\operatorname{Tr}(T_{22})\right)\operatorname{Id} \end{pmatrix}.
\end{aligned} \tag{4.3.1}$$

Here $\odot$ *is the entrywise (Hadamard) product of matrices. Suppose also that*

$$\xi_1'' > 0 \quad \textit{and} \quad \xi_2'' > 0.$$

Then each of these operators is flat, in the sense that for some κ *and all* N *we have*

$$T \geqslant 0 \quad \Longrightarrow \quad \frac{1}{\kappa(N-2)}\operatorname{Tr}(T) \leqslant \mathcal{S}_N[T] \leqslant \frac{\kappa}{N-2}\operatorname{Tr}(T)$$

(and similarly for $\mathcal{S}_N'$*). Furthermore, we have*

$$\sup_N \max(\|\mathcal{S}_N\|, \|\mathcal{S}_N'\|) < \infty, \tag{4.3.2}$$

$$\|\mathcal{S}_N - \mathcal{S}_N'\| = \mathrm{O}\Big(\frac{1}{N}\Big). \tag{4.3.3}$$

Proof. Since $\left|\frac{N_1}{N} - \gamma\right| = \mathrm{O}(\frac{1}{N})$ and $\xi_1'', \xi_2'' > 0$, we can find κ such that

$$\frac{1}{\kappa(N-2)} \leqslant \frac{N\xi_1''}{N_1^2}, \frac{N\xi_1'\xi_2'}{N_1N_2}, \frac{N\xi_2''}{N_2^2}, \frac{\xi_1''}{\gamma(N_1-1)}, \frac{\xi_1'\xi_2'}{\gamma(N_2-1)}, \frac{\xi_1'\xi_2'}{(1-\gamma)(N_1-1)}, \frac{\xi_2''}{(1-\gamma)(N_2-1)} \leqslant \frac{\kappa}{N-2}.$$

If $T \geqslant 0$, then $0 \leqslant \sigma \odot \operatorname{diag}(T) \leqslant \frac{\kappa}{N-2}\operatorname{Tr}(T)$; this suffices to prove flatness for both operators. Estimates like

$$\|\sigma \odot \operatorname{diag}(T)\| \leqslant \frac{\kappa}{N-2}\|T\| \tag{4.3.4}$$

and

$$\left|\frac{N\xi_1''}{N_1^2}\operatorname{Tr}(T_{11}) + \frac{N\xi_1'\xi_2'}{N_1N_2}\operatorname{Tr}(T_{22})\right| \leqslant \frac{\kappa_2}{N-2}(|\operatorname{Tr}(T_{11})| + |\operatorname{Tr}(T_{22})|) \leqslant 2\kappa\|T\|$$

establish (4.3.2). Finally, if we define the sequences

$$a_{11}^{(N)} = \frac{N\xi_1''(N-2)}{N_1^2} - \frac{\xi_1''(N-2)}{\gamma(N_1-1)}, \qquad a_{12}^{(N)} = \frac{N\xi_1'\xi_2'(N-2)}{N_1N_2} - \frac{\xi_1'\xi_2'(N-2)}{\gamma(N_2-1)},$$
$$a_{21}^{(N)} = \frac{N\xi_1'\xi_2'(N-2)}{N_1N_2} - \frac{\xi_1'\xi_2'(N-2)}{(1-\gamma)(N_1-1)} \qquad a_{22}^{(N)} = \frac{N\xi_2''(N-2)}{N_2^2} - \frac{\xi_2''(N-2)}{(1-\gamma)(N_2-1)},$$

then using (4.3.4) we conclude $\|\mathcal{S}_N - \mathcal{S}_N'\| \leqslant \max\{|a_{11}^{(N)}| + |a_{12}^{(N)}|, |a_{21}^{(N)}| + |a_{22}^{(N)}|\}$. But we assumed $\frac{N_1-1}{N-2} = \gamma$ in (4.2.1), which tells us $\max(|a_{11}^{(N)}|, |a_{12}^{(N)}|, |a_{21}^{(N)}|, |a_{22}^{(N)}|) = \mathrm{O}(1/N)$; this completes the proof of (4.3.3). □

Lemma 4.3.4. *The random matrices $H_N(u)$ satisfy the assumptions of [35, Corollary 1.9], and furthermore*

$$\liminf_{N\to\infty} \lambda_{\min}(W_N) \geqslant -2\sqrt{\sup_N \|\mathcal{S}_N\|} - 1 \quad a.s. \tag{4.3.5}$$

Proof. For the bounded-mean condition (MS), (4.2.1) tells us that $\frac{N}{N_1}$ and $\frac{N}{N_2}$ are bounded over N, so that

$$\sup_N \|A_N(u)\| = \sup_N \max\left\{\frac{N}{N_1}(|\alpha_1 u_1| + |\xi_1' u_0|), \frac{N}{N_2}(|\alpha_2 u_2| + |\xi_2' u_0|)\right\} = \mathrm{O}(\|u\|). \tag{4.3.6}$$

The mean-field-randomness condition (MF) is clear, since (dropping the superscript since there is

only one matrix) we have

$$s_{jk} = \begin{cases} \frac{N\xi_1''(1+\delta_{jk})}{N_1^2} & \text{if } j,k \in I_1, \\ \frac{N\xi_1'\xi_2'}{N_1 N_2} & \text{if } j \in I_1, k \in I_2 \text{ or } j \in I_2, k \in I_1, \\ \frac{N\xi_2''(1+\delta_{jk})}{N_2^2} & \text{if } j,k \in I_2. \end{cases}$$

Now we check the regularity (R) of the MDE solution. In this context, the operators $\mathscr{S}_i : \mathbb{C}^{N-2} \to \mathbb{C}$ defined by [35, (1.15)] have the form

$$\mathscr{S}_i[\mathbf{r}] = \begin{cases} \frac{N\xi_1''}{N_1^2} \sum_{k\in I_1}(1+\delta_{ik})r_k + \frac{N\xi_1'\xi_2'}{N_1 N_2}\sum_{k\in I_2} r_k & \text{if } i \in I_1, \\ \frac{N\xi_1'\xi_2'}{N_1 N_2} \sum_{k\in I_1} r_k + \frac{N\xi_2''}{N_2^2}\sum_{k\in I_2}(1+\delta_{ik})r_k & \text{if } i \in I_2. \end{cases}$$

The appropriate MDE [35, (1.16)] is a system of $N-2$ coupled scalar equations, with solution $\mathbf{m}(u,z) \in \mathbb{C}^{N-2}$, and we write μ_N for the measure thus obtained. To establish regularity of μ_N, we think of the $N-2$ coupled scalar equations equivalently as a single MDE over matrices in $\mathbb{C}^{(N-2)\times(N-2)}$, by defining $\mathcal{S}_N : \mathbb{C}^{(N-2)\times(N-2)} \to \mathbb{C}^{(N-2)\times(N-2)}$ by

$$\mathcal{S}_N[T] = \operatorname{diag}(\mathscr{S}_1[\operatorname{diag}(T)], \ldots, \mathscr{S}_{N-2}[\operatorname{diag}(T)]).$$

In fact, one can check that $\mathcal{S}_N$ is the same as the operator $\mathcal{S}_N$ defined in (4.3.1). Then we consider the problem

$$\operatorname{Id} + (z \operatorname{Id} - A_N(u) + \mathcal{S}_N[M_N(u,z)])M_N(u,z) = 0 \quad \text{subject to} \quad \operatorname{Im} M_N(u,z) > 0. \tag{4.3.7}$$

But now $M_N(u,z) := \operatorname{diag}(\mathbf{m}(u,z))$ exhibits a solution to (4.3.7), so we can think of $\mu_N(u)$ equivalently as the measure obtained by solving this matrix version of the MDE.

It is easy to show that each $\mathcal{S}_N$ preserves the cone of positive semidefinite matrices, and that $\mathcal{S}_N$ is self-adjoint with respect to the inner product $\langle R, T\rangle = \operatorname{Tr}(R^*T)$. The other regularity properties

of $\mathcal{S}_N$ established in Lemma 4.3.3 let us apply [5, Propositions 2.1, 2.2], which give (a) that each $\mu_N(u)$ admits a density $\mu_N(u,\cdot)$ with respect to Lebesgue measure; (b) that each $\mu_N(u)$ is supported in $[-\kappa(u),\kappa(u)]$, where $\kappa(u) = \sup_N \|A_N(u)\| + 2(\sup_N \|\mathcal{S}_N\|)^{1/2}$ satisfies $\sup_{u\in B_R(0)} \kappa(u) < \infty$ from (4.3.2) and (4.3.6); and (c) that each $\mu_N(u,\cdot)$ is Hölderian, with a Hölder exponent that is universal and a Hölder constant that is uniform over $u \in B_R(0)$. These three conditions ensure that the densities are bounded, uniformly over $u \in B_R(0)$, which finishes checking the regularity assumption (R).

To check (4.3.5), we note that $W_N = H_N(0)$, and that by the above discussion $\mu_N(0)$ is supported in $[-2\sqrt{\sup_N \|\mathcal{S}_N\|}, 2\sqrt{\sup_N \|\mathcal{S}_N\|}]$. Then [6, Theorem 2.4, Remark 2.5(v)] gives

$$\mathbb{P}\left(\lambda_{\min}(W_N) \leqslant -2\sqrt{\sup_N \|\mathcal{S}_N\|} - 1\right) \leqslant \frac{C}{N^{100}}$$

for some constant C, which suffices. □

Lemma 4.3.5. *The measures $\mu_\infty(u)$ discussed in Remark 4.2.2 are well-defined. They admit densities that are bounded and compactly supported locally uniformly in u, and for each R there exists κ with*

$$\sup_{u\in B_R(0)} \mathrm{W}_1(\mu_N(u), \mu_\infty(u)) \leqslant N^{-\kappa}. \tag{4.3.8}$$

Furthermore, there exists $C > 0$ such that

$$\mathbb{E}[|\det(H_N(u))|] \leqslant (C\max(\|u\|, 1))^N. \tag{4.3.9}$$

Finally, for every R and ε we have

$$\lim_{N\to\infty} \frac{1}{N\log N} \log\left[\sup_{u\in B_R(0)} \mathbb{P}(d_{\mathrm{BL}}(\hat{\mu}_{H_N(u)}, \mu_\infty(u)) > \varepsilon)\right] = -\infty. \tag{4.3.10}$$

Proof. With $\mathcal{S}'_N$ as in (4.3.1), consider the following MDE over matrices in $\mathbb{C}^{(N-2)\times(N-2)}$:

$$\mathrm{Id} + (z\,\mathrm{Id} - A'_N(u) + \mathcal{S}'_N[M'_N(u,z)])M'_N(u,z) = 0 \quad \text{subject to} \quad \operatorname{Im} M'_N(u,z) > 0. \tag{4.3.11}$$

One can check that $\mathcal{S}'_N$ preserves the cone of positive semidefinite matrices and that it is self-adjoint with respect to the inner product $\langle R, T\rangle = \operatorname{Tr}(R^*T)$. Thus this problem has a unique solution $M'_N(u,z)$.

In fact we can write $M'_N(u,z)$ much more explicitly. In (4.2.6) we wrote an MDE-type problem for two N-independent scalars $m_1(u,z)$ and $m_2(u,z)$. Now we prove existence and uniqueness of solutions to that problem: Since $\mathcal{S}'_N$ maps into diagonal matrices, we can see directly from the MDE (4.3.11) that $M'_N(u,z)$ must be diagonal. By looking at the MDE componentwise, we see that the entries on the diagonal can only take two values, which we will call $m_1(u,z)$ (for the first $N_1 - 1$ entries) and $m_2(u,z)$ (for the last $N_2 - 1$ entries). With this information, writing (4.3.11) out in components shows that $\{m_1(u,z), m_2(u,z)\}$ is a solution to (4.2.6). Uniqueness for (4.2.6) follows from uniqueness for (4.3.11), since one can check that

$$M'_N(u,z) = \operatorname{diag}\left(\underbrace{m_1(u,z), m_1(u,z), \ldots}_{N_1-1 \text{ times}}, \underbrace{m_2(u,z), m_2(u,z), \ldots}_{N_2-1 \text{ times}}\right)$$

exhibits a solution to (4.3.11) whenever $\{m_1(u,z), m_2(u,z)\}$ solves (4.2.6). Thus (4.2.1) tells us that

$$\frac{1}{N-2}\operatorname{Tr}(M'_N(u,z)) = \frac{N_1 - 1}{N-2} m_1(u,z) + \frac{N_2 - 1}{N-2} m_2(u,z) = \gamma m_1(u,z) + (1-\gamma) m_2(u,z)$$

is actually independent of N, and we write $\mu_\infty(u)$ for the measure with this Stieltjes transform.

Using the regularity of $\mathcal{S}'_N$ established in Lemma 4.3.3, the same arguments as in the proof of Lemma 4.3.4 tell us that each $\mu_\infty(u)$ admits a compactly supported Hölderian density $\mu_\infty(u,\cdot)$ with respect to Lebesgue measure, and that the support, the Hölder constant, and the Hölder coefficient can all be taken uniform over $u \in B_R(0)$.

Now we prove the distance estimate (4.3.8). The general result [35, Proposition 3.1] reduces this problem to estimating the difference between the Stieltjes transforms, and [36, Lemma 3.1] provides a general technique for doing this, assuming inputs which we verified in Lemmas 4.3.3 and 4.3.4.

The proof of the determinant estimate (4.3.9) follows [36, Lemma 4.4], using (4.3.6). The proof of the concentration estimate (4.3.10) follows [36, Lemma 4.5]. □

Lemma 4.3.6. *For every $\varepsilon > 0$ and $R > 0$, we have*

$$\lim_{N\to\infty} \inf_{u\in B_R(0)} \mathbb{P}(\mathrm{Spec}(H_N(u)) \subset [\mathtt{l}(\mu_\infty(u)) - \varepsilon, \mathtt{r}(\mu_\infty(u)) + \varepsilon]) = 1 \tag{4.3.12}$$

and in fact the extreme eigenvalues of $H_N(u)$ converge in probability to the endpoints of $\mu_\infty(u)$.

Proof. In the proof of Lemma 4.3.4, we showed that the measures μ_N are exactly those given by the MDE for the matrices $H_N(u)$. In the same way, one can check that the measures μ_∞ are exactly those given by the MDE for the matrices $H'_N(u)$; this is why we introduced those matrices. The rest of the argument is exactly as in the proof of [36, Lemma 4.6]: it uses the local law of Alt *et al.* [6] to localize the spectrum of $H'_N(u)$ near the support of $\mu_\infty(u)$, then Lemma 4.3.1 to relate H_N to H'_N, and finally (4.3.10) to show that the extreme eigenvalues of H_N do not push inside the support of $\mu_\infty(u)$. □

Lemma 4.3.7. *With $\mathcal{G}_{+\varepsilon}$ as defined in [35, (4.5)] and $\mathcal{G}$ as defined in (4.2.4), we have that each $\mathcal{G}_{+\varepsilon}$ is convex, that $\mathcal{G}_{+1}$ has positive measure, and that*

$$\overline{\bigcup_{\varepsilon>0} \mathcal{G}_{+\varepsilon}} = \mathcal{G} \quad \text{and} \quad \overline{H_t \cap \left(\bigcup_{\varepsilon>0} \mathcal{G}_{+\varepsilon}\right)} = H_t \cap \mathcal{G} \text{ for all } t.$$

Proof. Convexity for $\mathcal{G}_{+\varepsilon}$ is proved exactly as in [36, Lemma 4.7].

For simplicity, we restrict ourselves to u in the quarter space

$$\mathcal{Q} = \{(u_0, u_1, u_2) : u_1 \geqslant 0, u_2 \geqslant 0\}.$$

For $u \in \mathcal{Q}$ we have

$$\lambda_{\min}(A_N(u)) = \min\left\{\frac{N}{N_1}(\alpha_1 u_1 - \xi_1' u_0), \frac{N}{N_2}(\alpha_2 u_2 - \xi_2' u_0)\right\} \geqslant u_0 \min\left\{-\frac{2\xi_1'}{\gamma}, -\frac{2\xi_2'}{1-\gamma}\right\} =: -\kappa_{\text{bsg}} u_0.$$

Combining this with (4.3.5), we find

$$\liminf_{N\to\infty} \lambda_{\min}(H_N(u)) \geqslant -\kappa_{\text{bsg}} u_0 - 2\sqrt{\sup_N \|\mathcal{S}_N\|} - 1$$

for $u \in \mathcal{Q}$. Along with the convergence in probability of $\lambda_{\min}(H_N(u))$ to $\mathtt{l}(\mu_\infty(u))$ of Lemma 4.3.6, this shows that $\mathcal{G}_{+1}$ has positive measure.

Finally, we note that the inclusion $\cup_{\varepsilon>0}\mathcal{G}_{+\varepsilon} \subset \mathcal{G}$ is clear, and that $\mathcal{G}$ is closed by [35, Lemma 4.6]. To show the reverse inclusion, write $e_1 = (1,0,0)$; then for $\delta > 0$ we have $A_N(u - \delta e_1) \geqslant A_N(u) + \frac{\kappa_{\text{bsg}}}{4}\delta \operatorname{Id}$, so that by the convergence in probability of Lemma 4.3.6 we have $\mathtt{l}(\mu_\infty(u - \delta e_1)) \geqslant \mathtt{l}(\mu_\infty(u)) + \frac{\kappa_{\text{bsg}}}{4}\delta$. This completes the proof of the equality $\overline{\cup_\varepsilon \mathcal{G}_{+\varepsilon}} = \mathcal{G}$. The version intersected with H_t is an exercise in point-set topology, since $\cup_\varepsilon \mathcal{G}_{+\varepsilon}$ is convex as a union of nested convex sets, H_t is a half-space, and their intersection has non-empty interior by the arguments above. □

Proof of Theorem 4.2.1. By the discussion after the proof of Lemma 4.3.2, to show (4.2.3) it suffices to show

$$\lim_{N\to\infty} \frac{1}{N} \log \int_{H_t} e^{-N\frac{\|u\|^2}{2}} \mathbb{E}[|\det(H_N(u))|]\, \mathrm{d}u = \sup_{u\in H_t} \mathcal{S}_{\text{bsg}}[u],$$
$$\lim_{N\to\infty} \frac{1}{N} \log \int_{\mathbb{R}^3} e^{-N\frac{\|u\|^2}{2}} \mathbb{E}[|\det(H_N(u))|]\, \mathrm{d}u = \sup_{u\in \mathbb{R}^3} \mathcal{S}_{\text{bsg}}[u].$$

This is a direct consequence of [35, Theorem 4.1], whose conditions we have checked in the preceding lemmas, with the choices $\alpha = 1/2$, $p = 2$ (recall $N = (N-2)+2$ is two more than the size of $H_N(u)$), and $\mathfrak{D} = \mathbb{R}^3$ or $\mathfrak{D} = H_t$.

Similarly, to prove (4.2.5), it suffices to show

$$\limsup_{N\to\infty}\frac{1}{N}\log\int_{H_t} e^{-N\frac{\|u\|^2}{2}}\mathbb{E}[|\det(H_N(u))|\mathbf{1}_{H_N(u)}\geqslant 0]\,\mathrm{d}u=\sup_{u\in\mathcal{G}\cap\mathcal{H}_t}\mathcal{S}_{\mathrm{bsg}}[u],$$
$$\limsup_{N\to\infty}\frac{1}{N}\log\int_{\mathbb{R}^3} e^{-N\frac{\|u\|^2}{2}}\mathbb{E}[|\det(H_N(u))|\mathbf{1}_{H_N(u)}\geqslant 0]\,\mathrm{d}u=\sup_{u\in\mathcal{G}}\mathcal{S}_{\mathrm{bsg}}[u].$$

This is a direct consequence of [35, Theorem 4.5], whose conditions we have also just checked, with the same choices of parameters. □

Proof of Corollary 4.2.4. Directly from the MDE (4.2.6), we obtain the symmetry

$$\mu_\infty(-u,\lambda)=\mu_\infty(u,-\lambda).$$

In particular, $\mu_\infty(0,\lambda)$ is an even function of λ, so its left edge is strictly negative, hence $u=0$ is not an element of $\mathcal{G}_{\mathrm{pure}}$. Since $\mathcal{G}_{\mathrm{pure}}$ has the form $(-\infty,-E_\infty(p,q,\gamma)]$ we conclude $-E_\infty(p,q,\gamma)<0$. Since $\Sigma^{\min}(t)=\text{constant}+\sup_{u\in\mathcal{G}_{\mathrm{pure}}\cap(-\infty,t]}\mathcal{S}_{\mathrm{bsg}}[u]$ from (4.2.5), we conclude that $\Sigma^{\min}(t)$ stabilizes at $t=-E_\infty(p,q,\gamma)$.

For each ε, consider the half-space

$$\widetilde{H}_\varepsilon=\{(u_0,u_1,u_2)\in\mathbb{R}^3:u_0\geqslant -E_\infty(p,q,\gamma)+\varepsilon\}.$$

By Markov's and a Kac-Rice argument, we have

$$\begin{aligned}\mathbb{P}(\mathrm{Crt}_N^{\min}((-E_\infty(p,q,\gamma)+\varepsilon,\infty))\geqslant 1)&\leqslant\mathbb{E}[\mathrm{Crt}_N^{\min}((-E_\infty(p,q,\gamma)+\varepsilon,\infty))]]\\&=f(N_1,N_2)\int_{\widetilde{H}_\varepsilon}e^{-N\frac{\|u\|^2}{2}}\mathbb{E}[|\det(H_N(u))|\mathbf{1}_{H_N(u)\geqslant 0}]\,\mathrm{d}u.\end{aligned}$$

In the companion paper [35, (4.5)] we considered a sequence of nested sets $\mathcal{G}_{-\delta}$ defined by

$$\mathcal{G}_{-\delta}=\{u\in\mathbb{R}^m:\mu_\infty(u)((-\infty,-\delta))\leqslant\delta\}.$$

Since $\mu_\infty(u)$ depends only on u_0 in the pure case we are currently considering, each $\mathcal{G}_{-\delta}$ is of the form $\mathcal{G}_{-\delta} = \{u_0 \times \mathbb{R}^2 : u_0 \in \mathcal{G}_{\text{pure},-\delta}\}$ for some set $\mathcal{G}_{\text{pure},-\delta} \subset \mathbb{R}$. In fact, we claim that $\mathcal{G}_{\text{pure},-\delta}$ is an interval of the form

$$\mathcal{G}_{\text{pure},-\delta} = (-\infty, f(\delta)]. \tag{4.3.13}$$

Assume this claim momentarily. From the definitions one can see that $\cap_{\delta>0}\mathcal{G}_{-\delta} = \mathcal{G}$, and thus (4.2.7) tells us that $\lim_{\delta\downarrow 0} f(\delta) = -E_\infty(p,q,\gamma)$. Hence there exists a small $\delta = \delta(\varepsilon) > 0$ with $f(\delta) < -E_\infty(p,q,\gamma) + \varepsilon$. For this δ we therefore have $\widetilde{H}_\varepsilon \subset (\mathcal{G}_{-\delta})^c$, but we showed in [35, Lemma 4.7] that

$$\lim_{N\to\infty} \frac{1}{N} \log \int_{(\mathcal{G}_{-\delta})^c} e^{-N\frac{\|u\|^2}{2}} \mathbb{E}[|\det(H_N(u))| \mathbf{1}_{H_N(u)\geqslant 0}]\, \mathrm{d}u = -\infty$$

for every $\delta > 0$. This completes the proof, modulo (4.3.13).

Now we prove (4.3.13). Since $\mu_\infty((u_0,u_1,u_2))$ depends on u_0 only, we abuse notation and write $\mu_\infty(u_0)$. Notice that $\mu_\infty(u_0)$ is the limiting empirical measure of the random matrix $W_N + u_0 B_N$, where

$$B_N = A_N(1,0,0) = -\begin{pmatrix} \frac{N\xi_1'}{N_1} & 0 \\ 0 & \frac{N\xi_2'}{N_2} \end{pmatrix}$$

has strictly negative eigenvalues. The Courant-Fischer variational characterization of eigenvalues gives that, for each $i \in [\![1, N]\!]$, the ith eigenvalue of $W_N + u_0 B_N$ is a non-increasing function of u_0. Thus

$$\frac{1}{N}\#\{i : \lambda_i(W_N + u_0 B_N) < -\delta\}$$

is almost surely non-decreasing in u_0, hence its $N \to +\infty$ limit $\mu_\infty(u_0)((-\infty,-\delta))$ is also non-decreasing in u_0. This shows that $\mathcal{G}_{-\delta}$ is a single interval containing arbitrarily large negative values. From its definition and continuity of the map $u \mapsto \mu_\infty(u)$ (see the proof of [35, Lemma 4.6]) one can see that it is closed, which completes the proof of (4.3.13). □

Remark 4.3.8. *We remark briefly on the restriction $\xi_1'' > 0$ and $\xi_2'' > 0$, which is equivalent to "neither a pure $(1,q)$ spin nor a pure $(p,1)$ spin." In order to apply our Laplace-method arguments, we need the measures $\mu_\infty(u)$ to admit densities. We know of two strategies to show that a measure*

induced by the MDE admits a density: Either check that the stability operator $\mathcal{S}$ in the MDE is flat and import results of [5] (which is our strategy here), or "by hand," meaning manipulate the MDE in a clever way to show that the solution $M_N(u,z)$ satisfies $\sup_{N,u,z} \|M_N(u,z)\| < \infty$ (which is our strategy for the "elastic-manifold" model in [36]).

If $\min(\xi_1'', \xi_2'') = 0$, the stability operators $\mathcal{S}_N$ and $\mathcal{S}_N'$ are not flat: For example, if $\xi_1'' = 0$, then

$$\mathcal{S}_N\left[\begin{pmatrix} T_{11} & 0 \\ 0 & 0 \end{pmatrix}\right] = \begin{pmatrix} 0 & 0 \\ 0 & \frac{N\xi_1'\xi_2'}{N_1N_2}\operatorname{Tr}(T_{11})\operatorname{Id} + \frac{N\xi_2''}{N_2^2}\operatorname{diag}(T_{11}) \end{pmatrix} \not\geq \frac{1}{\kappa(N-2)}\operatorname{Tr}(T).$$

The missing piece is thus to establish regularity "by hand," which we do not know how to do for this model.

Chapter 5

Large deviations for extreme eigenvalues of deformed Wigner random matrices

This chapter is essentially borrowed from [121], which appeared *Electronic Journal of Probability*.

5.1 Introduction

5.1.1 Deformed ensembles: typical behavior.

In this paper, our goal is to prove a large deviation principle (LDP) for the largest eigenvalue of the random matrix

$$X_N = \frac{W_N}{\sqrt{N}} + D_N. \tag{5.1.1}$$

Here $\frac{W_N}{\sqrt{N}}$ lies in a particular class of real or complex Wigner matrices. Specifically, we will ask that the laws of the entries of W_N have sub-Gaussian Laplace transforms with certain variances, and that these laws satisfy concentration properties. The archetypal examples of this class are the

Gaussian ensembles (GOE and GUE). We also assume that D_N is a deterministic matrix whose empirical spectral measure tends to a deterministic limit μ_D and whose extreme eigenvalues tend to the edges of μ_D. In all of our proofs we will assume that D_N is diagonal, but by rotational invariance, our results hold for the deformed Gaussian models even when D_N is not diagonal. More details on our assumptions will be given in Section 5.2.

If we write $\lambda_1(M) \leqslant \cdots \leqslant \lambda_N(M)$ for the eigenvalues of a self-adjoint matrix M and $\hat{\mu}_M = \frac{1}{N}\sum_{i=1}^N \delta_{\lambda_i(M)}$ for its empirical measure, it is well-known that

$$\hat{\mu}_{X_N} \to \rho_{\mathrm{sc}} \boxplus \mu_D,$$

both almost surely and in expectation, where ρ_{sc} is the semicircle law normalized as $\rho_{\mathrm{sc}}(\mathrm{d}x) = \frac{1}{2\pi}\sqrt{(4-x^2)_+}\,\mathrm{d}x$ and $\mu \boxplus \nu$ is the free convolution of the probability measures μ and ν [131, 153].

If μ is a compactly supported measure on $\mathbb{R}$, we write $\mathtt{l}(\mu)$ and $\mathtt{r}(\mu)$ for the left and right endpoints, respectively, of its support. For some special cases of our model, it is known that

$$\lambda_N(X_N) \to \mathtt{r}(\rho_{\mathrm{sc}} \boxplus \mu_D) \quad \text{almost surely.}$$

New cases will be a corollary of our large deviation principle; see Remark 5.2.6 below for details.

Our model also exhibits edge universality for many choices of D_N; that is, the fluctuations of $\lambda_N(X_N)$, rescaled appropriately, are known to follow the Tracy-Widom distribution. This was first established by [140] for the deformed GUE, if $\hat{\mu}_{D_N} \to \mu_D$ quickly ($d(\hat{\mu}_{D_N}, \mu_D) = O(N^{-2/3-\varepsilon})$ is enough, where d is defined in Equation (5.1.5)) and without outliers. The convergence-rate assumption was removed by [65], which also allowed a finite number of outliers in a controlled way, under a technical assumption implying that μ_D does not decay too quickly near its edges. The assumption of Gaussianity was removed by [115], under a similar technical assumption on μ_D.

5.1.2 History of large deviations in random matrix theory.

The history of LDPs for random matrix theory is fairly sparse. The first result, from [42], is

for the empirical measure of the Gaussian ensembles. The first LDP for the largest eigenvalue of a random matrix ensemble, namely for the GOE, appeared in [37]. We mention also [79] for the largest eigenvalue of thin sample covariance matrices, and [54] for the empirical measure and [14] for the largest eigenvalue of Wigner matrices whose entries have tails heavier than Gaussian.

There are also several results for the large deviations of deformed random matrices. For example, the paper [104] studied large deviations of the empirical measure of full-rank deformations of Gaussian ensembles, making rigorous a prediction from [119]. The largest eigenvalue of a rank-one deformation of a Gaussian ensemble was studied by [118]; this result was recovered as the time-one marginal of a large deviation principle for Hermitian Brownian motions in [70]. Finite-rank deformations, rather than rank-one deformations, were covered in [45].

Our work builds on the recent papers [99] and [102]. These works use techniques discussed below to establish LDPs for extreme eigenvalues, treating respectively sharp sub-Gaussian Wigner matrices and the free-convolution model $A + UBU^*$ (with U Haar orthogonal or Haar unitary). This method was also adapted in [50] to study joint large deviations of the largest eigenvalue and of one component of the corresponding eigenvector for rank-one deformations of Gaussian ensembles. Very recently, [15] adapted this method to study non-sharp sub-Gaussian Wigner matrices; see Remark 5.2.2 below for a precise explanation of this terminology.

5.1.3 Large deviations for ensembles with full-rank deformations.

In many large-deviations proofs, one wants to tilt measures by a Laplace transform. The insight of the paper [99] was that the appropriate Laplace transform in our context is the so-called *(rank-one) spherical integral*

$$\mathbb{E}_e[e^{N\theta\langle e, Me\rangle}]. \tag{5.1.2}$$

Here M is an $N \times N$ self-adjoint matrix, $\theta \geqslant 0$ is the argument of the Laplace transform, and the integration $\mathbb{E}_e$ is over vectors e uniform on the unit sphere $\mathbb{S}^{N-1}$ (we take $\mathbb{S}^{N-1} \subset \mathbb{R}^N$ if M is real, or $\mathbb{S}^{N-1} \subset \mathbb{C}^N$ if M is complex, so that (5.1.2) is real). If M is a random matrix, then (5.1.2) is a

random variable. This is a special case of the famous Harish-Chandra/Itzykson/Zuber integral.

For an LDP for the model (5.1.1), we encounter two technical challenges. If we write $\mathbb{P}_N$ for the law of X_N and $\mathbb{E}_{X_N}$ for the corresponding expectation (and define $\mathbb{E}_{W_N}$ in the obvious way), then the first challenge is the computation of

$$\lim_{N\to\infty} \frac{1}{N} \log \mathbb{E}_{X_N}[\mathbb{E}_e[e^{N\theta\langle e, X_N e\rangle}]] = \lim_{N\to\infty} \frac{1}{N} \log \mathbb{E}_e[\mathbb{E}_{W_N}[e^{\sqrt{N}\theta\langle e, W_N e\rangle}] \cdot e^{N\theta\langle e, D_N e\rangle}]. \tag{5.1.3}$$

The term $\mathbb{E}_{X_N}[\mathbb{E}_e[e^{N\theta\langle e, X_N e\rangle}]]$ appears as a normalization constant when tilting the measure, so its logarithmic asymptotics appear as part of the rate function. To understand these asymptotics when W_N is not Gaussian, we use the method of [99, Lemma 3.2] to understand $\mathbb{E}_{W_N}[e^{\sqrt{N}\theta\langle e, W_N e\rangle}]$ pointwise for unit vectors e that are delocalized in an appropriate sense. We combine this with the new result (see Lemma 5.4.4 below)

$$\text{for } \theta \text{ small enough depending on } \mu_D, \qquad \lim_{N\to\infty} \frac{1}{N} \log \left[\frac{\mathbb{E}_e[\mathbf{1}_{e \text{ delocalized}} e^{N\theta\langle e, D_N e\rangle}]}{\mathbb{E}_e[e^{N\theta\langle e, D_N e\rangle}]} \right] = 0.$$

The qualifier "for θ small enough" means that, via this argument, we can only obtain large-deviations asymptotics of events that localize $\lambda_N(X_N)$ below some critical threshold x_c, which depends on the deformation μ_D only. We show $x_c \geqslant \mathtt{r}(\rho_{\mathrm{sc}} \boxplus \mu_D)$ with strict inequality except in degenerate cases, and that x_c can be infinite. For example, $x_c = +\infty$ when μ_D is the uniform measure on an interval. For the Gaussian ensembles, the limit in (5.1.3) is directly computable for every $\theta \geqslant 0$ without recourse to this delocalization problem, so our results for those models are stronger.

The second difficulty (in some respects the main one) is that we need a concentration result of the form

$$\lim_{N\to\infty} \frac{1}{N} \log \mathbb{P}_N(d(\hat{\mu}_{X_N}, \rho_{\mathrm{sc}} \boxplus \mu_D) > N^{-\kappa}) = -\infty \tag{5.1.4}$$

for $\kappa > 0$ small enough, where d is defined in (5.1.5). This result is Lemma 5.5.3 below. With $\rho_{\mathrm{sc}} \boxplus \mu_D$ replaced with $\mathbb{E}[\hat{\mu}_{X_N}]$, this is standard concentration of linear statistics [103], easily

extended to our model. To approximate $\mathbb{E}[\hat{\mu}_{X_N}]$ with $\rho_{\mathrm{sc}} \boxplus \mu_D$, we use local laws for deformed ensembles [116, 115, 75]. Our argument is slightly technical, since these local laws let us approximate $\mathbb{E}[\hat{\mu}_{X_N}]$, not directly by $\rho_{\mathrm{sc}} \boxplus \mu_D$, but by a measure close to $\rho_{\mathrm{sc}} \boxplus \hat{\mu}_{D_N}$, so several intermediate comparisons are needed.

The organization of the paper is as follows: In Section 5.2, we state our assumptions and main result with commentary and examples. In Section 5.3, we provide background on spherical integrals, introduce the tilted measures, and provide a high-level overview of the technique as well as proofs of weak-large-deviations upper and lower bounds. These arguments rely on several key lemmas, the proofs of which make up the remaining three sections. In Section 5.4, we address the first technical issue discussed above. In Section 5.5, we prove exponential tightness for our model, then address the second technical issue discussed above. In Section 5.6, we establish properties of the rate function. Throughout, our results are stated for both the real and complex cases, but we only give proofs in the real case. The proofs in the complex case require only minor modifications.

Conventions. We use the shorthand β for the symmetry class at hand: $\beta = 1$ refers to real symmetric matrices and $\beta = 2$ refers to complex Hermitian matrices. Our norm $\|M\|$ on matrices is the operator norm $\|M\| = \sup_{\|u\|_2=1} \|Mu\|_2$. We define

$$\|f\|_{\mathrm{Lip}} = \sup_{x \neq y} \frac{|f(x) - f(y)|}{|x - y|}$$

for test functions $f : \mathbb{R} \to \mathbb{R}$, and our metric d on probability measures will be the Dudley distance (also called the bounded-Lipschitz distance), given by

$$d(\mu, \nu) = \sup\left\{ \left| \int f \, \mathrm{d}(\mu - \nu) \right| : \|f\|_{\mathrm{Lip}} + \|f\|_{L^\infty} \leqslant 1 \right\}. \tag{5.1.5}$$

Recall that this distance metrizes weak convergence.

Finally, we recall the Stieltjes transform and the Voiculescu R-transform of a compactly supported probability measure. If μ is a probability measure on $\mathbb{R}$ the convex hull of whose support is

$[a, b]$, then we will normalize its Stieltjes transform G_μ as

$$G_\mu(y) = \int \frac{\mu(\mathrm{d}t)}{y - t}.$$

If we write $G_\mu(a) = \lim_{y \uparrow a} G_\mu(y)$ and $G_\mu(b) = \lim_{y \downarrow b} G_\mu(y)$, then it can be shown that G_μ is a bijection from $\mathbb{R} \setminus [a, b]$ to $(G_\mu(a), G_\mu(b)) \setminus \{0\}$. We will write

$$K_\mu : (G_\mu(a), G_\mu(b)) \setminus \{0\} \to \mathbb{R} \setminus [a, b]$$

for its functional inverse, and write

$$R_\mu(y) = K_\mu(y) - \frac{1}{y}$$

for its Voiculescu R-transform, which linearizes free convolution: $R_{\mu \boxplus \nu} = R_\mu + R_\nu$.

5.2 Main result

5.2.1 Assumptions. We first present our assumptions on D_N, which will be made throughout, even though we will only state them in the presentation of the main results.

Assumption 1. *The matrix D_N is real, diagonal, and deterministic, and its empirical measure $\hat{\mu}_{D_N}$ tends weakly as $N \to \infty$ to a compactly supported probability measure μ_D. Furthermore,*

$$\lambda_N(D_N) \to \mathtt{r}(\mu_D),$$
$$\lambda_1(D_N) \to \mathtt{l}(\mu_D).$$

Assumption 2. *There exist $C > 0$ and $\varepsilon_0 > 0$ such that*

$$d(\hat{\mu}_{D_N}, \mu_D) \leqslant CN^{-\varepsilon_0}.$$

Remark 5.2.1. *We emphasize that μ_D is allowed to be quite poorly behaved. For example, it can be singular with respect to Lebesgue measure. It can also have disconnected support. Notice that Assumption 2 is fairly mild. For example, if μ_D has a density and the entries of D_N are the $\frac{1}{N}$-quantiles of μ_D, then in fact $d(\hat{\mu}_{D_N}, \mu_D) = O(\frac{1}{N})$. If the entries of D_N were obtained from i.i.d. random variables, we would have $d(\hat{\mu}_{D_N}, \mu_D) = O(\frac{1}{\sqrt{N}})$.*

In fact, the proof of Lemma 5.5.6 below shows that, instead of Assumption 2, it suffices to bound the difference between the Stieltjes transforms of $\hat{\mu}_{D_N}$ and μ_D at distance $N^{-\delta}$ from the real line, for $\delta > 0$ small enough.

We will write the Laplace transform of a measure μ on $\mathbb{C}$ as

$$T_\mu(t) := \int e^{\Re(z\bar{t})}\mu(\mathrm{d}z).$$

If in fact μ is supported on $\mathbb{R}$ and t is real, this reduces to the familiar

$$T_\mu(t) = \int e^{tx}\mu(\mathrm{d}x).$$

We assume that $\frac{W_N}{\sqrt{N}}$ is a Wigner matrix, by which we mean that its entries are independent up to the self-adjoint condition. Our assumptions on the Wigner part are named, rather than numbered, to emphasize that our results apply under either of them, rather than both of them.

Gaussian Hypothesis. The matrix $\frac{W_N}{\sqrt{N}}$ is distributed according to the Gaussian Orthogonal Ensemble if $\beta = 1$, or the Gaussian Unitary Ensemble if $\beta = 2$. (That is, the law of W_N on the space of symmetric/Hermitian matrices has density proportional to $\exp(-\beta \operatorname{tr}(W_N^2)/4)$.)

SSGC Hypothesis. (This labelling stands for "sharp sub-Gaussian and concentrates." It matches the assumptions of [99].)

Write $\mu_{i,j}^N$ for the law of the (i,j)th entry of W_N.

1. Assume **both** of the following.

   – The first and second moments match those of the relevant Gaussian ensemble. In our normalization, this means that for every $N \in \mathbb{N}$ and $i,j \in [\![1,N]\!]$, if $\beta = 1$ we have

   $$\int x\mu_{i,j}^N(\mathrm{d}x) = 0, \quad \int x^2\mu_{i,j}^N(\mathrm{d}x) = 1 + \delta_{ij},$$

   whereas if $\beta = 2$ and $i \neq j$ we have

   $$\int \Re(z)\mu_{i,j}^N(\mathrm{d}z) = \int \Im(z)\mu_{i,j}^N(\mathrm{d}z) = \int \Re(z)\Im(z)\mu_{i,j}^N(\mathrm{d}z) = 0,$$
   $$\int \Re(z)^2\mu_{i,j}^N(\mathrm{d}z) = \int \Im(z)^2\mu_{i,j}^N(\mathrm{d}z) = \frac{1}{2}.$$

   If $\beta = 2$, then $\mu_{i,i}^N$ is supported on $\mathbb{R}$, with $\int x\mu_{i,i}^N(\mathrm{d}x) = 0$ and $\int x^2\mu_{i,i}^N(\mathrm{d}x) = 1$.

   – For every $N \in \mathbb{N}$ and $i,j \in [\![1,N]\!]$, the measure $\mu_{i,j}^N$ has a *sharp sub-Gaussian Laplace transform*:

   $$\text{for all } \begin{cases} t \in \mathbb{R} \text{ if } \beta = 1 \\ t \in \mathbb{C} \text{ if } \beta = 2 \end{cases}, \quad T_{\mu_{i,j}^N}(t) \leqslant \exp\left(\frac{|t|^2(1+\delta_{ij})}{2\beta}\right). \tag{5.2.1}$$

2. In addition, assume **one** of the following concentration-type hypotheses.

   – There exists a constant c independent of N such that, for all $N \in \mathbb{N}$ and all $i,j \in [\![1,N]\!]$, the law $\mu_{i,j}^N$ satisfies a log-Sobolev inequality with constant c.

   – There exists a compact set K independent of N (real if $\beta = 1$, or complex if $\beta = 2$) such that, for all $N \in \mathbb{N}$ and all $i,j \in [\![1,N]\!]$, the law $\mu_{i,j}^N$ is supported in K.

Remark 5.2.2. *A list of examples satisfying the SSGC Hypothesis is provided in [99]. Among these examples are real matrices whose entries follow the Rademacher law $\frac{1}{2}(\delta_{-1} + \delta_{+1})$ or the uniform law on $[-\sqrt{3},\sqrt{3}]$ (appropriately rescaled on the diagonal).*

In the literature, it is common to call a centered measure μ on $\mathbb{R}$ with unit variance sub-Gaussian *whenever*

$$A := 2 \sup_{t \in \mathbb{R}} \frac{1}{t^2} \log T_\mu(t)$$

is finite. We emphasize that we are asking for more: in (5.2.1) we require $A = 1$ (off the diagonal, with appropriate modifications otherwise), and following [99] we call such measures sharp sub-Gaussian. *This is a strict subclass; for example, the law of $\frac{1}{p}BG$, where $B \sim$ Bernoulli(p) and $G \sim \mathcal{N}(0,1)$ are independent, has unit variance but $A = 1/p$. This example appears in [15], which treats the general case $A > 1$, with zero deformation.*

5.2.2 Main result

Definition 5.2.3. *For a compactly supported measure ν, a parameter $\theta \geqslant 0$, and a real number $\mathscr{M} \geqslant \mathtt{r}(\nu)$, define*

$$J^{(\beta)}(\nu, \theta, \mathscr{M}) = \begin{cases} \frac{\beta}{2} \int_0^{\frac{2}{\beta}\theta} R_\nu(t)\, \mathrm{d}t & \text{if } 0 \leqslant \frac{2}{\beta}\theta \leqslant G_\nu(\mathscr{M}), \\ \theta\mathscr{M} - \frac{\beta}{2}\Big[1 + \log\Big(\frac{2}{\beta}\theta\Big)\Big] - \frac{\beta}{2}\int \log(\mathscr{M} - y)\nu(\mathrm{d}y) & \text{if } \frac{2}{\beta}\theta \geqslant G_\nu(\mathscr{M}). \end{cases} \tag{5.2.2}$$

(If $\mathscr{M} = \mathtt{r}(\nu)$, we recall our convention $G_\nu(\mathtt{r}(\nu)) = \lim_{y \downarrow \mathtt{r}(\nu)} G_\nu(y)$, which is possibly infinite.) In Section 5.3.1 we will explain how this function arises as the limit of appropriately normalized spherical integrals.

For $x \geqslant \mathtt{r}(\rho_{sc} \boxplus \mu_D)$ and $\theta \geqslant 0$, we define

$$I^{(\beta)}(x, \theta) = J^{(\beta)}(\rho_{sc} \boxplus \mu_D, \theta, x) - \frac{\theta^2}{\beta} - J^{(\beta)}(\mu_D, \theta, \mathtt{r}(\mu_D))$$

and then set

$$I^{(\beta)}(x) = \begin{cases} +\infty & \text{if } x < \mathtt{r}(\rho_{sc} \boxplus \mu_D), \\ \sup_{\theta \geqslant 0} I^{(\beta)}(x, \theta) & \text{if } x \geqslant \mathtt{r}(\rho_{sc} \boxplus \mu_D). \end{cases}$$

We will show below that

$$I^{(2)}(x) = 2I^{(1)}(x)$$

for all measures μ_D.

To state our result, we will need the following critical threshold.

Definition 5.2.4. *Given the compactly supported measure μ_D, define the real number x_c by*

$$x_c = x_c(\mu_D) = \begin{cases} \mathtt{r}(\mu_D) + G_{\mu_D}(\mathtt{r}(\mu_D)) & \text{if } G_{\mu_D}(\mathtt{r}(\mu_D)) < +\infty, \\ +\infty & \text{otherwise.} \end{cases}$$

It will be shown in Proposition 5.6.1 below that $x_c \geqslant \mathtt{r}(\rho_{sc} \boxplus \mu_D)$, with equality if and only if an inequality involving the Stieltjes transform of μ_D degenerates.

The main result of the paper is the following:

Theorem 5.2.5. *Suppose that Assumptions 1 and 2 hold.*

1. *If the Gaussian Hypothesis holds, then the law of the largest eigenvalue $\lambda_N(X_N)$ satisfies a large deviation principle at speed N with the good rate function $I^{(\beta)}(x)$. By rotational invariance, we have the same result when D_N is not diagonal but simply symmetric (if $\beta = 1$) or Hermitian (if $\beta = 2$) and satisfies the rest of the requirements of Assumption 1.*

2. *If instead the SSGC Hypothesis holds, then the law of the largest eigenvalue $\lambda_N(X_N)$ satisfies what we will call a "restricted large deviation principle on $(-\infty, x_c)$" at speed N with the good rate function $I^{(\beta)}(x)$. In fact the restriction is just for the lower bound; the upper bound is unrestricted. This means the following:*

   – *For every closed set $F \subset \mathbb{R}$, we have*

$$\limsup_{N\to\infty} \frac{1}{N} \log \mathbb{P}_N(\lambda_N(X_N) \in F) \leqslant -\inf_{x\in F} I^{(\beta)}(x). \tag{5.2.3}$$

– *For every open set $G \subset (-\infty, x_c)$, we have*

$$\liminf_{N\to\infty} \frac{1}{N} \log \mathbb{P}_N(\lambda_N(X_N) \in G) \geqslant - \inf_{x\in G} I^{(\beta)}(x). \tag{5.2.4}$$

3. *In particular, if the SSGC Hypothesis holds and μ_D is such that $x_c = +\infty$, then the law of the largest eigenvalue $\lambda_N(X_N)$ satisfies a large deviation principle at speed N with the good rate function $I^{(\beta)}(x)$ in the usual sense.*

Remark 5.2.6. *See Proposition 5.6.1 below for a more in-depth study of the function $I^{(\beta)}(x)$. There, it is shown that $I^{(\beta)}(x)$ has a unique minimizer at $x = \mathtt{r}(\rho_{sc} \boxplus \mu_D)$, where it takes the value zero. In particular, if the Gaussian Hypothesis holds, or if the SSGC Hypothesis holds and μ_D is such that $x_c = +\infty$, then*

$$\lambda_N(X_N) \to \mathtt{r}(\rho_{sc} \boxplus \mu_D) \quad \textit{almost surely.} \tag{5.2.5}$$

This result appears to be new in the real case when $\rho_{sc} \boxplus \mu_D$ is multicut, and in the complex non-Gaussian case when $\rho_{sc} \boxplus \mu_D$ is multicut and $(D_N)_{N=1}^{\infty}$ has "internal outliers" between the connected components of $\operatorname{supp}(\mu_D)$ that persist as $N \to \infty$. (Recall that we forbid "external outliers" by assuming $\lambda_N(D_N) \to \mathtt{r}(\mu_D)$ and $\lambda_1(D_N) \to \mathtt{l}(\mu_D)$.) In the literature Equation (5.2.5) appears as an easy corollary of edge universality results, or as a special case of BBP results when the deforming matrix D_N has no external outliers. For example, it follows from [65] for deformed GUE, possibly multicut with internal outliers, under some assumptions about the decay rate of μ_D near its edges; from [115] for general real or complex noise if μ_D is such that $\rho_{sc} \boxplus \mu_D$ is supported on a single interval with square-root decay at its two edges; and from [33] in the complex (and possibly multicut) case with no outliers. Of course, all of these papers achieve much more.

Remark 5.2.7. *The proof of the "restricted LDP," i.e., of Equations (5.2.3) and (5.2.4), follows in the classical way from estimates of small-ball probabilities via a weak large deviation principle and exponential tightness, except that we can only lower-bound small-ball probabilities*

$\mathbb{P}_N(|\lambda_N(X_N) - x| < \delta)$ *for* $x < x_c$ *rather than* $x \in \mathbb{R}$. *However, we can upper-bound these probabilities for all* x *(see Theorem 5.3.4); this is the reason for the different restrictions on* F *and* G *in* (5.2.3) *and* (5.2.4).

Remark 5.2.8. *Of course, one would prefer to write the rate function non-variationally, and we can do this when the argument is at or above the critical threshold* $x_c(\mu_D)$. *Proposition 5.6.1 shows that, for all* $x > \mathtt{r}(\rho_{sc} \boxplus \mu_D)$, *the supremum in the definition of* $I^{(\beta)}(x)$ *is achieved at a unique* $\theta_x^{(\beta)}$. *For* $x \geqslant x_c$ *(which is relevant for the Gaussian case), this* $\theta_x^{(\beta)}$ *is given explicitly as* $\theta_x^{(\beta)} = \frac{\beta}{2}(x - \mathtt{r}(\mu_D))$; *thus if* $x \geqslant x_c(\mu_D)$,

$$I^{(\beta)}(x) = \frac{\beta}{2}\left[\frac{(x - \mathtt{r}(\mu_D))^2}{2} - \int \log(x-y)(\rho_{sc} \boxplus \mu_D)(\mathrm{d}y) + \int \log(\mathtt{r}(\mu_D) - y)\mu_D(\mathrm{d}y)\right].$$

(If $x_c(\mu_D) < \infty$, *then* $\int \log(\mathtt{r}(\mu_D) - y)\mu_D(\mathrm{d}y) < \infty$.*) But for subcritical* x *values,* $\theta_x^{(\beta)}$ *is defined implicitly in the proof of Proposition 5.6.1 as the unique solution of the constrained problem*

$$\frac{2}{\beta}\theta_x^{(\beta)} + K_{\mu_D}\left(\frac{2}{\beta}\theta_x^{(\beta)}\right) = x \quad \textit{subject to} \quad \theta_x^{(\beta)} \in \left(\frac{\beta}{2}G_{\rho_{sc}\boxplus\mu_D}(\mathtt{r}(\rho_{sc} \boxplus \mu_D)), \frac{\beta}{2}G_{\mu_D}(\mathtt{r}(\mu_D))\right). \quad (5.2.6)$$

We have not found a way to solve this constrained problem explicitly, nor to write $I^{(\beta)}(x, \theta_x^{(\beta)})$ *explicitly at its solution. If the domain of* $\theta_x^{(\beta)}$ *in the constraint were instead* $(0, \frac{\beta}{2}G_{\rho_{sc}\boxplus\mu_D}(\mathtt{r}(\rho_{sc} \boxplus \mu_D)))$, *the equation would simplify to* $K_{\rho_{sc}\boxplus\mu_D}(\frac{2}{\beta}\theta_x^{(\beta)}) = x$, *which has the solution* $\theta_x^{(\beta)} = \frac{\beta}{2}G_{\rho_{sc}\boxplus\mu_D}(x)$. *But* $K_{\rho_{sc}\boxplus\mu_D}(\cdot)$ *is not generally guaranteed to exist for arguments larger than* $G_{\rho_{sc}\boxplus\mu_D}(\mathtt{r}(\rho_{sc} \boxplus \mu_D))$, *and even when extendable it may not be globally invertible.*

Thus our rate function remains implicit for subcritical x *values. Nevertheless, in some simple cases the constrained problem can be solved explicitly; two examples are given below in Sections 5.2.3 and 5.2.4.*

Remark 5.2.9. *If $D_N = 0$, then $x_c = +\infty$,*

$$I^{(\beta)}(x) = \begin{cases} +\infty & \text{if } x < 2 \\ \sup_{\theta \geqslant 0}\left\{ J^{(\beta)}(\rho_{sc}, \theta, x) - \frac{\theta^2}{\beta} \right\} & \text{if } x \geqslant 2 \end{cases} = \begin{cases} +\infty & \text{if } x < 2 \\ \frac{\beta}{2} \int_2^x \sqrt{t^2 - 4}\, \mathrm{d}t & \text{if } x \geqslant 2, \end{cases}$$

and we recover [99, Theorems 1.5 and 1.6], which in particular includes the classical LDP for the Gaussian ensembles. (The last equality in the above display is true by [99, Section 4.1].) Notice that we get the same rate function if D_N is not identically zero but rather $\|D_N\| \to 0$ sufficiently quickly.

Remark 5.2.10. *One wants to recover large deviations for BBP-type problems, so it is tempting to conjecture that, if the largest eigenvalue of D_N tends not to $\mathtt{r}(\mu_D)$ but to some $\rho > \mathtt{r}(\mu_D)$, then an LDP should hold for $\lambda_N(X_N)$ at speed N with the good rate function*

$$\tilde{I}^{(\beta)}(x) = \begin{cases} +\infty & \text{if } x < \mathtt{r}(\rho_{sc} \boxplus \mu_D) \\ \sup_{\theta \geqslant 0}\{ J^{(\beta)}(\rho_{sc} \boxplus \mu_D, \theta, x) - \frac{\theta^2}{\beta} - J^{(\beta)}(\mu_D, \theta, \rho) \} & \text{otherwise.} \end{cases}$$

But, at least for certain simple situations, such a conjecture would be wrong. For example, suppose that $\frac{W_N}{\sqrt{N}}$ is distributed according to the GOE (if $\beta = 1$) or the GUE (if $\beta = 2$), that $\mu_D = \delta_0$ (so that $\rho_{sc} \boxplus \mu_D = \rho_{sc}$), and that D_N has $N - 1$ zero eigenvalues with one spike at, say, 2 for concreteness. Then it is known [118, Theorem 1.2] that $\lambda_N(X_N)$ satisfies an LDP at speed N with the good rate function

$$\hat{I}^{(\beta)}(x) = \begin{cases} +\infty & x < 2 \\ \frac{\beta}{4} \int_{\frac{5}{2}}^x \sqrt{z^2 - 4}\, \mathrm{d}z - \beta\left(x - \frac{5}{2}\right) + \frac{\beta}{8}\left[x^2 - \left(\frac{5}{2}\right)^2 \right] & x \geqslant 2. \end{cases}$$

(The published rate function has a typo; it is corrected in the v2 arXiv posting. We also normalize the semicircle law differently.) Notice that this vanishes uniquely at $x = \frac{5}{2}$, which lies outside

$\mathrm{supp}(\rho_{sc})$ – *this model is past the BBP phase transition. But in this situation we can compute*

$$\tilde{I}^{(\beta)}(x) = \begin{cases} +\infty & x < 2 \\ 0 & 2 \leqslant x \leqslant \frac{5}{2} \\ \hat{I}^{(\beta)}(x) & x \geqslant \frac{5}{2} \end{cases}$$

It is likely that our method could be extended as in [102] to models where $\lim_{N\to\infty} \lambda_N(D_N)$ *is a spike below the BBP threshold, i.e., such that still* $\lambda_N(X_N) \to \mathtt{r}(\rho_{sc} \boxplus \mu_D)$ *almost surely. But a new idea is needed beyond the BBP threshold.*

5.2.3 First example ($x_c < \infty$)*.* If

$$\mu_D(\mathrm{d}x) = \frac{1}{2\pi\sigma^2}\sqrt{4\sigma^2 - x^2}\mathbf{1}_{x\in[-2\sigma,2\sigma]}\,\mathrm{d}x$$

for some parameter $\sigma > 0$, then $\rho_{\mathrm{sc}} \boxplus \mu_D$ is again semicircular, scaled so its support lies in $[-2\sqrt{\sigma^2+1}, 2\sqrt{\sigma^2+1}]$. The constrained equation (5.2.6) can be solved explicitly, and writing $\mathtt{r} = \mathtt{r}(\rho_{\mathrm{sc}} \boxplus \mu_D) = 2\sqrt{\sigma^2+1}$ and $x_c = 2\sigma + \frac{1}{\sigma}$ we can calculate

$$I^{(\beta)}(x) = \begin{cases} +\infty & \text{if } x < \mathtt{r} \\ \beta\left[\frac{x\sqrt{x^2-4(1+\sigma^2)}}{4(1+\sigma^2)} + \log\left(\frac{2\sqrt{1+\sigma^2}}{x+\sqrt{x^2-4(1+\sigma^2)}}\right)\right] & \text{if } \mathtt{r} \leqslant x \leqslant x_c \\ \beta\left[\frac{(x-2\sigma)^2}{4} + \frac{x\sqrt{x^2-4(1+\sigma^2)}-x^2}{8(1+\sigma^2)} + \frac{1}{2}\log\left(\frac{2\sigma}{x+\sqrt{x^2-4(1+\sigma^2)}}\right) + \frac{1}{2}\right] & \text{if } x \geqslant x_c. \end{cases}$$

Notice that $I^{(\beta)}(x)$ is C^2 but no better at x_c, which is perhaps surprising. Figure 5.1 plots this function when $\beta = 1$ and $\sigma = 1$ (i.e., when μ_D is the usual semicircle law supported on $[-2, 2]$). Here $\mathtt{r}(\rho_{\mathrm{sc}} \boxplus \mu_D) = 2\sqrt{2} \approx 2.83$, $x_c = 3$, and $I^{(\beta)}(x_c) \approx 0.03 \cdot \beta$. Under the SSGC Hypothesis, we would be able to estimate, say, $\mathbb{P}_N(\lambda_N \in (2.9, 2.95))$ but not $\mathbb{P}_N(\lambda_N \in (2.9, 3.1))$.

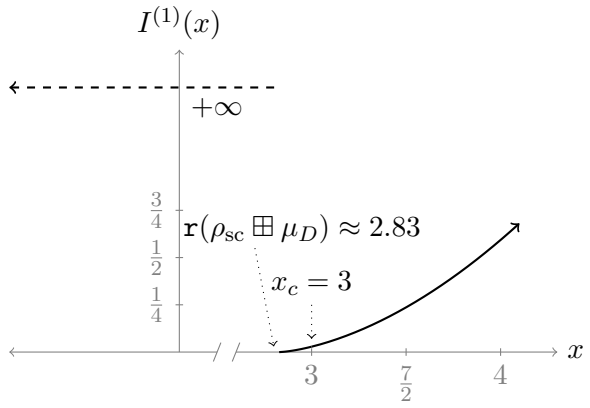


Figure 5.1: Sketch of the rate function when $\beta = 1$ and $\mu_D = \rho_{\mathrm{sc}}$.

5.2.4 Second example ($x_c = \infty$). Now suppose

$$\mu_D = \frac{1}{2}(\delta_{-a} + \delta_{+a})$$

for some parameter $a > 0$. Here $x_c(\mu_D) = a + G_{\mu_D}(a) = +\infty$, so all x are subcritical; that is, we can estimate any probability $\mathbb{P}_N(\lambda_N \in A)$ under either the SSGC Hypothesis or the Gaussian Hypothesis. Our computations use the known result

$$\mathtt{r}(\rho_{\mathrm{sc}} \boxplus \mu_D) = \frac{(4a^2 - 1 + \sqrt{8a^2+1})^{3/2}}{2\sqrt{2}a(\sqrt{8a^2+1} - 1)} =: \mathtt{r}(a). \tag{5.2.7}$$

In the physics literature this dates back to [157, Equations (55), (56)]; it was established in the mathematical literature in [51, Equations (3.5), (3.6)] (for $a > 1$), [8, Section 1] (for $a < 1$), and [52, Section 7] (for $a = 1$). The latter three papers establish that the measure $\rho_{\mathrm{sc}} \boxplus \mu_D$ undergoes a phase transition at $a = 1$. When $a > 1$, the support of $\rho_{\mathrm{sc}} \boxplus \mu_D$ consists of two intervals; when $a = 1$, these intervals meet at zero, where the density has cubic-root decay; and when $a < 1$ the support is a single interval, on the interior of which the density is strictly positive. (This set of three papers also establishes universality of correlation functions.) We emphasize that our results apply to all $a > 0$.

The details of the computations are given in Appendix C, but the result is this: With

$$\mathtt{c}(a) = \frac{(-1+4a^2+\sqrt{1+8a^2})^{3/2}}{3\sqrt{2}a(-1+\sqrt{1+8a^2})} - \frac{\sqrt{(1+8a^2)(-1-4a^2+8a^4+\sqrt{1+8a^2})}}{3\sqrt{2}a(-1+\sqrt{1+8a^2})}, \tag{5.2.8}$$

$$\mathtt{d}(y) = \mathtt{d}(a,y) = \frac{9y+18a^2y-2y^3}{\sqrt{-4(3-3a^2-y^2)^3-(9y+18a^2y-2y^3)^2}}, \tag{5.2.9}$$

we have

$$\begin{aligned} I^{(\beta)}(x) = \frac{\beta}{4}\Bigg[&\left[\frac{2}{3}\left[x-\sqrt{-3+3a^2+x^2}\sin\left(\frac{1}{3}\arctan(\mathtt{d}(x))\right)\right]\right]^2 \\ &+\log\left[\left(\frac{x}{3}+\left[\frac{2\sqrt{-3+3a^2+x^2}}{3}\sin\left(\frac{1}{3}\arctan(\mathtt{d}(x))\right)\right]\right)^2-a^2\right] \\ &-2\int_{\mathtt{r}(a)}^{x}\frac{2}{3}\left[t-\sqrt{-3+3a^2+t^2}\sin\left(\frac{\pi}{3}-\frac{1}{3}\arctan(\mathtt{d}(t))\right)\right]\mathrm{d}t \\ &-(\mathtt{c}(a)^2+\log((\mathtt{r}(a)-\mathtt{c}(a))^2-a^2))\Bigg] \end{aligned}$$

for $x > \mathtt{r}(\rho_{\mathrm{sc}} \boxplus \mu_D)$. Figure 5.2 plots this function at the critical parameter $a = 1$ (so that $\mathtt{r}(\rho_{\mathrm{sc}} \boxplus \mu_D) = \frac{3\sqrt{3}}{2}$) when $\beta = 1$.

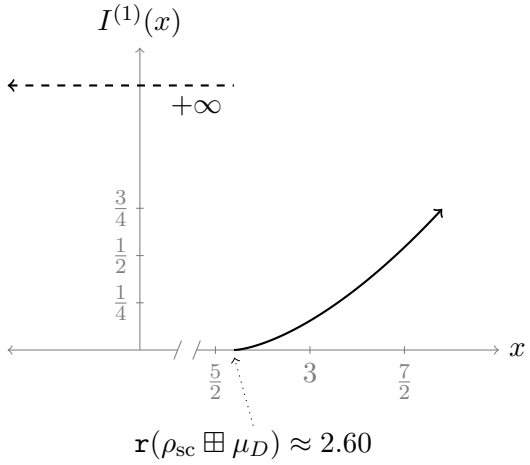


Figure 5.2: Sketch of the rate function when $\beta = 1$ and $\mu_D = \frac{1}{2}(\delta_1 + \delta_{-1})$.

Question 5.2.11. *Does the mechanism driving the deviations $\{\lambda_N(X_N) \approx x\}$ change as x passes*

the critical threshold x_c? Specifically, can one formalize and prove the notion that, with large probability, while the eigenvector corresponding to λ_N is delocalized under the above event for subcritical x values, it localizes for supercritical x values?

5.3 Proof overview

5.3.1 Spherical integrals. Given a self-adjoint $N \times N$ matrix X and $\theta \geqslant 0$, consider

$$I_N^{(\beta)}(X, \theta) = \mathbb{E}_{e,\beta}[e^{N\theta\langle e, Xe\rangle}],$$
$$J_N^{(\beta)}(X, \theta) = \frac{1}{N} \log I_N^{(\beta)}(X, \theta).$$

Recall that $\mathbb{E}_{e,\beta}$ is integration over vectors e uniform on the unit sphere, understood as $\mathbb{S}^{N-1} \subset \mathbb{R}^N$ if $\beta = 1$ or $\mathbb{S}^{N-1} \subset \mathbb{C}^N$ if $\beta = 2$, so that $I_N^{(\beta)}(X, \theta)$ is real and nonnegative for both symmetry classes. We emphasize again that $\mathbb{E}_{e,\beta}$ only averages over the unit sphere, so if X is random then $I_N^{(\beta)}(X, \theta)$ and $J_N^{(\beta)}(X, \theta)$ are random variables.

If $\{X_N\}$ is such that $\hat{\mu}_{X_N}$ has a weak limit ν, then we might hope that $J_N^{(\beta)}(X_N, \theta)$ also has a limit depending on ν and θ. This is so; but the limit also depends on $\lambda_N(X_N)$ if θ is sufficiently large. This should not be surprising, since the integrand $e^{N\theta\langle e, Xe\rangle}$ is maximized near the eigenvector corresponding to $\lambda_N(X)$, especially for larger θ values. Indeed, we have the following result.

Proposition 5.3.1. *[101, Theorem 6] Suppose that the sequence $(A_N)_{N=1}^{\infty}$ of self-adjoint matrices is such that $\hat{\mu}_{A_N} \to \nu$ weakly for some compactly-supported measure ν, that $\lambda_1(A_N)$ has a finite limit, and that $\lambda_N(A_N) \to \mathscr{M}$ for some real number $\mathscr{M}$. (Notice that we are not assuming that $\mathscr{M}$ is the right edge of ν, but of course we must have $\mathscr{M} \geqslant \mathtt{r}(\nu)$.) If $\theta \geqslant 0$, then*

$$\lim_{N\to\infty} J_N^{(\beta)}(A_N, \theta) = J^{(\beta)}(\nu, \theta, \mathscr{M}),$$

where $J^{(\beta)}(\nu, \theta, \mathscr{M})$ *is as in* (5.2.2).

5.3.2 Tilted measures and weak large deviations. Our general strategy will be to show a weak large deviation principle, as well as exponential tightness. In the proof of the weak-large-deviations lower bound for our measure of interest, we will actually need a weak-large-deviations *upper bound* for the following family of measures.

Definition 5.3.2. *Given* $\theta \geqslant 0$*, we consider the "tilted" measure* $\mathbb{P}_N^\theta$ *on* $N \times N$ *matrices (symmetric if* $\beta = 1$*, or Hermitian if* $\beta = 2$*) whose density with respect to the law* $\mathbb{P}_N$ *of* X_N *is given by*

$$\frac{\mathrm{d}\mathbb{P}_N^\theta}{\mathrm{d}\mathbb{P}_N}(X) = \frac{I_N^{(\beta)}(X, \theta)}{\mathbb{E}_{X_N}(I_N^{(\beta)}(X_N, \theta))}.$$

Notice from the definition of $I_N^{(\beta)}$ *that* $\mathbb{P}_N^0 = \mathbb{P}_N$.

We will need the following asymptotics of the free energy for this measure, with proof in Section 5.4.

Proposition 5.3.3. *Given the compactly supported measure* μ_D*, define the threshold*

$$\theta_c^{(\beta)} = \theta_c^{(\beta)}(\mu_D) = \begin{cases} \frac{\beta}{2} G_{\mu_D}(\mathtt{r}(\mu_D)) & \text{if } G_{\mu_D}(\mathtt{r}(\mu_D)) < +\infty, \\ +\infty & \text{otherwise.} \end{cases} \tag{5.3.1}$$

Under the Gaussian Hypothesis, choose any $\theta \geqslant 0$*; or, under the SSGC Hypothesis, choose any* $0 \leqslant \theta < \theta_c^{(\beta)}$*. Then*

$$\lim_{N \to \infty} \frac{1}{N} \log \mathbb{E}_{X_N}[I_N^{(\beta)}(X_N, \theta)] = \frac{\theta^2}{\beta} + J^{(\beta)}(\mu_D, \theta, \mathtt{r}(\mu_D)).$$

The reason for the appearance of $\theta_c^{(\beta)}$ and x_c is this: Under the SSGC Hypothesis, we can only give lower bounds for $\mathbb{E}_{X_N}[e^{N\theta\langle e, X_N e\rangle}] = e^{N\theta\langle e, D_N e\rangle}\mathbb{E}_{W_N}[e^{\sqrt{N}\theta\langle e, W_N e\rangle}]$ when e is delocalized, since we only have lower bounds for the Laplace transforms of the entries of W_N near zero. Informally,

to understand the normalization constant in $\mathbb{P}_N^\theta$ we therefore need θ to be such that

$$\mathbb{E}_{e,\beta}[\mathbf{1}_{e \text{ delocalized}} e^{N\theta\langle e, D_N e\rangle}] \approx \mathbb{E}_{e,\beta}[e^{N\theta\langle e, D_N e\rangle}]$$

at exponential scale, which we can only prove for $\theta < \theta_c^{(\beta)}$ (in fact it probably fails for larger θ). To establish the weak LDP lower bound, we need to show that the event $\{\lambda_N(X_N) \approx x\}$ is likely under $\mathbb{P}_N^\theta$ for some $\theta = \theta_x$. Under the SSGC Hypothesis, this is possible only if x is such that $\theta_x < \theta_c^{(\beta)}$, and this turns out to be true if and only if $x < x_c$, where x_c is as in Definition 5.2.4.

We split up the weak-large-deviations upper and lower bounds as follows:

Theorem 5.3.4. *First, let $x < \mathtt{r}(\rho_{sc} \boxplus \mu_D)$. Under either Hypothesis, choose any $\theta \geqslant 0$. Then*

$$\lim_{\delta \to 0} \limsup_{N \to \infty} \frac{1}{N} \log \mathbb{P}_N^\theta(|\lambda_N(X_N) - x| \leqslant \delta) = -\infty.$$

Second, let $x \geqslant \mathtt{r}(\rho_{sc} \boxplus \mu_D)$. Under the Gaussian Hypothesis, choose any $\theta \geqslant 0$; or, under the SSGC Hypothesis, choose any $0 \leqslant \theta < \theta_c^{(\beta)}$. Then

$$\limsup_{\delta \to 0} \limsup_{N \to \infty} \frac{1}{N} \log \mathbb{P}_N^\theta(|\lambda_N(X_N) - x| \leqslant \delta) \leqslant -(I^{(\beta)}(x) - I^{(\beta)}(x, \theta)).$$

Notice that $I^{(\beta)}(x, 0) = 0$ for all measures μ_D and all $x \geqslant \mathtt{r}(\rho_{sc} \boxplus \mu_D)$. Thus when $\theta = 0$ we recover the weak large deviation upper bound for the measure of primary interest, under either Hypothesis.

Theorem 5.3.5. *Under the Gaussian Hypothesis, choose any $x \in \mathbb{R}$; or, under the SSGC Hypothesis, choose any $x < x_c$. Then*

$$\liminf_{\delta \to 0} \liminf_{N \to \infty} \frac{1}{N} \log \mathbb{P}_N(|\lambda_N(X_N) - x| < \delta) \geqslant -I^{(\beta)}(x).$$

5.3.3 Outline. When estimating $\frac{1}{N} \log \mathbb{P}_N(|\lambda_N(X_N) - x| \leqslant \delta)$ by tilting by spherical integrals, one wants to estimate $J_N^{(\beta)}(X_N, \theta)$ on the event $\{|\lambda_N(X_N) - x| \leqslant \delta\}$. To localize $J_N^{(\beta)}(X_N, \theta)$, one

needs to control $\hat{\mu}_{X_N}$. Therefore one wants to find a set

$$\mathcal{A}_{x,\delta}^M \subset \{|\lambda_N(X_N) - x| \leqslant \delta\}$$

of matrices with controlled empirical measures (which will turn out to depend on some $M \gg 1$) satisfying both of the following:

– On the one hand, $\mathcal{A}_{x,\delta}^M$ is a continuity set for spherical integrals, in the sense that we have a good enough understanding of $J_N^{(\beta)}(T,\theta)$ for $T \in \mathcal{A}_{x,\delta}^M$ to be able to estimate

$$\frac{1}{N}\log \mathbb{P}_N(\mathcal{A}_{x,\delta}^M) \approx e^{-NI^{(\beta)}(x)}.$$

– On the other hand, $\mathcal{A}_{x,\delta}^M$ is not too much smaller than $\{|\lambda_N(X_N) - x| \leqslant \delta\}$, in the sense that

$$\frac{1}{N}\log \mathbb{P}_N(|\lambda_N(X_N) - x| \leqslant \delta) \approx \frac{1}{N}\log \mathbb{P}_N(\mathcal{A}_{x,\delta}^M).$$

The next subsection first details the continuity result of [118], which helps us choose $\mathcal{A}_{x,\delta}^M$ while satisfying the first point, then states a proposition which we need to show that our choice satisfies the second point.

5.3.4 Continuity of spherical integrals

Proposition 5.3.6. *[118, Proposition 2.1] For any $\theta > 0$ and any $\kappa > 0$, there exists a function $g_{\kappa,\theta} : \mathbb{R}^+ \to \mathbb{R}^+$ going to zero at zero such that, for any $\delta > 0$ and N large enough, if B_N and B'_N are sequences of matrices such that $d(\hat{\mu}_{B_N}, \hat{\mu}_{B'_N}) < N^{-\kappa}$, $|\lambda_N(B_N) - \lambda_N(B'_N)| < \delta$, $\sup_N \|B_N\| < \infty$, and $\sup_N \|B'_N\| < \infty$, then we have*

$$\left|J_N^{(\beta)}(B_N,\theta) - J_N^{(\beta)}(B'_N,\theta)\right| < g_{\kappa,\theta}(\delta).$$

This suggests that we introduce the following deterministic sets of $N \times N$ symmetric matrices.

Fix once and for all a κ satisfying Proposition 5.3.9, below, and write g_θ for $g_{\frac{\kappa}{2},\theta}$; then for any $x \in \mathbb{R}$, $\delta > 0$, and $M > 0$, let

$$\mathcal{A}_{x,\delta}^M = \{X : |\lambda_N(X) - x| < \delta, d(\hat{\mu}_X, \rho_{\mathrm{sc}} \boxplus \mu_D) < N^{-\kappa}, \text{ and } \|X\| \leqslant M\}.$$

In the next few results, we discretize the measure $\rho_{\mathrm{sc}} \boxplus \mu_D$ so that we can apply Proposition 5.3.6 and control $J_N^{(\beta)}(X_N, \theta)$ uniformly for $X_N \in \mathcal{A}_{x,\delta}^M$.

Lemma 5.3.7. *Fix $x \geqslant \mathtt{r}(\rho_{sc} \boxplus \mu_D)$ and $M \geqslant \max(x, |\mathtt{l}(\rho_{sc} \boxplus \mu_D)|)$. Then there exists a sequence of deterministic matrices B'_N with the following properties:*

- *$\lambda_N(B'_N) = x$,*
- *$\sup_{N \geqslant 1} \|B'_N\| \leqslant M$, and*
- *$d(\hat{\mu}_{B'_N}, \rho_{sc} \boxplus \mu_D) = O(1/N)$.*

Proof. Given N, define the $\frac{1}{N}$ quantiles $\{\gamma_j\}_{j=1}^N = \{\gamma_j^{(N)}\}_{j=1}^N$ of the measure $\rho_{\mathrm{sc}} \boxplus \mu_D$ implicitly by

$$\frac{j}{N} = (\rho_{\mathrm{sc}} \boxplus \mu_D)((-\infty, \gamma_j)).$$

(This is possible since $\rho_{\mathrm{sc}} \boxplus \mu_D$ admits a density [49, Corollary 2].) Then let

$$B'_N = \mathrm{diag}(\gamma_1, \ldots, \gamma_{N-1}, x).$$

The distance estimate is easy to show, since $d(\cdot, \cdot)$ is defined with respect to bounded-Lipschitz test functions. □

Corollary 5.3.8. *For every $\theta \geqslant 0$, $x \geqslant \mathtt{r}(\rho_{sc} \boxplus \mu_D)$, $\delta > 0$, and $M > \max(x + \delta, |\mathtt{l}(\rho_{sc} \boxplus \mu_D)|)$, we have*

$$\limsup_{N \to \infty} \sup_{B_N \in \mathcal{A}_{x,\delta}^M} \left| J_N^{(\beta)}(B_N, \theta) - J^{(\beta)}(\rho_{sc} \boxplus \mu_D, \theta, x) \right| \leqslant g_\theta(\delta).$$

Proof. Let $\{B'_N\}_{N=1}^{\infty}$ be as in Lemma 5.3.7, so that $d(\hat{\mu}_{B'_N}, \rho_{\mathrm{sc}} \boxplus \mu_D) \leqslant N^{-\kappa}$ for N sufficiently large, with κ fixed as above. Then whenever $B_N \in \mathcal{A}_{x,\delta}^M$ we have $d(\hat{\mu}_{B_N}, \hat{\mu}_{B'_N}) \leqslant 2N^{-\kappa} \leqslant N^{-\frac{\kappa}{2}}$, and $|\lambda_N(B_N) - \lambda_N(B'_N)| \leqslant \delta$, so that by Proposition 5.3.6 and by our definition of g_θ

$$\sup_{B_N \in \mathcal{A}_{x,\delta}^M} \left| J_N^{(\beta)}(B_N, \theta) - J_N^{(\beta)}(B'_N, \theta) \right| \leqslant g_\theta(\delta)$$

for N sufficiently large. In addition, by Proposition 5.3.1 we have

$$\lim_{N\to\infty} \left| J_N^{(\beta)}(B'_N, \theta) - J^{(\beta)}(\rho_{\mathrm{sc}} \boxplus \mu_D, \theta, x) \right| = 0.$$

The result follows. □

On the other hand, the result below shows that the restrictions we added to $\{X : |\lambda_N(X) - x| < \delta\}$ to arrive at $\mathcal{A}_{x,\delta}^M$ have probability negligibly close to 1 at the exponential scale. Notice that the first point is exponential tightness. The proof will make up Section 5.5.

Proposition 5.3.9. *Assume either the Gaussian Hypothesis or the SSGC Hypothesis.*

1. *For every $\theta \geqslant 0$ we have*

$$\lim_{M\to\infty} \limsup_{N\to\infty} \frac{1}{N} \log \mathbb{P}_N^\theta(\|X_N\| > M) = -\infty.$$

2. *There exists $\gamma > 0$ such that, for any $0 < \kappa < \gamma$ and any $\theta \geqslant 0$,*

$$\lim_{N\to\infty} \frac{1}{N} \log \mathbb{P}_N^\theta(d(\hat{\mu}_{X_N}, \rho_{sc} \boxplus \mu_D) > N^{-\kappa}) = -\infty.$$

Theorem 5.2.5 follows in the classical way from the exponential tightness above, the weak LDP upper bound (Theorem 5.3.4), and the weak LDP lower bound (Theorem 5.3.5). We now prove the latter two.

5.3.5 *The proof of the weak LDP upper bound*

Lemma 5.3.10. *Fix $y \geqslant \mathtt{r}(\rho_{sc} \boxplus \mu_D)$ and $M > \max(y, |\mathtt{l}(\rho_{sc} \boxplus \mu_D)|)$. Under the Gaussian Hypothesis, choose any $\theta \geqslant 0$; or, under the SSGC Hypothesis, choose any $0 \leqslant \theta < \theta_c^{(\beta)}$. Then*

$$\limsup_{\delta \to 0} \limsup_{N \to \infty} \frac{1}{N} \log \mathbb{P}_N^{\theta}(\mathcal{A}_{y,\delta}^M) \leqslant -(I^{(\beta)}(y) - I^{(\beta)}(y, \theta)).$$

Proof. For any $\theta' \geqslant 0$, we have

$$\begin{aligned}
\mathbb{P}_N^{\theta}(\mathcal{A}_{y,\delta}^M) &= \frac{1}{\mathbb{E}_{X_N}[I_N^{(\beta)}(X_N, \theta)]} \mathbb{E}_{X_N}\left[\mathbf{1}_{X_N \in \mathcal{A}_{y,\delta}^M} I_N^{(\beta)}(X_N, \theta) \frac{I_N^{(\beta)}(X_N, \theta')}{I_N^{(\beta)}(X_N, \theta')}\right] \\
&\leqslant \frac{\mathbb{E}_{X_N}[I_N^{(\beta)}(X_N, \theta')]}{\mathbb{E}_{X_N}[I_N^{(\beta)}(X_N, \theta)]} \left(\sup_{X \in \mathcal{A}_{y,\delta}^M} I_N^{(\beta)}(X, \theta)\right) \left(\sup_{X \in \mathcal{A}_{y,\delta}^M} \frac{1}{I_N^{(\beta)}(X, \theta')}\right).
\end{aligned}$$

Fix $\varepsilon > 0$. By Corollary 5.3.8 and Lemmas 5.4.1 (applied to θ', which is any nonnegative number, hence the need for Lemma 5.4.1) and 5.4.2 (applied to θ, which is subcritical if necessary), if $M > y + \delta$ (true for small enough δ since $M > y$) and for N sufficiently large depending on θ, θ', and ε, we thus have

$$\frac{1}{N} \log \mathbb{P}_N^{\theta}(\mathcal{A}_{y,\delta}^M) \leqslant I^{(\beta)}(y, \theta) - I^{(\beta)}(y, \theta') + 2g_{\theta}(\delta) + 2g_{\theta'}(\delta) + \varepsilon.$$

By taking $N \to \infty$, then $\delta \downarrow 0$, then $\varepsilon \downarrow 0$, we obtain

$$\limsup_{\delta \downarrow 0} \limsup_{N \to \infty} \frac{1}{N} \log \mathbb{P}_N^{\theta}(\mathcal{A}_{y,\delta}^M) \leqslant -(I^{(\beta)}(y, \theta') - I^{(\beta)}(y, \theta))$$

which gives us the result by optimizing over θ'. □

Proof of Theorem 5.3.4. We first focus on the case when $x < \mathtt{r}(\rho_{\mathrm{sc}} \boxplus \mu_D)$. For such an x, if δ is so small that $x + \delta < \mathtt{r}(\rho_{\mathrm{sc}} \boxplus \mu_D) - \delta$, then whenever $|\lambda_N(X_N) - x| \leqslant \delta$, the empirical spectral measure $\hat{\mu}_{X_N}$ does not charge $(\mathtt{r}(\rho_{\mathrm{sc}} \boxplus \mu_D) - \delta, \mathtt{r}(\rho_{\mathrm{sc}} \boxplus \mu_D))$. Hence $d(\hat{\mu}_{X_N}, \rho_{\mathrm{sc}} \boxplus \mu_D) \geqslant f(\delta)$ for

some positive function f. Thus for such δ and for N large enough we have

$$\frac{1}{N}\log \mathbb{P}_N^\theta(|\lambda_N(X_N)-x|\leqslant\delta)\leqslant\frac{1}{N}\log\mathbb{P}_N^\theta(d(\hat{\mu}_{X_N},\rho_{\mathrm{sc}}\boxplus\mu_D)>N^{-\kappa})$$

which suffices in light of Proposition 5.3.9. Thus in the following it remains only to consider $x\geqslant \mathtt{r}(\rho_{\mathrm{sc}}\boxplus\mu_D)$.

Fix $\theta\geqslant 0$, $\delta>0$, $x>\mathtt{r}(\rho_{\mathrm{sc}}\boxplus\mu_D)$, and a sufficiently large M. Then we have

$$\mathbb{P}_N^\theta(\lambda_N\in[x-\delta,x+\delta])\leqslant\mathbb{P}_N^\theta(\mathcal{A}_{x,2\delta}^M)+\mathbb{P}_N^\theta(d(\hat{\mu}_{X_N},\rho_{\mathrm{sc}}\boxplus\mu_D)>N^{-\kappa})+\mathbb{P}_N^\theta(\|X_N\|>M).$$

An application of Proposition 5.3.9 gives us

$$\begin{aligned}&\limsup_{N\to\infty}\frac{1}{N}\log\mathbb{P}_N^\theta(\lambda_N\in[x-\delta,x+\delta])\\&\leqslant\max\Big(\limsup_{N\to\infty}\frac{1}{N}\log\mathbb{P}_N^\theta(\mathcal{A}_{x,2\delta}^M),\limsup_{N\to\infty}\frac{1}{N}\log\mathbb{P}_N^\theta(\|X_N\|>M)\Big).\end{aligned}$$

By taking $\delta\downarrow 0$ and applying Lemma 5.3.10, we obtain

$$\begin{aligned}&\limsup_{\delta\downarrow0}\limsup_{N\to\infty}\frac{1}{N}\log\mathbb{P}_N^\theta(\lambda_N\in[x-\delta,x+\delta])\\&\leqslant\max\Big(-(I^{(\beta)}(x)-I^{(\beta)}(x,\theta)),\limsup_{N\to\infty}\frac{1}{N}\log\mathbb{P}_N^\theta(\|X_N\|>M)\Big).\end{aligned}$$

Finally we obtain the result by taking $M\to\infty$ and applying again Proposition 5.3.9. □

5.3.6 The proof of the weak LDP lower bound. The following lemma relies on results about the rate function which will be established in Section 5.6.

Lemma 5.3.11. *Under the Gaussian Hypothesis, choose any $x\geqslant\mathtt{r}(\rho_{sc}\boxplus\mu_D)$; or, under the SSGC Hypothesis, choose any $\mathtt{r}(\rho_{sc}\boxplus\mu_D)\leqslant x<x_c$. Then there exists $\theta_x^{(\beta)}>0$ such that, for any M*

sufficiently large depending on x and any $\delta > 0$ sufficiently small depending on x, we have

$$\lim_{N\to\infty} \frac{1}{N} \log \mathbb{P}_N^{\theta_x^{(\beta)}}(\mathcal{A}_{x,\delta}^M) = 0.$$

If $x < x_c$, then $\theta_x^{(\beta)} < \theta_c^{(\beta)}$.

Proof. Fix $x \geqslant \mathtt{r}(\rho_{\mathrm{sc}} \boxplus \mu_D)$, and let $\theta_x^{(\beta)}$ be such that

$$I^{(\beta)}(x) = \sup_{\theta \geqslant 0} I^{(\beta)}(x, \theta) = I^{(\beta)}(x, \theta_x^{(\beta)}).$$

Proposition 5.6.1 below shows that this exists and is unique (except at $x = \mathtt{r}(\rho_{\mathrm{sc}} \boxplus \mu_D)$, where we choose one of many possible $\theta_x^{(\beta)}$ values by convention), and that $\theta_x^{(\beta)} < \theta_c^{(\beta)}$ whenever $x < x_c$. We claim that in fact $\mathbb{P}_N^{\theta_x^{(\beta)}}(\mathcal{A}_{x,\delta}^M) = 1 - o(1)$; to prove this, by Proposition 5.3.9 it suffices to show

$$\limsup_{N\to\infty} \frac{1}{N} \log \mathbb{P}_N^{\theta_x^{(\beta)}}(\lambda_N \notin [x - \delta, x + \delta]) < 0$$

for δ small enough. Since $\{\lambda_N < \mathtt{r}(\rho_{\mathrm{sc}} \boxplus \mu_D) - 1\} \subset \{d(\hat{\mu}_{X_N}, \rho_{\mathrm{sc}} \boxplus \mu_D) > \varepsilon\}$ for some ε, and since the law of λ_N is exponentially tight under $\mathbb{P}_N^{\theta_x}$, we need only show that for K large enough

$$\limsup_{N\to\infty} \frac{1}{N} \log \mathbb{P}_N^{\theta_x^{(\beta)}}(\lambda_N \in [\mathtt{r}(\rho_{\mathrm{sc}} \boxplus \mu_D) - 1, x - \delta] \cup [x + \delta, K]) < 0.$$

But Theorem 5.3.4 shows a weak large deviation upper bound for $\mathbb{P}_N^{\theta_x^{(\beta)}}$ with the rate function $J_x^{(\beta)}(y) = I^{(\beta)}(y) - I^{(\beta)}(y, \theta_x^{(\beta)})$, which Proposition 5.6.1 below shows is nonnegative and vanishes uniquely at $y = x$. (This theorem applies, since $\theta_x^{(\beta)}$ is less than $\theta_c^{(\beta)}$ if necessary.) Since $[\mathtt{r}(\rho_{\mathrm{sc}} \boxplus \mu_D) - 1, x - \delta] \cup [x + \delta, K]$ is a compact set that does not contain x, this suffices. □

Proof of Theorem 5.3.5. If $x < \mathtt{r}(\rho_{\mathrm{sc}} \boxplus \mu_D)$, then $I^{(\beta)}(x) = +\infty$, and there is nothing to prove. Thus we will assume in the following that $x \geqslant \mathtt{r}(\rho_{\mathrm{sc}} \boxplus \mu_D)$.

Whenever $X \in \mathcal{A}_{x,\delta}^M$, by Corollary 5.3.8 we have

$$J_N^{(\beta)}(X, \theta_x^{(\beta)}) \leqslant 2g_{\theta_x^{(\beta)}}(\delta) + J^{(\beta)}(\rho_{\mathrm{sc}} \boxplus \mu_D, \theta_x^{(\beta)}, x)$$

for N sufficiently large. In addition, for every $\varepsilon > 0$, Lemma 5.3.11 tells us that for N sufficiently large depending on ε we have $\mathbb{P}_N^{\theta_x^{(\beta)}}(\mathcal{A}_{x,\delta}^M) \geqslant e^{-N\varepsilon}$.

We wish to use Proposition 5.3.3 to conclude that, for N sufficiently large depending on ε and on $\theta_x^{(\beta)}$, we also have

$$\mathbb{E}_{X_N}[I_N(X_N, \theta_x^{(\beta)})] \geqslant e^{N(\frac{\theta^2}{\beta} + J^{(\beta)}(\mu_D, \theta_x^{(\beta)}, \mathbf{r}(\mu_D)) - \varepsilon)}.$$

Under the Gaussian Hypothesis, this is permissible for every x; under the SSGC Hypothesis, our restriction $x < x_c$ tells us by Lemma 5.3.11 that $\theta_x^{(\beta)} < \theta_c^{(\beta)}$, so that Proposition 5.3.3 indeed applies.

Thus

$$\begin{aligned}
\mathbb{P}_N(\mathcal{A}_{x,\delta}^M) &\geqslant \frac{\mathbb{E}_{X_N}[\mathbf{1}_{X_N \in \mathcal{A}_{x,\delta}^M} I_N^{(\beta)}(X_N, \theta_x^{(\beta)})]}{\mathbb{E}_{X_N}[I_N^{(\beta)}(X_N, \theta_x^{(\beta)})]} \mathbb{E}_{X_N}[I_N^{(\beta)}(X_N, \theta_x^{(\beta)})] e^{-N \sup_{X \in \mathcal{A}_{x,\delta}^M} J_N^{(\beta)}(X, \theta_x^{(\beta)})} \\
&\geqslant \mathbb{P}_N^{\theta_x^{(\beta)}}(\mathcal{A}_{x,\delta}^M) e^{N(\frac{(\theta_x^{(\beta)})^2}{\beta} + J^{(\beta)}(\mu_D, \theta_x^{(\beta)}, \mathbf{r}(\mu_D)) - \varepsilon)} e^{-N \sup_{X \in \mathcal{A}_{x,\delta}^M} J_N^{(\beta)}(X, \theta_x^{(\beta)})} \\
&\geqslant e^{-N\varepsilon} e^{N((\frac{\theta_x^{(\beta)})^2}{\beta} + J^{(\beta)}(\mu_D, \theta_x^{(\beta)}, \mathbf{r}(\mu_D)) - \varepsilon)} e^{-N(J^{(\beta)}(\rho_{\mathrm{sc}} \boxplus \mu_D, \theta_x^{(\beta)}, x) + 2g_{\theta_x^{(\beta)}}(\delta))} \\
&= e^{-N(I^{(\beta)}(x) + 2\varepsilon + 2g_{\theta_x^{(\beta)}}(\delta))}.
\end{aligned}$$

Thus, fixing some M sufficiently large, we obtain

$$\liminf_{N \to \infty} \frac{1}{N} \log \mathbb{P}_N(|\lambda_N(X_N) - x| < \delta) \geqslant \liminf_{N \to \infty} \frac{1}{N} \log \mathbb{P}_N(\mathcal{A}_{x,\delta}^M) \geqslant -(I^{(\beta)}(x) + 2\varepsilon + 2g_{\theta_x^{(\beta)}}(\delta))$$

and since this is true for every $\varepsilon > 0$ we can take the limit as $\delta \downarrow 0$ to conclude. □

5.4 Free energy expansion

In this section we prove Proposition 5.3.3, recalling that we state the results for $\beta \in \{1, 2\}$ but give proofs only for $\beta = 1$. (In particular, in the proofs we drop β from notations, writing $I_N(\cdot, \cdot)$ for $I_N^{(\beta)}(\cdot, \cdot)$ and so on.)

Proof under the Gaussian Hypothesis. For the remainder of this paper, we introduce the notation

$$D_N = \operatorname{diag}(d_1, \dots, d_N) = \operatorname{diag}(d_1^{(N)}, \dots, d_N^{(N)}).$$

For every unit vector e, we have

$$\begin{aligned} \mathbb{E}_{X_N}[e^{N\theta\langle e, X_N e\rangle}] &= \left[\prod_{i<j} T_{\mu_{i,j}^N}(2\sqrt{N}\theta e_i e_j)\right]\left[\prod_{i=1}^N T_{\mu_{i,i}^N}(\sqrt{N}\theta e_i^2) e^{N\theta d_i e_i^2}\right] \\ &= \left[\prod_{i<j} \exp\left(2N\theta^2 e_i^2 e_j^2\right)\right]\left[\prod_{i=1}^N \exp\left(N\theta^2 e_i^4 + N\theta d_i e_i^2\right)\right] \\ &= \exp\left(N\theta^2\right)\exp(N\theta\langle e, D_N e\rangle). \end{aligned} \tag{5.4.1}$$

Integrating over $\mathbb{S}^{N-1}$, we find

$$\mathbb{E}_{X_N}[I_N(X_N, \theta)] = e^{N\theta^2} I_N(D_N, \theta),$$

so Proposition 5.3.3 follows from Proposition 5.3.1. □

The proof under the SSGC Hypothesis is more involved and will take up the remainder of this section. We separate the upper and lower bounds as follows.

Lemma 5.4.1. *Under the SSGC Hypothesis, for any $\theta \geqslant 0$ and any N we have*

$$\mathbb{E}_{X_N}[I_N^{(\beta)}(X_N, \theta)] \leqslant e^{N\frac{\theta^2}{\beta}} I_N^{(\beta)}(D_N, \theta).$$

In particular, by Proposition 5.3.1, for any $\theta \geqslant 0$ we have

$$\limsup_{N\to\infty} \frac{1}{N} \log \mathbb{E}_{X_N}[I_N^{(\beta)}(X_N, \theta)] \leqslant \frac{\theta^2}{\beta} + J^{(\beta)}(\mu_D, \theta, \mathtt{r}(\mu_D)).$$

Lemma 5.4.2. *Under the SSGC Hypothesis, for any $0 \leqslant \theta < \theta_c^{(\beta)}$ we have*

$$\liminf_{N\to\infty} \frac{1}{N} \log \mathbb{E}_{X_N}[I_N^{(\beta)}(X_N, \theta)] \geqslant \frac{\theta^2}{\beta} + J^{(\beta)}(\mu_D, \theta, \mathtt{r}(\mu_D)).$$

The proof of the lower bound will use the following two technical results.

Lemma 5.4.3. *Under the SSGC Hypothesis, for every $\delta > 0$ there exists $\varepsilon(\delta) > 0$ such that, for every $N \in \mathbb{N}$, every $i, j \in [\![1, N]\!]$, and every $t \in \mathbb{R}$ with $|t| \leqslant \varepsilon(\delta)$ if $\beta = 1$ (or every $t \in \mathbb{C}$ with $|t| \leqslant \varepsilon(\delta)$ if $\beta = 2$),*

$$T_{\mu_{i,j}^N}(t) \geqslant \exp\left((1-\delta)\frac{|t|^2(1+\delta_{ij})}{2\beta}\right).$$

Lemma 5.4.4. *For any $0 \leqslant \theta < \theta_c$ we have*

$$\lim_{N\to\infty} \frac{1}{N} \log \left(\frac{\mathbb{E}_{e,\beta}\left[\mathbf{1}_{\|e\|_\infty \leqslant N^{-\frac{3}{8}}} e^{N\theta\langle e, D_N e\rangle}\right]}{I_N^{(\beta)}(D_N, \theta)} \right) = 0. \tag{5.4.2}$$

Proof of Lemma 5.4.1. This is the same as the proof under the Gaussian Hypothesis, except that the second equality in (5.4.1) is replaced by an upper bound, due to the upper-bound assumption (5.2.1) on Laplace transforms. □

Proof of Lemma 5.4.2. Fix $\delta > 0$, and let $\varepsilon = \varepsilon(\delta)$ be as in Lemma 5.4.3, proved below. Whenever the unit vector e is such that $\|e\|_\infty \leqslant N^{-3/8}$, we have

$$\max_{i,j} \left|2\sqrt{N}\theta e_i e_j\right| \leqslant \varepsilon(\delta), \qquad \max_i \left|\sqrt{N}\theta e_i^2\right| \leqslant \varepsilon(\delta)$$

for $N \geqslant N_0(\delta)$. (The proof below will work with any exponent strictly between $-1/2$ and $-1/4$; but since the exponent does not appear in the final result, we have chosen $-3/8$ for definiteness.)

Thus the lower bound on the Laplace transform of Lemma 5.4.3 gives us, for such vectors e,

$$\begin{aligned}\mathbb{E}_{X_N}[e^{N\theta\langle e,X_N e\rangle}] &= \left[\prod_{i<j} T_{\mu^N_{i,j}}(2\sqrt{N}\theta e_i e_j)\right]\left[\prod_{i=1}^N T_{\mu^N_{i,i}}(\sqrt{N}\theta e_i^2)e^{N\theta d_i e_i^2}\right]\\ &\geqslant \left[\prod_{i<j} e^{(1-\delta)2N\theta^2 e_i^2 e_j^2}\right]\left[\prod_{i=1}^N e^{(1-\delta)N\theta^2 e_i^4 + N\theta d_i e_i^2}\right]\\ &= e^{(1-\delta)N\theta^2} e^{N\theta\langle e, D_N e\rangle}.\end{aligned}$$

Therefore

$$\begin{aligned}\mathbb{E}_{X_N}[I_N(X_N,\theta)] &= \mathbb{E}_e[\mathbb{E}_{X_N}[e^{N\theta\langle e,X_N e\rangle}]] \geqslant \mathbb{E}_e\left[\mathbf{1}_{\|e\|_\infty \leqslant N^{-3/8}}\mathbb{E}_{X_N}[I_N(X_N,\theta)]\right]\\ &\geqslant e^{(1-\delta)N\theta^2}\mathbb{E}_e\left[\mathbf{1}_{\|e\|_\infty\leqslant N^{-\frac{3}{8}}} e^{N\theta\langle e,D_N e\rangle}\right]\\ &= e^{(1-\delta)N\theta^2}\frac{\mathbb{E}_e\left[\mathbf{1}_{\|e\|_\infty\leqslant N^{-\frac{3}{8}}} e^{N\theta\langle e,D_N e\rangle}\right]}{I_N(D_N,\theta)} I_N(D_N,\theta).\end{aligned}$$

Thus Lemma 5.4.4, which is proved below, and Proposition 5.3.1 give us

$$\liminf_{N\to\infty}\frac{1}{N}\log\mathbb{E}_{X_N}[I_N(X_N,\theta)] \geqslant (1-\delta)\theta^2 + J(\mu_D,\theta,\mathtt{r}(\mu_D))$$

for every $\delta > 0$. □

Proof of Lemma 5.4.3. Let $\mu \neq \delta_0$ be a centered measure on $\mathbb{R}$, and write $\mu(f)$ for the integral of a function f against μ. Whenever $x \in \mathbb{R}$, we have $e^x \geqslant 1 + x + \frac{x^2}{2} + \frac{x^3}{6}$; thus

$$T_\mu(t) \geqslant 1 + \frac{t^2\mu(x^2)}{2} + \frac{t^3\mu(x^3)}{6} \geqslant 1 + \frac{t^2\mu(x^2)}{2} - \frac{|t|^3\mu(|x|^3)}{6}.$$

Now it is standard that the bound $T_\mu(t) \leqslant \exp(\frac{t^2\mu(x^2)}{2})$ implies

$$\mu(|x|^3) \leqslant 3(2\mu(x^2))^{3/2}\Gamma(3/2) \leqslant 8\mu(x^2)^{3/2}.$$

Then the result follows from the limit

$$\lim_{t\to 0}\frac{1}{|t|}\left[\frac{\log\left[1+\frac{t^2\mu(x^2)}{2}-\frac{8|t|^3\mu(x^2)^{3/2}}{6}\right]}{\left(\frac{t^2\mu(x^2)}{2}\right)}-1\right]=-\frac{8\sqrt{\mu(x^2)}}{3}.$$

The speed of convergence in this limit can only depend on μ through $\mu(x^2)$; thus in the result we may choose $\varepsilon(\delta)$ uniformly in the distributions $\mu_{i,j}^N$. □

Proof of Lemma 5.4.4. This builds on the proof of Lemma 14 in [101]. Notice that the upper bound in Equation (5.4.2) is for free; we only need to show the lower bound.

It is well known that

$$(e_1,\dots,e_N)\stackrel{d}{=}\left(\frac{g_1}{\|g\|_2},\dots,\frac{g_N}{\|g\|_2}\right)$$

where $g=(g_1,\dots,g_N)$ is a standard Gaussian vector in $\mathbb{R}^N$. The idea is to work in this Gaussian representation, relying on the fact that $\|g\|$ will concentrate around $\sqrt{N}$.

Towards this end, we rewrite our desired inequality as

$$\liminf_{N\to\infty}\frac{1}{N}\log\frac{\mathbb{E}\left[\mathbf{1}_{\frac{\|g\|_\infty}{\|g\|_2}\leqslant N^{-3/8}}\exp\left(N\theta\frac{\sum_{i=1}^N d_i g_i^2}{\sum_{i=1}^N g_i^2}\right)\right]}{\mathbb{E}\left[\exp\left(N\theta\frac{\sum_{i=1}^N d_i g_i^2}{\sum_{i=1}^N g_i^2}\right)\right]}\geqslant 0.$$

Since standard Gaussian measure is isotropic, we may and will assume for the remainder of this proof that the d_i's are ordered as $d_1\geqslant\cdots\geqslant d_N$. Write $v=v_N$ for the unique solution in $(d_1-\frac{1}{2\theta},+\infty)$ of the equation

$$\frac{1}{2\theta}\frac{1}{N}\sum_{i=1}^N\frac{1}{v+\frac{1}{2\theta}-d_i}=1.$$

(This exists and is unique because the left-hand side is a strictly decreasing positive function of $v\in(d_1-\frac{1}{2\theta},+\infty)$, tending to infinity as $v\downarrow d_1-\frac{1}{2\theta}$ and tending to zero as $v\to\infty$.)

Let us pause to collect some facts about v. If we write

$$d_{\max}=d_{\max}(N_0)=\sup_{N\geqslant N_0}\left(\max_{i=1}^N|d_i|\right)$$

for N_0 large enough, then we have [118, Fact 2.4(3)] that $v \leqslant d_1 \leqslant d_{\max}$, and by definition $v \geqslant d_1 - \frac{1}{2\theta} \geqslant -d_{\max} - \frac{1}{2\theta}$, so

$$|v| \leqslant d_{\max} + \frac{1}{2\theta}. \tag{5.4.3}$$

Furthermore, the proof of [101, Theorem 2] shows that, since $\theta < \theta_c$, there exists some small $\eta > 0$ such that

$$\text{for all } i, \qquad 1 + 2\theta v - 2\theta d_i \geqslant \eta. \tag{5.4.4}$$

By the proof of [101, Lemma 14] (for the first inequality) and Equation (5.4.3) (for the second), for every $0 < \kappa < \frac{1}{2}$ and N large enough depending on κ, we have

$$\begin{aligned} \frac{1}{\mathbb{E}\left[\exp\left(N\theta \frac{\sum_{i=1}^N d_i g_i^2}{\sum_{i=1}^N g_i^2}\right)\right]} &\geqslant \frac{1}{2}\prod_{i=1}^N \left[\sqrt{1+2\theta v - 2\theta d_i}\right] e^{-N\theta v - N^{1-\kappa}\theta(|v| + d_{\max})} \\ &\geqslant \frac{1}{2}\prod_{i=1}^N \left[\sqrt{1+2\theta v - 2\theta d_i}\right] e^{-N\theta v - N^{1-\kappa}\theta(2d_{\max} + \frac{1}{2\theta})}. \end{aligned} \tag{5.4.5}$$

For $0 < \kappa < \frac{1}{2}$, we introduce the event $A_N(\kappa) = \left\{\left|\frac{\|g\|_2^2}{N} - 1\right| \leqslant N^{-\kappa}\right\}$. Now the same arguments from [101, Lemma 14], along with Equation (5.4.3), give

$$\begin{aligned} &\mathbb{E}\left[\mathbf{1}_{\frac{\|g\|_\infty}{\|g\|_2} \leqslant N^{-3/8}} \exp\left(N\theta \frac{\sum_{i=1}^N d_i g_i^2}{\sum_{i=1}^N g_i^2}\right)\right] \\ &\geqslant \mathbb{E}\left[\mathbf{1}_{A_N(\kappa)} \mathbf{1}_{\frac{\|g\|_\infty}{\|g\|_2} \leqslant N^{-3/8}} \exp\left(N\theta \frac{\sum_{i=1}^N d_i g_i^2}{\sum_{i=1}^N g_i^2}\right)\right] \\ &\geqslant e^{N\theta v - N^{1-\kappa}\theta(d_{\max} + |v|)} \mathbb{E}\left[\mathbf{1}_{A_N(\kappa)} \mathbf{1}_{\frac{\|g\|_\infty}{\|g\|_2} \leqslant N^{-3/8}} \exp\left(\sum_{i=1}^N \theta(d_i - v) g_i^2\right)\right] \\ &\geqslant e^{N\theta v - N^{1-\kappa}\theta(2d_{\max} + \frac{1}{2\theta})} \prod_{i=1}^N \left[\frac{1}{\sqrt{1+2\theta v - 2\theta d_i}}\right] P_N^v\left(A_N(\kappa), \frac{\|g\|_\infty}{\|g\|_2} \leqslant N^{-3/8}\right), \end{aligned} \tag{5.4.6}$$

where $P_N^v = P_N^{v, D_N, \theta}$ is the probability measure on $\mathbb{R}^N$ defined by

$$P_N^v(\mathrm{d}g_1, \ldots, \mathrm{d}g_N) = \frac{1}{\sqrt{2\pi}^N} \prod_{i=1}^N \left[\sqrt{1+2\theta v - 2\theta d_i} e^{-\frac{1}{2}(1+2\theta v - 2\theta d_i) g_i^2} \, \mathrm{d}g_i\right].$$

By Equations (5.4.5) and (5.4.6), we are done if we can show that

$$\lim_{N\to\infty} P_N^v\left(A_N(\kappa), \frac{\|g\|_\infty}{\|g\|_2} \leqslant N^{-3/8}\right) = 1.$$

The proof of [101, Lemma 14] shows that, for our choice of v and since we have chosen $\theta < \theta_c$, we have

$$P_N^v(A_N(\kappa)^c) = o(1),$$

so it remains only to bound

$$\begin{aligned} P_N^v\left(A_N(\kappa), \frac{\|g\|_\infty}{\|g\|_2} > N^{-3/8}\right) &\leqslant \sum_{i=1}^N P_N^v\left(A_N(\kappa), \frac{|g_i|^2}{\|g\|_2^2} \geqslant N^{-3/4}\right) \\ &\leqslant \sum_{i=1}^N P_N^v\left(|g_i| \geqslant \sqrt{(N - N^{1-\kappa})N^{-3/4}}\right) \\ &\leqslant \sum_{i=1}^N P_N^v\left(|g_i| \geqslant \frac{1}{2}N^{1/8}\right) \end{aligned}$$

for N large enough depending on κ. But now we observe that

$$\tilde{g}_i = \sqrt{1 + 2\theta v - 2\theta d_i} g_i$$

are i.i.d. standard normal variables under P_N^v, so that by Equation (5.4.4) we have

$$\begin{aligned} \sum_{i=1}^N P_N^v\left(|g_i| \geqslant \frac{1}{2}N^{1/8}\right) &= \sum_{i=1}^N P_N^v\left(|\tilde{g}_i| \geqslant \frac{1}{2}N^{1/8}\sqrt{1 + 2\theta v - 2\theta d_i}\right) \\ &\leqslant N P_N^v\left(|\tilde{g}_i| \geqslant \frac{\sqrt{\eta}}{2}N^{1/8}\right) \leqslant N \exp\left(-\frac{\eta}{8}N^{1/4}\right) \end{aligned}$$

which is $o(1)$. This concludes the proof. □

5.5.1 Proof overview. The proof of Proposition 5.3.9 is broken into the following three lemmata. We emphasize that Lemma 5.5.3 is perhaps the main technical difficulty of the present paper, and could be useful by itself. As throughout the paper, proofs are only written for $\beta = 1$ and thus we drop β from all notations.

Lemma 5.5.1. *If Proposition 5.3.9 holds for $\theta = 0$, then it holds for all $\theta > 0$. (For the second point, the same $\gamma > 0$ works for all $\theta \geqslant 0$.)*

Lemma 5.5.2. *For any $K > 2d_{\max}$,*

$$\mathbb{P}_N(\lambda_N(X_N) > K) \leqslant 4\exp\left(N\left(5 - \frac{K}{8\sqrt{2}}\right)\right),$$
$$\mathbb{P}_N(\lambda_1(X_N) < -K) \leqslant 4\exp\left(N\left(5 - \frac{K}{8\sqrt{2}}\right)\right).$$

In particular, the first point of Proposition 5.3.9 is true for $\theta = 0$.

Lemma 5.5.3. *Under Assumption 2, the second point of Proposition 5.3.9 is true for $\theta = 0$: There exists $\gamma > 0$ such that, for any $0 < \kappa < \gamma$,*

$$\lim_{N\to\infty} \frac{1}{N} \log \mathbb{P}_N(d(\hat{\mu}_{X_N}, \rho_{sc} \boxplus \mu_D) > N^{-\kappa}) = -\infty.$$

Note that this result is the only place in the paper where we use Assumption 2.

5.5.2 Proof of Lemma 5.5.1.

Fix $\theta > 0$. Lemma 5.4.2 gives sharp lower bounds on

$$\mathbb{E}_{X_N}[I_N(X_N, \theta)]$$

for subcritical θ values, but here we need a much weaker lower bound for all positive θ values. To-

wards this end, notice that whenever μ is a centered measure on $\mathbb{R}$, Jensen's gives us $\inf_{t\in\mathbb{R}} T_\mu(t) \geqslant 1$. Thus for every unit vector e we have

$$\begin{aligned}\mathbb{E}_{X_N}[e^{N\theta\langle e, X_N e\rangle}] &= \left[\prod_{i<j} T_{\mu^N_{i,j}}(2\sqrt{N}\theta e_i e_j)\right]\left[\prod_{i=1}^N T_{\mu^N_{i,i}}(\sqrt{N}\theta e_i^2) e^{N\theta d_i e_i^2}\right] \\ &\geqslant \prod_{i=1}^N e^{N\theta d_i e_i^2} \geqslant e^{-N\theta d_{\max}}.\end{aligned} \tag{5.5.1}$$

Now, whenever $A = A_N$ is a Borel subset of the space of $N \times N$ real matrices, Equation (5.5.1) and Cauchy-Schwarz give us, for N sufficiently large depending on θ,

$$\begin{aligned}\mathbb{P}^\theta_N(A) = \frac{\mathbb{E}_{X_N}[\mathbf{1}_{X_N\in A} I_N(X_N,\theta)]}{\mathbb{E}_{X_N}[I_N(X_N,\theta)]} &\leqslant e^{N\theta d_{\max}}\,\mathbb{E}_{X_N,e}[\mathbf{1}_{X_N\in A} e^{N\theta\langle e, X_N e\rangle}] \\ &\leqslant e^{N\theta d_{\max}}\sqrt{\mathbb{P}_N(A)\mathbb{E}_{X_N,e}[e^{2N\theta\langle e, X_N e\rangle}]} = e^{N\theta d_{\max}}\sqrt{\mathbb{P}_N(A)\mathbb{E}_{X_N}[I_N(X_N, 2\theta)]}.\end{aligned}$$

Thus for any sequence $\{A_N\}$ we have, from Lemma 5.4.1,

$$\limsup_{N\to\infty}\frac{1}{N}\log\mathbb{P}^\theta_N(A_N) \leqslant \theta d_{\max} + \frac{1}{2}\limsup_{N\to\infty}\frac{1}{N}\log\mathbb{P}_N(A_N) + \frac{(2\theta)^2 + J(\mu_D, 2\theta, \mathfrak{r}(\mu_D))}{2}.$$

This estimate gives us the following two points, from which we can verify the various claims of Proposition 5.3.9 by taking various choices of $\{A_N\}$ and $\{A_{M,N}\}$.

– If $\{A_N\}$ is such that $\lim_{N\to\infty}\frac{1}{N}\log\mathbb{P}_N(A_N) = -\infty$, then for all $\theta > 0$ we have

$$\limsup_{N\to\infty}\frac{1}{N}\log\mathbb{P}^\theta_N(A_N) = -\infty.$$

– If $\{A_{M,N}\}$ is such that $\lim_{M\to\infty}\limsup_{N\to\infty}\frac{1}{N}\log\mathbb{P}_N(A_{M,N}) = -\infty$, then for all $\theta > 0$ we have

$$\lim_{M\to\infty}\limsup_{N\to\infty}\frac{1}{N}\log\mathbb{P}^\theta_N(A_{M,N}) = -\infty.$$

5.5.3 Proof of Lemma 5.5.2.

For Lemma 5.5.2, notice that it suffices to bound $\mathbb{P}_N(\|X_N\| \geqslant K)$. But

$$\mathbb{P}_N(\|X_N\| > K) \leqslant \mathbb{P}_N\left(\left\|\frac{W_N}{\sqrt{N}}\right\| > \frac{K}{2}\right) + \mathbb{P}_N\left(\|D_N\| > \frac{K}{2}\right)$$

and the second term vanishes for K large enough, so we only need to control the first term. But this was done in [99, Lemma 1.9]. The constants are slightly worse for the $\beta = 2$ estimate, and we phrase Lemma 5.5.2 in terms of these worse constants.

5.5.4 Proof of Lemma 5.5.3.

Lemma 5.5.4. *With C and ε_0 as in Assumption 2, then for any $\eta \leqslant 1$ we have*

$$\sup_{E\in\mathbb{R}}\left|G_{\rho_{sc}\boxplus\hat{\mu}_{D_N}}(E+i\eta) - G_{\rho_{sc}\boxplus\mu_D}(E+i\eta)\right| \leqslant \frac{8\sqrt{C}N^{-\frac{\varepsilon_0}{2}}}{\eta^2}.$$

Proof. By recalling the definition of the Dudley distance and by calculating the L^∞ norm and Lipschitz constants of the function $y \mapsto \frac{1}{E+i\eta-y}$, we find that

$$\left|G_{\rho_{\mathrm{sc}}\boxplus\hat{\mu}_{D_N}}(E+i\eta) - G_{\rho_{\mathrm{sc}}\boxplus\mu_D}(E+i\eta)\right| \leqslant \frac{2}{\eta^2}d(\rho_{\mathrm{sc}}\boxplus\hat{\mu}_{D_N}, \rho_{\mathrm{sc}}\boxplus\mu_D),$$

uniformly in $E \in \mathbb{R}$.

Now we control $d(\rho_{\mathrm{sc}}\boxplus\hat{\mu}_{D_N}, \rho_{\mathrm{sc}}\boxplus\mu_D)$ in terms of $d(\hat{\mu}_{D_N}, \mu_D)$. Write d_L for the Lévy distance between probability measures

$$d_L(\mu,\nu) = \inf\{\varepsilon > 0 : \mu(A) \leqslant \nu(A^\varepsilon) + \varepsilon \text{ for all Borel A}\}.$$

Then it is classical [71, Corollary 11.6.5, Theorem 11.3.3] that, whenever μ and ν are probability measures on $\mathbb{R}$,

$$\frac{1}{2}d(\mu,\nu) \leqslant d_L(\mu,\nu) \leqslant 2\sqrt{d(\mu,\nu)}.$$

On the other hand, [48, Proposition 4.13] says that

$$d_L(\rho_{\mathrm{sc}} \boxplus \hat{\mu}_{D_N}, \rho_{\mathrm{sc}} \boxplus \mu_D) \leqslant d_L(\rho_{\mathrm{sc}}, \rho_{\mathrm{sc}}) + d_L(\hat{\mu}_{D_N}, \mu_D) = d_L(\hat{\mu}_{D_N}, \mu_D).$$

Putting these together, we obtain

$$d(\rho_{\mathrm{sc}} \boxplus \mu_D, \rho_{\mathrm{sc}} \boxplus \hat{\mu}_{D_N}) \leqslant 2d_L(\rho_{\mathrm{sc}} \boxplus \mu_D, \rho_{\mathrm{sc}} \boxplus \hat{\mu}_{D_N}) \leqslant 2d_L(\hat{\mu}_{D_N}, \mu_D) \leqslant 4\sqrt{d(\hat{\mu}_{D_N}, \mu_D)}.$$

This finishes the proof by Assumption 2. □

Lemma 5.5.5. *Fix some $A > 0$ independent of N. If $\delta > 0$ is chosen sufficiently small, then*

$$\int_{-A}^{A} \left| \mathbb{E}_{X_N}[G_{\hat{\mu}_{X_N}}(E + iN^{-\delta})] - G_{\rho_{sc} \boxplus \mu_D}(E + iN^{-\delta}) \right| \mathrm{d}E = O(N^{2\delta - \min(0.99, \frac{\varepsilon_0}{2})}).$$

We first give an informal overview of the proof. We will compare $\mathbb{E}_{X_N}[G_{\hat{\mu}_{X_N}}(\cdot)]$ and $G_{\rho_{\mathrm{sc}} \boxplus \mu_D}(\cdot)$ via three intermediate comparisons. First, we will import a local law to show that with high probability and for appropriate z values,

$$G_{\hat{\mu}_{X_N}}(z) \approx -\frac{1}{N} \operatorname{tr} M_{\mathrm{MDE}}(z)$$

where the matrix $M_{\mathrm{MDE}}(z) = M_{N,\mathrm{MDE}}(z)$ exactly solves a matrix equation called the Matrix Dyson Equation (MDE). (The negatives appear since the convention in the local-law literature is to define the Stieltjes transform of a measure as $\int \frac{\mu(\mathrm{d}y)}{z-y}$ instead of our $\int \frac{\mu(\mathrm{d}y)}{y-z}$. We have preferred to stick to that convention when working in that vein, so that the reader can more easily cross-reference.) Then we will show that a matrix $M_{\mathrm{Wig}}(z) = M_{N,\mathrm{Wig}}(z)$ whose normalized trace is exactly $-G_{\rho_{\mathrm{sc}} \boxplus \hat{\mu}_{D_N}}$ approximately solves the MDE; standard arguments about the so-called stability of the MDE will then show

$$-\frac{1}{N} \operatorname{tr} M_{\mathrm{MDE}}(z) \approx -\frac{1}{N} \operatorname{tr} M_{\mathrm{Wig}}(z) = G_{\rho_{\mathrm{sc}} \boxplus \hat{\mu}_{D_N}}(z).$$

Finally, we will use Lemma 5.5.4 to show

$$G_{\rho_{\mathrm{sc}} \boxplus \hat{\mu}_{D_N}}(z) \approx G_{\rho_{\mathrm{sc}} \boxplus \mu_D}(z).$$

Notice that all quantities here, except for $G_{\hat{\mu}_{X_N}}$, are deterministic.

Proof of Lemma 5.5.5. Throughout, we write $z = E + i\eta$. Later, we will decide how to choose $\eta = \eta(N)$.

For a matrix $M \in \mathbb{C}^{N \times N}$, we define its imaginary part as $\Im(M) = \frac{1}{2i}[M - M^*]$. Whenever $\mathcal{S} : \mathbb{C}^{N \times N} \to \mathbb{C}^{N \times N}$ is a linear operator preserving the set $\{M : \Im(M) > 0\}$ that is self-adjoint with respect to the inner product $\langle R, T \rangle = \operatorname{tr}(R^* T)$, it is known [106] that the following constrained equation admits a unique solution:

$$0 = \mathrm{Id} + (z\,\mathrm{Id} - D_N + \mathcal{S}[M(z)])M(z) \qquad \text{subject to} \qquad \Im(M(z)) = \frac{1}{2i}[M(z) - M^*(z)] > 0. \quad (5.5.2)$$

Furthermore, $M(z)$ is a holomorphic matrix-valued function of z. In particular, we will be interested in the unique solutions to this equation corresponding to two operators $\mathcal{S}$:

$$\begin{aligned} \mathcal{S}_{\mathrm{MDE}}[M] = \frac{1}{N} \operatorname{tr}(M)\,\mathrm{Id} + \frac{1}{N} M^T \qquad &\text{induces the solution} \qquad M_{\mathrm{MDE}}(z), \\ \mathcal{S}_{\mathrm{Wig}}[M] = \frac{1}{N} \operatorname{tr}(M)\,\mathrm{Id} \qquad &\text{induces the solution} \qquad M_{\mathrm{Wig}}(z). \end{aligned}$$

By rearranging (5.5.2) and taking the normalized trace, one sees that $s(z) = -\frac{1}{N} \operatorname{tr} M_{\mathrm{Wig}}(z)$ satisfies the Pastur equation

$$s(z) = G_{\hat{\mu}_{D_N}}(z - s(z))$$

which characterizes the Stieltjes transform of $\rho_{\mathrm{sc}} \boxplus \hat{\mu}_{D_N}$ ([131], see also [130, Section 2.2]). Hence

$$-\frac{1}{N} \operatorname{tr} M_{\mathrm{Wig}}(z) = G_{\rho_{\mathrm{sc}} \boxplus \hat{\mu}_{D_N}}(z).$$

– For any $\delta > 0$, write $\mathbb{H} = \{z = E + i\eta \in \mathbb{C} : \eta > 0\}$ and define the complex domain

$$\mathcal{D}^{\delta}_{\text{far}} = \{z \in \mathbb{H} : |z| \leqslant N^{100}, \eta \geqslant N^{-\delta}\}.$$

(The notation reminds us that points in this domain are relatively far from the real line; typically in local laws the optimal scale is $\eta \gg \frac{1}{N}$.) Then [75, Theorem 2.1] tells us that there is a universal constant $c > 0$ such that, for any sufficiently small $\varepsilon > 0$, there exists $C = C(\varepsilon)$ such that

$$\mathbb{P}\left(\left|G_{\hat{\mu}_{X_N}}(z) + \frac{1}{N}\operatorname{tr}(M_{\text{MDE}}(z))\right| \leqslant \frac{N^{\varepsilon}}{N} \quad \text{in} \quad \mathcal{D}^{c\varepsilon}_{\text{far}}\right) \geqslant 1 - CN^{-100}.$$

Since $\frac{1}{N}\operatorname{tr}(M_{\text{MDE}}(z))$ is known by [5, Proposition 2.1] to be the Stieltjes transform of some measure, we also have the trivial bounds

$$\left|G_{\hat{\mu}_{X_N}}(E + i\eta)\right| \leqslant \frac{1}{\eta} \quad \text{and} \quad \left|\frac{1}{N}\operatorname{tr}(M_{\text{MDE}}(E + i\eta))\right| \leqslant \frac{1}{\eta}.$$

If $\eta = N^a$ for some $-c\varepsilon < a < 0$, then for N sufficiently large we have $\{E + i\eta : |E| \leqslant A\} \subset \mathcal{D}^{c\varepsilon}_{\text{far}}$; thus whenever $|E| \leqslant A$ and η is as above we have

$$\mathbb{E}_{X_N}\left|G_{\hat{\mu}_{X_N}}(E + i\eta) + \frac{1}{N}\operatorname{tr} M_{\text{MDE}}(E + i\eta)\right| \leqslant \frac{N^{\varepsilon}}{N} + \frac{2C}{\eta}N^{-100},$$

so that

$$\int_{-A}^{A}\left|\mathbb{E}_{X_N}[G_{\hat{\mu}_{X_N}}(E + i\eta)] + \frac{1}{N}\operatorname{tr} M_{\text{MDE}}(E + i\eta)\right| \mathrm{d}E \leqslant 2A\left(\frac{N^{\varepsilon}}{N} + \frac{2C}{\eta}N^{-100}\right).$$

– The following type of stability analysis is standard in the MDE literature; our exact line of argument follows most closely that of [6]. By the definition of M_{Wig} and since $\mathcal{S}_{\text{MDE}}[M] =$

$\mathcal{S}_{\mathrm{Wig}}[M] + \frac{1}{N}M^T$, we have

$$M_{\mathrm{Wig}}^{-1}(z) = z\,\mathrm{Id} - D_N + \mathcal{S}_{\mathrm{MDE}}[M_{\mathrm{Wig}}(z)] - \underbrace{\frac{M_{\mathrm{Wig}}^T(z)}{N}}_{=:E(z)}.$$

As the notation suggests, we will consider $E(z)$ as an error term, so that $M_{\mathrm{Wig}}(z)$ approximately solves Equation (5.5.2) with $\mathcal{S} = \mathcal{S}_{\mathrm{MDE}}$. Indeed, from (4.1) in [5] we have, for every $z \in \mathbb{H}$,

$$\max\{\|M_{\mathrm{MDE}}(z)\|, \|M_{\mathrm{Wig}}(z)\|\} \leqslant \frac{1}{\eta}. \tag{5.5.3}$$

(Recall $\|\cdot\|$ is the operator norm induced by the standard Euclidean norm.) In particular, we have

$$\|E(z)\| \leqslant \frac{1}{N\eta}. \tag{5.5.4}$$

Now manipulations of the MDE like those leading up to [6, (4.25)] yield the quadratic inequality

$$\begin{aligned}&\|M_{\mathrm{Wig}}(z) - M_{\mathrm{MDE}}(z)\| \\ &\leqslant \|\mathscr{L}^{-1}(z)\|\|M_{\mathrm{MDE}}(z)\|(\|E(z)\|\|M_{\mathrm{Wig}}(z)\| + \|\mathcal{S}_{\mathrm{MDE}}\|\|M_{\mathrm{Wig}}(z) - M_{\mathrm{MDE}}(z)\|^2),\end{aligned}$$

where $\mathscr{L}(z) : \mathbb{C}^{N\times N} \to \mathbb{C}^{N\times N}$ is the invertible operator

$$\mathscr{L}(z)[T] = T - M_{\mathrm{MDE}}(z)\mathcal{S}_{\mathrm{MDE}}[T]M_{\mathrm{MDE}}(z)$$

and norms on operators from $\mathbb{C}^{N\times N}$ to itself are operator norms with respect to $\|\cdot\|$. Using (5.5.3), (5.5.4), and the estimate $\|\mathcal{S}_{\mathrm{MDE}}\| \leqslant 2$, this simplifies to

$$\|M_{\mathrm{Wig}}(z) - M_{\mathrm{MDE}}(z)\| \leqslant \frac{\|\mathscr{L}^{-1}(z)\|}{\eta}\left(\frac{1}{N\eta^2} + 2\|M_{\mathrm{Wig}}(z) - M_{\mathrm{MDE}}(z)\|^2\right). \tag{5.5.5}$$

From [6, (3.23), (3.22), Convention 3.5] combined with (5.5.3), there exists a constant C such

that, for all z,

$$\|\mathscr{L}^{-1}(z)\| \leqslant C\left(1 + \frac{1}{\eta^2} + \frac{\|M_{\mathrm{MDE}}^{-1}(z)\|^9}{\eta^{13}}\right). \tag{5.5.6}$$

We will use this in two regimes, depending on whether $\eta \geqslant 1$ or $\eta \leqslant 1$.

– **Step 1 ($\eta \geqslant 1$):** Taking norms on both sides of (5.5.2) and using (5.5.3), we obtain

$$\|M_{\mathrm{MDE}}^{-1}(z)\| \leqslant |z| + d_{\max} + 2.$$

Recall we are integrating over E in some $[-A, A]$; for such E, we have $|z| \leqslant \eta\sqrt{1 + A^2}$, so (using (5.5.6)) there exist constants C', C'' such that

$$\sup_{|E|\leqslant A, \eta \geqslant 1} \frac{\|M_{\mathrm{MDE}}^{-1}(z)\|}{\eta} \leqslant C', \qquad \sup_{|E|\leqslant A, \eta \geqslant 1} \|\mathscr{L}^{-1}(z)\| \leqslant C''. \tag{5.5.7}$$

Now fix $E \in [-A, A]$ and consider the functions $f_N : (0, \infty) \to \mathbb{R}$ and $g_N^{\pm} : [1, \infty) \to \mathbb{R}$ given by

$$f_N(\eta) = \|M_{\mathrm{Wig}}(E + i\eta) - M_{\mathrm{MDE}}(E + i\eta)\|,$$
$$g_N^{\pm}(\eta) = \frac{\eta}{4C''}\left(1 \pm \sqrt{1 - \frac{8(C'')^2}{N\eta^4}}\right).$$

For $\eta \geqslant \sqrt{8C''}$, the bound (5.5.3) gives us

$$f_N(\eta) \leqslant \frac{2}{\eta} \leqslant \frac{\eta}{4C''} < g_N^{+}(\eta).$$

But the quadratic inequality (5.5.5) with the estimate (5.5.7) inserted tells us that

$$f_N(\eta) \in [0, g_N^{-}(\eta)] \cup [g_N^{+}(\eta), \infty).$$

Furthermore, since $M_{\mathrm{MDE}}(z)$ and $M_{\mathrm{Wig}}(z)$ are holomorphic functions of z, we have that

$f_N(\cdot)$ is continuous; and since $g_N^-(\eta) < g_N^+(\eta)$ for all $\eta > 1$, we conclude $f_N(\eta) \leqslant g_N^-(\eta)$ for all $\eta \geqslant 1$.

– **Step 2 ($\eta \leqslant 1$):** Taking norms on both sides of (5.5.2) and using (5.5.3), we obtain

$$\|M_{\mathrm{MDE}}^{-1}(z)\| \leqslant |z| + d_{\max} + \frac{2}{\eta}.$$

Arguments like those above then give

$$\sup_{|E|\leqslant A, \eta \leqslant 1} \frac{\|\mathscr{L}^{-1}(z)\|}{\eta^{-22}} \leqslant C'''$$

for some new constant C''', which we again insert back in the quadratic inequality (5.5.5). This lets us bound $f_N(\eta)$ with respect to the new functions $h_N^{\pm}(\eta) : [N^{-1/50}, 1] \to \mathbb{R}$ given by

$$h_N^{\pm}(\eta) = \frac{\eta^{23}}{4C'''}\left(1 \pm \sqrt{1 - \frac{8(C''')^2}{N\eta^{48}}}\right),$$

but at $\eta = 1$ and for N large enough we have (using $1 - \sqrt{1-x} \leqslant x$)

$$f_N(1) \leqslant g_N^-(1) \leqslant \frac{2C''}{N} < \frac{1}{4C'''} = h_N^+(1).$$

Thus

$$f_N(\eta) \leqslant h_N^-(\eta) \leqslant \frac{2C'''}{N\eta^{25}},$$

uniformly over $E \in [-A, A]$; hence

$$\int_{-A}^{A} \left| \frac{1}{N} \operatorname{tr}(M_{\mathrm{MDE}}(E + iN^{-\delta}) - M_{\mathrm{Wig}}(E + iN^{-\delta})) \right| \mathrm{d}E \leqslant \frac{4AC'''}{N^{1-25\delta}}$$

for sufficiently small $\delta > 0$.

– If $\eta \leqslant 1$ then Lemma 5.5.4 gives us

$$\begin{aligned}&\int_{-A}^{A}\left|-\frac{1}{N}\operatorname{tr} M_{\mathrm{Wig}}(E+i\eta)-G_{\rho_{\mathrm{sc}}\boxplus\mu_D}(E+i\eta)\right|\mathrm{d}E\\&=\int_{-A}^{A}\left|G_{\rho_{\mathrm{sc}}\boxplus\hat{\mu}_{D_N}}(E+i\eta)-G_{\rho_{\mathrm{sc}}\boxplus\mu_D}(E+i\eta)\right|\mathrm{d}E\leqslant 16A\sqrt{C}\frac{N^{-\frac{\varepsilon_0}{2}}}{\eta^2}.\end{aligned}$$

Combining these estimates, we have the following result: If $\eta = N^{-\delta}$ and δ is sufficiently small, then every assumption we made on η in the above bounds is satisfied and, for all sufficiently small $\varepsilon > 0$,

$$\begin{aligned}\int_{-A}^{A}\left|\mathbb{E}_{X_N}[G_{\hat{\mu}_{X_N}}(E+i\eta)]-G_{\rho_{\mathrm{sc}}\boxplus\mu_D}(E+i\eta)\right|\mathrm{d}E&=O\left(\frac{N^{\varepsilon}}{N}+\frac{1}{\eta N^{100}}+\frac{1}{N\eta^{25}}+\frac{1}{N^{\frac{\varepsilon_0}{2}}\eta^2}\right)\\&=O\Big(N^{2\delta-\min(0.99,\frac{\varepsilon_0}{2})}\Big).\end{aligned}$$

This concludes the proof. □

Lemma 5.5.6. *Write*

$$\begin{aligned}F_{X_N}(x)&=\hat{\mu}_{X_N}((-\infty,x]),\\F_{\rho_{sc}\boxplus\mu_D}(x)&=(\rho_{sc}\boxplus\mu_D)((-\infty,x]).\end{aligned}$$

Then there exists some $\gamma > 0$ such that

$$\sup_x|\mathbb{E}_{X_N}[F_{X_N}(x)]-F_{\rho_{sc}\boxplus\mu_D}(x)|=O(N^{-\gamma}).$$

Proof. To apply a standard method for bounding Kolmogorov-Smirnov distances, we must first show

$$\int_{-\infty}^{\infty}|\mathbb{E}_{X_N}[F_{X_N}(x)]-F_{\rho_{\mathrm{sc}}\boxplus\mu_D}(x)|\,\mathrm{d}x<\infty. \tag{5.5.8}$$

Since $\mathbb{E}_{X_N}[F_{X_N}]$ and $F_{\rho_{\mathrm{sc}}\boxplus\mu_D}$ both take values in $[0,1]$, it suffices to find $M>0$ such that

$$\int_{|x|>M}|\mathbb{E}_{X_N}[F_{X_N}(x)]-F_{\rho_{\mathrm{sc}}\boxplus\mu_D}(x)|\,\mathrm{d}x<\infty.$$

Furthermore, since $\rho_{\mathrm{sc}}\boxplus\mu_D$ is compactly supported, we may take M so large that $F_{\rho_{\mathrm{sc}}\boxplus\mu_D}(x)$ vanishes for $x<-M$ and is identically one for $x>M$. Now,

$$\mathbb{E}_{X_N}[F_{X_N}(x)]=\frac{1}{N}\mathbb{E}_{X_N}\left[\sum_{j=1}^{N}\mathbf{1}_{\lambda_j(X_N)<x}\right]=\frac{1}{N}\sum_{j=1}^{N}\mathbb{P}_N(\lambda_j(X_N)<x)\leqslant\mathbb{P}_N(\lambda_1(X_N)<x)\tag{5.5.9}$$

so that, by Lemma 5.5.2,

$$\int_{-\infty}^{-M}\mathbb{E}_{X_N}[F_{X_N}(x)]\,\mathrm{d}x\leqslant\int_{-\infty}^{-M}4e^{N(5+\frac{x}{8\sqrt{2}})}\,\mathrm{d}x=\frac{32\sqrt{2}}{N}\exp\left[N\left(5-\frac{M}{8\sqrt{2}}\right)\right]<\infty.$$

Similarly,

$$\int_{M}^{\infty}(1-\mathbb{E}_{X_N}[F_{X_N}(x)])\,\mathrm{d}x<\infty$$

which finishes the proof of (5.5.8).

Thus we may import [19, Theorem 2.2], which says that, for any choice of $\eta>0$ and $B>0$, we have

$$\begin{aligned}\sup_x|\mathbb{E}_{X_N}[F_{X_N}(x)]-F_{\rho_{\mathrm{sc}}\boxplus\mu_D}(x)|\leqslant\Bigg[&\frac{1}{\eta}\sup_x\int_{|y|\leqslant5\eta}|F_{\rho_{\mathrm{sc}}\boxplus\mu_D}(x+y)-F_{\rho_{\mathrm{sc}}\boxplus\mu_D}(x)|\,\mathrm{d}y\\&+\frac{2\pi}{\eta}\int_{|x|>B}|\mathbb{E}_{X_N}[F_{X_N}(x)]-F_{\rho_{\mathrm{sc}}\boxplus\mu_D}(x)|\,\mathrm{d}x\\&+\int_{-10B}^{10B}\left|\mathbb{E}_{X_N}[G_{\hat{\mu}_{X_N}}(E+i\eta)]-G_{\rho_{\mathrm{sc}}\boxplus\mu_D}(E+i\eta)\right|\mathrm{d}E\Bigg].\end{aligned}$$

We will control the three terms on the right-hand side in order. In the course these estimates we shall choose the parameters B and $\eta=\eta(N)$.

– Since the compactly supported measure $\rho_{\mathrm{sc}}\boxplus\mu_D$ has L^∞ density [49, Corollary 5], $F_{\rho_{\mathrm{sc}}\boxplus\mu_D}$

is Lipschitz, so we can control the first term by

$$\frac{1}{\eta}\sup_x \int_{|y|\leqslant 5\eta} |F_{\rho_{\mathrm{sc}}\boxplus\mu_D}(x+y) - F_{\rho_{\mathrm{sc}}\boxplus\mu_D}(x)|\,\mathrm{d}y \leqslant 25\eta \|F_{\rho_{\mathrm{sc}}\boxplus\mu_D}(x)\|_{\mathrm{Lip}}.$$

– Choose some $B > \max(|\mathtt{r}(\rho_{\mathrm{sc}} \boxplus \mu_D)|, |\mathtt{l}(\rho_{\mathrm{sc}} \boxplus \mu_D)|)$; then arguments as above show that

$$\frac{2\pi}{\eta}\int_{|x|>B} |\mathbb{E}_{X_N}[F_{X_N}(x)] - F_{\rho_{\mathrm{sc}}\boxplus\mu_D}(x)|\,\mathrm{d}x \leqslant \frac{2\pi}{\eta}\cdot\frac{64\sqrt{2}}{N}\exp\left[N\left(5 - \frac{B}{8\sqrt{2}}\right)\right].$$

Since we will ultimately choose $\eta = N^{-\delta}$ for some small $\delta > 0$, we can choose B so large that this decays exponentially fast.

– If we choose $\eta = N^{-\delta}$ for $\delta > 0$ sufficiently small, then Lemma 5.5.5 tells us that

$$\int_{-10B}^{10B} \left|\mathbb{E}_{X_N}[G_{\hat{\mu}_{X_N}}(E + iN^{-\delta})] - G_{\rho_{\mathrm{sc}}\boxplus\mu_D}(E + iN^{-\delta})\right| \mathrm{d}E = O(N^{2\delta - \frac{\varepsilon_0}{2}}).$$

We combine these to obtain

$$\sup_x |\mathbb{E}_{X_N}[F_{X_N}(x)] - F_{\rho_{\mathrm{sc}}\boxplus\mu_D}(x)| = O(N^{\max(-\delta, 2\delta - \frac{\varepsilon_0}{2})}).$$

□

Lemma 5.5.7. *Under either the Gaussian Hypothesis or the SSGC Hypothesis, there exist positive constants C_1 and C_2 (depending on the constants in those hypotheses) such that*

$$\mathbb{P}_N\left[d(\hat{\mu}_{X_N}, \mathbb{E}_{X_N}[\hat{\mu}_{X_N}]) \geqslant N^{-1/6}\right] \leqslant C_1 N^{1/4}\exp\left(-C_2 N^{7/6}\right).$$

Concentration results of this type are quite classical, using either the Herbst argument under the log-Sobolev assumption, or results of Talagrand under the compact-support assumption. Indeed, results of the former type are available "out of the box"; results of the latter type are available "off the shelf" when D_N vanishes. But when $D_N \neq 0$, the barrier to using existing results is that, even

if the entries of W_N are uniformly compactly supported, the diagonal entries of

$$\sqrt{N}X_N = W_N + \sqrt{N}D_N$$

are supported in boxes that, while of fixed size, may have centers tending to infinity. So we modify the existing proofs for this situation.

Proof of Lemma 5.5.7. Suppose first that we satisfy the log-Sobolev option of the SSGC Hypothesis, that is, that the laws of the entries of W_N satisfy a log-Sobolev inequality with a uniform constant. Since Gaussian measure satisfies the log-Sobolev inequality, the same statement is true under the Gaussian Hypothesis. Furthermore, one can see directly from the definition of the inequality that, if the law of the real random variable X satisfies the logarithmic Sobolev inequality with constant c, then for any deterministic $\alpha \in \mathbb{R}$ the law of $X + \alpha$ also satisfies the logarithmic Sobolev inequality with constant c. Thus the laws of the entries of $\sqrt{N}X_N$ satisfy a log-Sobolev inequality with uniform constant. This uniformity allows us to import the result [103, Corollary 1.4b], which tells us that there exist positive constants C_1 and C_2 such that, for any $\delta > 0$,

$$\mathbb{P}_N[d(\hat{\mu}_{X_N}, \mathbb{E}_{X_N}[\hat{\mu}_{X_N}]) \geqslant \delta] \leqslant \frac{C_1}{\delta^{3/2}} \exp\Big(-C_2 N^2 \delta^5\Big).$$

By choosing $\delta = N^{-1/6}$, this completes the proof under the Gaussian Hypothesis or under the log-Sobolev option of the SSGC Hypothesis.

Next, we turn to the compact-support option of the SSGC Hypothesis, and start by importing the following result.

Lemma 5.5.8. *[103, Theorem 1.3a] Fix some $(a_{i,j})_{i,j \leqslant N} \subset \mathbb{R}^N$, and suppose that there exists a compact set $K \subset \mathbb{R}$ such that the i,jth entry of $\sqrt{N}X_N$ is supported on the compact set $a_{i,j} + K = \{a_{i,j} + k : k \in K\}$. Write $\delta_1(N) = 8|K|\sqrt{\pi}/N$. Let $\mathcal{K} \subset \mathbb{R}$ be compact, and define the class of test functions*

$$\mathcal{F}_{lip,\mathcal{K}} = \{f : \mathrm{supp}(f) \subset \mathcal{K}, \|f\|_\infty + \|f\|_{Lip} \leqslant 1\}.$$

Then, for any $\delta \geqslant 4\sqrt{|\mathcal{K}|\delta_1(N)}$*, we have*

$$\mathbb{P}\left(\sup_{f\in\mathcal{F}_{lip,\mathcal{K}}} |\mathrm{tr}_N(f(X_N)) - \mathbb{E}[\mathrm{tr}_N(f(X_N))]| \geqslant \delta\right) \leqslant \frac{32|\mathcal{K}|}{\delta}\exp\left(-\frac{N^2}{16|K|^2}\left[\frac{\delta^2}{16|\mathcal{K}|} - \delta_1(N)\right]^2\right).$$

(This result was initially stated for centered entries, but by shifting the test function they use to apply [143, Theorem 6.6] the proof goes through.) The authors of [103] then extend this result to a supremum over all bounded Lipschitz functions, not just those that are compactly supported, but in the case that $\mathbb{E}[X_N] = 0$. Their arguments require a bound on $\frac{1}{N}\mathrm{tr}(X_N^2)$, which we replace for our model with

$$\frac{1}{N}\mathrm{tr}(X_N^2) \leqslant \sup\{|x|^2 : x \in K\} + d_{\max}^2 + 1,$$

which is true for N sufficiently large. Following their proofs but substituting this estimate, we obtain the following result, which is analogous to [103, Corollary 1.4a]:

Lemma 5.5.9. *Under the assumptions and notation of Lemma 5.5.8, write* $S = \sup\{|x|^2 : x \in K\}$ *and* $M = \sqrt{8(S + d_{\max}^2 + 1)}$*. Then for any* N *sufficiently large and for any* $\delta > 0$ *satisfying the implicit equation* $\delta > (128(M + \sqrt{\delta})\delta_1(N))^{2/5}$*, we have*

$$\mathbb{P}_{X_N}(d(\hat{\mu}_{X_N}, \mathbb{E}_{X_N}(\hat{\mu}_{X_N})) > \delta) \leqslant \frac{128(M+\sqrt{\delta})}{\delta^{3/2}}\exp\left(-\frac{N^2}{16|K|^2}\left[\frac{\delta^{5/2}}{128(M+\sqrt{\delta})} - \delta_1(N)\right]^2\right).$$

For N sufficiently large, $\delta = N^{-1/6}$ satisfies the implicit equation given in the lemma, and it is easy to show that

$$\left[\frac{\delta^{5/2}}{128(M+\sqrt{\delta})} - \delta_1(N)\right]^2 \geqslant \frac{N^{7/6}}{N^2(512M)^2}$$

for N large enough, which gives the desired result in this case. □

Proof of Lemma 5.5.3. By Lemma 5.5.7, if $\kappa < \frac{1}{6}$ we have

$$\begin{aligned}
&\mathbb{P}_N(d(\hat{\mu}_{X_N}, \rho_{\mathrm{sc}} \boxplus \mu_D) > N^{-\kappa}) \\
&\leqslant \mathbf{1}_{d(\mathbb{E}_{X_N}[\hat{\mu}_{X_N}], \rho_{\mathrm{sc}} \boxplus \mu_D) > \frac{N^{-\kappa}}{2}} + \mathbb{P}_N\left(d(\hat{\mu}_{X_N}, \mathbb{E}_{X_N}[\hat{\mu}_{X_N}]) > \frac{N^{-\kappa}}{2}\right) \\
&\leqslant \mathbf{1}_{d(\mathbb{E}_{X_N}[\hat{\mu}_{X_N}], \rho_{\mathrm{sc}} \boxplus \mu_D) > \frac{N^{-\kappa}}{2}} + C_1 N^{1/4} \exp\left(-\frac{C_2}{2} N^{7/6}\right).
\end{aligned}$$

Now we wish to estimate $d(\mathbb{E}_{X_N}[\hat{\mu}_{X_N}], \rho_{\mathrm{sc}} \boxplus \mu_D)$, in order to show that the above indicator vanishes. Towards this end, choose an arbitrary test function f with $\|f\|_{L^\infty} + \|f\|_{\mathrm{Lip}} \leqslant 1$.

First we estimate the tails. For M large enough, Equation (5.5.9) gives us

$$\begin{aligned}
&\left|\int_{-\infty}^{-M} f(x)(\mathbb{E}_{X_N}[\hat{\mu}_{X_N}] - (\rho_{\mathrm{sc}} \boxplus \mu_D))(\mathrm{d}x)\right| \\
&= \left|\int_{-\infty}^{-M} f(x)\mathbb{E}_{X_N}[\hat{\mu}_{X_N}](\mathrm{d}x)\right| \leqslant \|f\|_{L^\infty}\mathbb{E}_{X_N}[F_{X_N}(-M)] \\
&\leqslant \mathbb{E}_{X_N}[F_{X_N}(-M)] \leqslant \mathbb{P}_N(\lambda_1(X_N) < -M) \leqslant e^{-N}
\end{aligned}$$

where the last inequality follows from Lemma 5.5.2. Similarly,

$$\left|\int_{M}^{\infty} f(x)(\mathbb{E}_{X_N}[\hat{\mu}_{X_N}] - (\rho_{\mathrm{sc}} \boxplus \mu_D))(\mathrm{d}x)\right| \leqslant 1 - \mathbb{E}_{X_N}[F_{X_N}(M)] \leqslant e^{-N}.$$

Thus it remains to estimate $\left|\int_{-M}^{M} f(x)(\mathbb{E}_{X_N}[\hat{\mu}_{X_N}] - (\rho_{\mathrm{sc}} \boxplus \mu_D))(\mathrm{d}x)\right|$. We will do this by approximating f by a test function smooth enough to integrate by parts.

More precisely, suppose first that f is C^1 and that $\|f'\|_{L^\infty} \leqslant 1$. Then

$$\begin{aligned}
&\left|\int_{-M}^{M} f(x)(\mathbb{E}_{X_N}[\hat{\mu}_{X_N}] - (\rho_{\mathrm{sc}} \boxplus \mu_D))(\mathrm{d}x)\right| \\
&= \left|\int_{-M}^{M} f(x)\,\mathrm{d}(\mathbb{E}_{X_N}[F_{X_N}(x)] - F_{\rho_{\mathrm{sc}}\boxplus\mu_D}(x))\right| \\
&\leqslant (2M\|f'\|_{L^\infty} + \|f\|_{L^\infty} + \|f\|_{L^\infty})\|\mathbb{E}_{X_N}[F_{X_N}] - F_{\rho_{\mathrm{sc}}\boxplus\mu_D}\|_{L^\infty} \\
&\leqslant (2M+2)\|\mathbb{E}_{X_N}[F_{X_N}] - F_{\rho_{\mathrm{sc}}\boxplus\mu_D}\|_{L^\infty}.
\end{aligned}$$

Now suppose that f only satisfies $\|f\|_{L^\infty} + \|f\|_{\mathrm{Lip}} \leqslant 1$. Since $[-M, M]$ is a compact set independent of N, we may choose $g \in C^1$ with $\|g'\|_{L^\infty([-M,M])} \leqslant 1$ and

$$\|f - g\|_{L^\infty([-M,M])} \leqslant (M+1)\|\mathbb{E}_{X_N}[F_{X_N}] - F_{\rho_{\mathrm{sc}}\boxplus\mu_D}\|_{L^\infty}.$$

Thus

$$\begin{aligned}
\left|\int_{-M}^{M} (f(x) - g(x))(\mathbb{E}_{X_N}[\hat{\mu}_{X_N}] - (\rho_{\mathrm{sc}} \boxplus \mu_D))(\mathrm{d}x)\right| &\leqslant 2\|f - g\|_{L^\infty([-M,M])} \\
&\leqslant (2M+2)\|\mathbb{E}_{X_N}[F_{X_N}] - F_{\rho_{\mathrm{sc}}\boxplus\mu_D}\|_{L^\infty}.
\end{aligned}$$

Combining these and and optimizing over f, we have

$$d(\mathbb{E}_{X_N}[\hat{\mu}_{X_N}], \rho_{\mathrm{sc}} \boxplus \mu_D) \leqslant 2e^{-N} + 4(M+1)\|\mathbb{E}_{X_N}[F_{X_N}] - F_{\rho_{\mathrm{sc}}\boxplus\mu_D}\|_{L^\infty} = O(N^{-\gamma}),$$

where the last equality follows from Lemma 5.5.6. Thus if we choose $0 < \kappa < \gamma$, we have

$$\mathbf{1}_{d(\mathbb{E}_{X_N}[\hat{\mu}_{X_N}], \rho_{\mathrm{sc}}\boxplus\mu_D) > \frac{N-\kappa}{2}} = 0$$

for sufficiently large N; in particular this shows us that

$$\mathbb{P}_N(d(\hat{\mu}_{X_N}, \rho_{\mathrm{sc}} \boxplus \mu_D) > N^{-\kappa}) \leqslant C_1 N^{1/4} \exp\left(-\frac{C_2}{2} N^{7/6}\right)$$

from which point it is easy to conclude the proof. □

5.6 Properties of the rate function

The purpose of this section is to show that the supremum in the definition of

$$I^{(\beta)}(x) = \sup_{\theta \geqslant 0} I^{(\beta)}(x, \theta)$$

is achieved at a value $\theta_x^{(\beta)}$, which is unique (except for $x = \mathtt{r}(\rho_{\mathrm{sc}} \boxplus \mu_D)$, where it is chosen by convention) and which depends injectively on x. This implies that, in the large-deviation upper bound established for tilted measures in Theorem 5.3.4, the rate function has a unique zero; this property was crucial in the proof of Lemma 5.3.11 above. At the end of this section, we establish goodness of $I^{(\beta)}(\cdot)$.

Proposition 5.6.1. *For every $x > \mathtt{r}(\rho_{sc} \boxplus \mu_D)$ and for each $\beta = 1, 2$, there exists a unique $\theta \geqslant 0$, which we will write $\theta_x^{(\beta)}$, such that*

$$I^{(\beta)}(x) = \sup_{\theta \geqslant 0} I^{(\beta)}(x, \theta) = I^{(\beta)}(x, \theta_x^{(\beta)}).$$

Furthermore, $I^{(\beta)}(x)$ vanishes uniquely at $x = \mathtt{r}(\rho_{sc} \boxplus \mu_D)$; and if we define by convention

$$\theta^{(\beta)}_{\mathtt{r}(\rho_{sc} \boxplus \mu_D)} = \frac{\beta}{2} G_{\rho_{sc} \boxplus \mu_D}(\mathtt{r}(\rho_{sc} \boxplus \mu_D))$$

then the map $x \mapsto \theta_x^{(\beta)}$ on the domain $\{x \geqslant \mathtt{r}(\rho_{sc} \boxplus \mu_D)\}$ is injective. In particular, whenever $x \neq y$

are at least $\mathtt{r}(\rho_{sc} \boxplus \mu_D)$*, we have*

$$I^{(\beta)}(y) > I^{(\beta)}(y, \theta_x^{(\beta)}).$$

We also have

$$x_c \geqslant \mathtt{r}(\rho_{sc} \boxplus \mu_D)$$

with equality if and only if $G_{\mu_D}(\mathtt{r}(\mu_D)) = G_{\rho_{sc} \boxplus \mu_D}(\mathtt{r}(\rho_{sc} \boxplus \mu_D))$*. In addition, the optimizer for* x_c *is in fact* $\theta_c^{(\beta)}$ *as defined in* (5.3.1)*:*

$$\theta_{x_c}^{(\beta)} = \theta_c^{(\beta)} = \begin{cases} \frac{\beta}{2} G_{\mu_D}(\mathtt{r}(\mu_D)) & \text{if } G_{\mu_D}(\mathtt{r}(\mu_D)) < +\infty, \\ +\infty & \text{otherwise, by convention,} \end{cases}$$

and if $x < x_c$ *then* $\theta_x^{(\beta)} < \theta_{x_c}^{(\beta)}$*. Finally,*

$$I^{(2)} = 2I^{(1)}.$$

Proof. For the duration of this proof, we introduce the notation

$$\mu_D^{\mathrm{sc}} := \rho_{\mathrm{sc}} \boxplus \mu_D.$$

We now restrict ourselves to $\beta = 1$, dropping β from all notations, until the last section of the proof when we show $I^{(2)} = 2I^{(1)}$. It can be checked directly from the definition that that, for any compactly supported measure ν and any $\mathscr{M} \geqslant \mathtt{r}(\nu)$,

$$\frac{\partial}{\partial \theta} J(\nu, \theta, \mathscr{M}) = \begin{cases} R_\nu(2\theta) & \text{if } 0 \leqslant 2\theta \leqslant G_\nu(\mathscr{M}), \\ \mathscr{M} - \frac{1}{2\theta} & \text{if } 2\theta > G_\nu(\mathscr{M}). \end{cases}$$

Notice that this is a continuous function of θ. Furthermore, it is known [102, Lemma 6.1] that

$$G_{\mu_D^{\mathrm{sc}}}(\mathtt{r}(\mu_D^{\mathrm{sc}})) \leqslant \min(G_{\mu_D}(\mathtt{r}(\mu_D)), G_{\rho_{\mathrm{sc}}}(\mathtt{r}(\rho_{\mathrm{sc}}))) = \min(G_{\mu_D}(\mathtt{r}(\mu_D)), 1).$$

Since G_ν is decreasing on $(\mathtt{r}(\nu), +\infty)$, there are three (or two) phases of θ values:

$$\frac{\partial}{\partial\theta} I(x,\theta) = \begin{cases} 0 & \text{if } 0 \leqslant 2\theta \leqslant G_{\mu_D^{\mathrm{sc}}}(x), \\ x - 2\theta - K_{\mu_D}(2\theta) & \text{if } G_{\mu_D^{\mathrm{sc}}}(x) \leqslant 2\theta \leqslant G_{\mu_D}(\mathtt{r}(\mu_D)), \\ x - 2\theta - \mathtt{r}(\mu_D) & \text{if } 2\theta \geqslant G_{\mu_D}(\mathtt{r}(\mu_D)), \end{cases}$$

where the third case disappears if $G_{\mu_D}(\mathtt{r}(\mu_D)) = +\infty$ and the second case disappears if $x = \mathtt{r}(\mu_D^{\mathrm{sc}})$ and $G_{\mu_D^{\mathrm{sc}}}(\mathtt{r}(\mu_D^{\mathrm{sc}})) = G_{\mu_D}(\mathtt{r}(\mu_D))$. Notice that this is a continuous function of $\theta \geqslant 0$, and that, if $G_{\mu_D^{\mathrm{sc}}}(x) \leqslant 2\theta \leqslant G_{\mu_D^{\mathrm{sc}}}(\mathtt{r}(\mu_D^{\mathrm{sc}}))$, we can in fact write

$$\frac{\partial}{\partial\theta} I(x,\theta) = x - K_{\mu_D^{\mathrm{sc}}}(2\theta).$$

In general we have

$$G_{\mu_D}(\mathtt{r}(\mu_D)) \in [G_{\mu_D^{\mathrm{sc}}}(\mathtt{r}(\mu_D^{\mathrm{sc}})), +\infty].$$

For the purposes of our analysis, the endpoints of this interval are degenerate cases, and will be handled separately at the end. For now, assume that

$$G_{\mu_D}(\mathtt{r}(\mu_D)) \in (G_{\mu_D^{\mathrm{sc}}}(\mathtt{r}(\mu_D^{\mathrm{sc}})), +\infty).$$

Then $\partial_\theta I(x,\theta)$ has three non-degenerate piecewise sections, and $x_c < \infty$, where we recall the threshold

$$x_c = \begin{cases} G_{\mu_D}(\mathtt{r}(\mu_D)) + \mathtt{r}(\mu_D) & \text{if } G_{\mu_D}(\mathtt{r}(\mu_D)) < \infty, \\ +\infty & \text{otherwise.} \end{cases}$$

In the course of the casework, we will show that $x_c > \mathtt{r}(\mu_D^{\mathrm{sc}})$ in this nondegenerate regime.

- **Case 1 ($x < x_c$):** First we study θ in $(\frac{1}{2}G_{\mu_D^{\mathrm{sc}}}(x), \frac{1}{2}G_{\mu_D}(\mathtt{r}(\mu_D)))$ and write the function $\partial_\theta I(x,\theta)$ as

$$\theta \mapsto f_x(\theta) = x - 2\theta - K_{\mu_D}(2\theta)$$

defined on this interval. We have

$$\begin{aligned} f_x''(\theta) &= \frac{4G''_{\mu_D}(K_{\mu_D}(2\theta))}{\left(G'_{\mu_D}(K_{\mu_D}(2\theta))\right)^3} \\ &= -8\left(\int \frac{\mu_D(\mathrm{d}t)}{(K_{\mu_D}(2\theta)-t)^2}\right)^{-3}\left(\int \frac{\mu_D(\mathrm{d}t)}{(K_{\mu_D}(2\theta)-t)^3}\right) < 0 \end{aligned}$$

since $2\theta < G_{\mu_D}(\mathtt{r}(\mu_D))$, so that $K_{\mu_D}(2\theta) > \mathtt{r}(\mu_D)$ and $\int \frac{\mu_D(\mathrm{d}t)}{(K_{\mu_D}(2\theta)-t)^i} > 0$ for $i = 2, 3$. Thus f_x is strictly concave.

Let us find out where it is maximized. Since

$$f_x'(\theta) = 2\left(\frac{1}{\int \frac{\mu_D(\mathrm{d}t)}{(K_{\mu_D}(2\theta)-t)^2}} - 1\right),$$

we can rearrange

$$\begin{aligned} \mathtt{r}(\mu_D^{\mathrm{sc}}) &= K_{\mu_D^{\mathrm{sc}}}(G_{\mu_D^{\mathrm{sc}}}(\mathtt{r}(\mu_D^{\mathrm{sc}}))) \\ &= R_{\rho_{\mathrm{sc}}}(G_{\mu_D^{\mathrm{sc}}}(\mathtt{r}(\mu_D^{\mathrm{sc}}))) + K_{\mu_D}(G_{\mu_D^{\mathrm{sc}}}(\mathtt{r}(\mu_D^{\mathrm{sc}}))) \\ &= G_{\mu_D^{\mathrm{sc}}}(\mathtt{r}(\mu_D^{\mathrm{sc}})) + K_{\mu_D}(G_{\mu_D^{\mathrm{sc}}}(\mathtt{r}(\mu_D^{\mathrm{sc}}))) \end{aligned}$$

to obtain

$$f_x'\left(\frac{1}{2}G_{\mu_D^{\mathrm{sc}}}(\mathtt{r}(\mu_D^{\mathrm{sc}}))\right) = 2\left(\frac{1}{\int \frac{\mu_D(\mathrm{d}t)}{(\mathtt{r}(\mu_D^{\mathrm{sc}})-G_{\mu_D^{\mathrm{sc}}}(\mathtt{r}(\mu_D^{\mathrm{sc}}))-t)^2}} - 1\right).$$

But it is known that

$$\int \frac{\mu_D(\mathrm{d}t)}{(\mathtt{r}(\mu_D^{\mathrm{sc}}) - G_{\mu_D^{\mathrm{sc}}}(\mathtt{r}(\mu_D^{\mathrm{sc}})) - t)^2} = 1.$$

Indeed, using the notation and results of [64, Proposition 2.1] (although the ideas date back to [49]), the above statement is equivalent to the statement $v_{1,\mu_D}(F_{1,\mu_D}(\mathtt{r}(\mu_D^{\mathrm{sc}}))) = 0$. But

F_{1,μ_D} maps into $\overline{\{u+iv \in \mathbb{C}^+ : v > v_{1,\mu_D}(u)\}}$, and here

$$F_{1,\mu_D}(\mathtt{r}(\mu_D^{\mathrm{sc}})) = \mathtt{r}(\mu_D^{\mathrm{sc}}) - G_{\mu_D^{\mathrm{sc}}}(\mathtt{r}(\mu_D^{\mathrm{sc}}))$$

is real; furthermore $v_{1,\mu_D}(u)$ is a continuous function [49] of the real parameter u. These conditions force $v_{1,\mu_D}(F_{1,\mu_D}(\mathtt{r}(\mu_D^{\mathrm{sc}}))) = 0$. But this means that

$$f_x'\left(\frac{1}{2}G_{\mu_D^{\mathrm{sc}}}(\mathtt{r}(\mu_D^{\mathrm{sc}}))\right) = 0.$$

Thus we have shown that $f_x(\theta) = \partial_\theta I(x,\theta)$ is a strictly concave function on the open interval $(\frac{1}{2}G_{\mu_D^{\mathrm{sc}}}(x), \frac{1}{2}G_{\mu_D}(\mathtt{r}(\mu_D))$, taking a unique maximum value (which can be computed to be $x - \mathtt{r}(\mu_D^{\mathrm{sc}})$) at the point $\theta = \frac{1}{2}G_{\mu_D^{\mathrm{sc}}}(\mathtt{r}(\mu_D^{\mathrm{sc}}))$. Its value at the left endpoint of the interval is 0, and its value at the right endpoint of the interval is $x - x_c < 0$. In particular, since f_x is decreasing on $(\frac{1}{2}G_{\mu_D^{\mathrm{sc}}}(\mathtt{r}(\mu_D^{\mathrm{sc}})), \frac{1}{2}G_{\mu_D}(\mathtt{r}(\mu_D)))$, taking the value $x - \mathtt{r}(\mu_D^{\mathrm{sc}})$ on the left endpoint and value $x - x_c$ on the right endpoint, we have $x_c > \mathtt{r}(\mu_D^{\mathrm{sc}})$ as claimed.

Now if $\theta \geqslant \frac{1}{2}G_{\mu_D}(\mathtt{r}(\mu_D))$, then

$$\partial_\theta I(x,\theta) = x - 2\theta - \mathtt{r}(\mu_D) < x_c - G_{\mu_D}(\mathtt{r}(\mu_D)) - \mathtt{r}(\mu_D) = 0.$$

There are two subcases here:

- **Subcase a ($x = \mathtt{r}(\mu_D^{\mathrm{sc}})$):** Here, $f_x(\theta) = \partial_\theta I(x,\theta)$ takes maximum value $x - \mathtt{r}(\mu_D^{\mathrm{sc}}) = 0$ on the interval
$$\left(\frac{1}{2}G_{\mu_D^{\mathrm{sc}}}(\mathtt{r}(\mu_D^{\mathrm{sc}})), \frac{1}{2}G_{\mu_D}(\mathtt{r}(\mu_D))\right),$$
and is negative on the interval $[\frac{1}{2}G_{\mu_D}(\mathtt{r}(\mu_D)), +\infty)$. Thus I vanishes at $\mathtt{r}(\mu_D^{\mathrm{sc}})$.
- **Subcase b ($x > \mathtt{r}(\mu_D^{\mathrm{sc}})$):** Here, the value of the function f_x at $\theta = \frac{1}{2}G_{\mu_D^{\mathrm{sc}}}(\mathtt{r}(\mu_D^{\mathrm{sc}}))$ is

$x - \mathtt{r}(\mu_D^{\text{sc}}) > 0$. Thus it vanishes at a unique point θ_x in the interval

$$(\frac{1}{2}G_{\mu_D^{\text{sc}}}(\mathtt{r}(\mu_D^{\text{sc}})), \frac{1}{2}G_{\mu_D}(\mathtt{r}(\mu_D))).$$

For such values of x, then, $I(x,\theta)$ vanishes for $\theta \in [0, \frac{1}{2}G_{\mu_D^{\text{sc}}}(x)]$; strictly increases for $\theta \in (\frac{1}{2}G_{\mu_D^{\text{sc}}}, \theta_x)$; and strictly decreases for $\theta \in (\theta_x, +\infty)$. In particular $I(x) > 0$ for such x values.

– **Case 2 ($x \geqslant x_c$):** Here we can explicitly write

$$\theta_x = \frac{1}{2}(x - \mathtt{r}(\mu_D)). \tag{5.6.1}$$

The function f_x defined above is still strictly concave on its domain and still vanishes at the left endpoint of this domain, but now its value at the right endpoint is nonnegative; thus $I(x,\theta)$ is strictly increasing for $\theta \in (\frac{1}{2}G_{\mu_D^{\text{sc}}}(x), \frac{1}{2}G_{\mu_D}(\mathtt{r}(\mu_D)))$. A simple analysis of $\partial_\theta I(x,\theta)$ for $\theta \geqslant \frac{1}{2}G_{\mu_D}(\mathtt{r}(\mu_D))$ shows that θ_x as defined above is, as claimed, the unique θ value that maximizes $I(x,\theta)$, and $I(x) > 0$.

In particular notice that

$$\theta_{x_c} = \frac{1}{2}(x_c - \mathtt{r}(\mu_D)) = \frac{1}{2}G_{\mu_D}(\mathtt{r}(\mu_D)).$$

It remains only to show that $x_1 \neq x_2 \implies \theta_{x_1} \neq \theta_{x_2}$. If $x_1 < x_c \leqslant x_2$, then θ_{x_1} and θ_{x_2} as constructed above lie in disjoint intervals, so cannot be equal; and if $x_c \leqslant x_1, x_2$ then we can see $\theta_{x_1} \neq \theta_{x_2}$ from our explicit formula (5.6.1). Thus we only need consider $x_1 < x_2 < x_c$. If $x_1 = \mathtt{r}(\mu_D^{\text{sc}})$, then $\theta_{x_1} = \frac{1}{2}G_{\mu_D^{\text{sc}}}(\mathtt{r}(\mu_D^{\text{sc}})) < \theta_{x_2}$ by construction; thus we can assume $\mathtt{r}(\mu_D^{\text{sc}}) < x_1 < x_2 < x_c$. But then θ_{x_1} and θ_{x_2} are defined on the common interval $(\frac{1}{2}G_{\mu_D^{\text{sc}}}(\mathtt{r}(\mu_D^{\text{sc}})), \frac{1}{2}G_{\mu_D}(\mathtt{r}(\mu_D)))$ as the unique points satisfying

$$2\theta_{x_1} + K_{\mu_D}(2\theta_{x_1}) = x_1 \neq x_2 = 2\theta_{x_2} + K_{\mu_D}(2\theta_{x_2}).$$

Thus we must have $\theta_{x_1} \neq \theta_{x_2}$.

Now we explain the necessary adjustments in the degenerate cases.

– **Degenerate Case 1 ($G_{\mu_D}(\mathtt{r}(\mu_D)) = G_{\mu_D^{\mathrm{sc}}}(\mathtt{r}(\mu_D^{\mathrm{sc}}))$):**

The proof of [102, Lemma 6.1] shows that $\omega(\mathtt{r}(\mu_D^{\mathrm{sc}})) \geqslant \mathtt{r}(\mu_D)$, where ω is defined (see [64, Proposition 2.1]) as $\omega(z) = z - G_{\mu_D^{\mathrm{sc}}}(z)$; hence

$$\begin{aligned} x_c &= \mathtt{r}(\mu_D) + G_{\mu_D}(\mathtt{r}(\mu_D)) = \mathtt{r}(\mu_D) + G_{\mu_D^{\mathrm{sc}}}(\mathtt{r}(\mu_D^{\mathrm{sc}})) \\ &= \mathtt{r}(\mu_D) + \mathtt{r}(\mu_D^{\mathrm{sc}}) - \omega(\mathtt{r}(\mu_D^{\mathrm{sc}})) \leqslant \mathtt{r}(\mu_D^{\mathrm{sc}}) \end{aligned}$$

and all x are "at least critical."

– **Degenerate Subcase a ($x = \mathtt{r}(\mu_D^{\mathrm{sc}})$):** Then we only have

$$\begin{aligned} &\partial_\theta I(\mathtt{r}(\mu_D^{\mathrm{sc}}), \theta) \\ &= \begin{cases} 0 & \text{if } 0 \leqslant 2\theta \leqslant G_{\mu_D}(\mathtt{r}(\mu_D)) \\ \mathtt{r}(\mu_D^{\mathrm{sc}}) - 2\theta - \mathtt{r}(\mu_D) & \text{if } 2\theta \geqslant G_{\mu_D}(\mathtt{r}(\mu_D)). \end{cases} \end{aligned}$$

From the first line of this display and from the equality

$$G_{\mu_D}(\mathtt{r}(\mu_D)) = G_{\mu_D^{\mathrm{sc}}}(\mathtt{r}(\mu_D^{\mathrm{sc}}))$$

we have

$$\begin{aligned} 0 &= R_{\mu_D^{\mathrm{sc}}}(G_{\mu_D^{\mathrm{sc}}}(\mathtt{r}(\mu_D^{\mathrm{sc}}))) - G_{\mu_D}(\mathtt{r}(\mu_D)) - R_{\mu_D}(G_{\mu_D}(\mathtt{r}(\mu_D))) \\ &= \mathtt{r}(\mu_D^{\mathrm{sc}}) - G_{\mu_D}(\mathtt{r}(\mu_D)) - \mathtt{r}(\mu_D). \end{aligned} \tag{5.6.2}$$

On the one hand, (5.6.2) tells us that

$$x_c = \mathtt{r}(\mu_D^{\mathrm{sc}})$$

so that by convention

$$\theta_{x_c} = \frac{1}{2} G_{\mu_D^{\mathrm{sc}}}(\mathtt{r}(\mu_D^{\mathrm{sc}})) = \frac{1}{2} G_{\mu_D}(\mathtt{r}(\mu_D))$$

as claimed. On the other hand, if $2\theta \geqslant G_{\mu_D}(\mathtt{r}(\mu_D))$ then (5.6.2) tells us that

$$\partial_\theta I(\mathtt{r}(\mu_D^{\mathrm{sc}}, \theta)) = \mathtt{r}(\mu_D^{\mathrm{sc}}) - 2\theta - \mathtt{r}(\mu_D) \leqslant \mathtt{r}(\mu_D^{\mathrm{sc}}) - G_{\mu_D}(\mathtt{r}(\mu_D)) - \mathtt{r}(\mu_D) = 0.$$

So $\partial_\theta I(\mathtt{r}(\mu_D^{\mathrm{sc}})) \leqslant 0$ for all θ and $I(\mathtt{r}(\mu_D^{\mathrm{sc}})) = 0$ as claimed.

– **Degenerate Subcase b ($x > \mathtt{r}(\mu_D^{\mathrm{sc}})$):** Then f_x as above is defined and strictly concave on a nondegenerate interval; it vanishes at the left endpoint of this interval; it takes a positive maximum (namely $x - \mathtt{r}(\mu_D^{\mathrm{sc}})$) at the right endpoint of this interval. Thus the analysis of Case 2 above holds to show that θ_x is given by Equation (5.6.1).

The argument above for injectivity goes through, since Equation (5.6.1) works for all x values.

– **Degenerate Case 2 ($G_{\mu_D}(\mathtt{r}(\mu_D)) = +\infty$):** Here $x_c = +\infty$, and all x values are subcritical. The function f_x from Case 1 is then defined and strictly concave on the interval $(\frac{1}{2} G_{\mu_D^{\mathrm{sc}}}(x), +\infty)$. It has a unique maximum at $\frac{1}{2} G_{\mu_D^{\mathrm{sc}}}(\mathtt{r}(\mu_D^{\mathrm{sc}}))$, where its value is positive; and strict concavity tells us $\lim_{\theta \to +\infty} f_x(\theta) = -\infty$; thus f_x still has a unique zero on its domain, which we still call θ_x. The argument above for injectivity goes through.

Now we reintroduce β to all notations and show $I^{(2)} = 2I^{(1)}$. If $x < x_c$, then $\theta_x^{(\beta)}$ is defined implicitly by

$$\frac{2}{\beta}\theta_x^{(\beta)} + K_{\mu_D}\left(\frac{2}{\beta}\theta_x^{(\beta)}\right) = x \quad \text{subject to} \quad \frac{2}{\beta}\theta_x^{(\beta)} \in \left(G_{\mu_D^{\mathrm{sc}}}(\mathtt{r}(\mu_D^{\mathrm{sc}})), G_{\mu_D}(\mathtt{r}(\mu_D))\right).$$

(We showed this for $\beta = 1$, and the extension to $\beta = 2$ is similar.) If $x \geqslant x_c$, then we have

$$\frac{2}{\beta}\theta_x^{(\beta)} = x - \mathtt{r}(\mu_D).$$

Notice that $\frac{2}{\beta}\theta_x^{(\beta)}$ is independent of β. But then the definition (5.2.2) gives us

$$J^{(2)}(\nu, \theta_x, \mathscr{M}) = 2J^{(1)}(\nu, \theta_x, \mathscr{M})$$

from which the claim follows. □

Proposition 5.6.2. *The function $I^{(\beta)}(\cdot)$ is a good rate function.*

Proof. First, for any compactly-supported measure μ and any $\lambda \geqslant \mathtt{r}(\mu)$, we have $J^{(\beta)}(\mu, 0, \lambda) = 0$; hence $I^{(\beta)}(x)$ is nonnegative.

For every fixed θ, dominated convergence tells us that $J^{(\beta)}(\rho_{\mathrm{sc}} \boxplus \mu_D, \theta, x)$ is a continuous function of $x > \mathtt{r}(\rho_{\mathrm{sc}})$; hence $I^{(\beta)}(\cdot)$ is lower semi-continuous at such x values. It is also lower semi-continuous for $x < \mathtt{r}(\rho_{\mathrm{sc}} \boxplus \mu_D)$, where its value is infinite. Finally, since $I^{(\beta)}(\cdot)$ is nonnegative and vanishes at $\mathtt{r}(\rho_{\mathrm{sc}} \boxplus \mu_D)$, it is also lower semi-continuous there.

Hence $I^{(\beta)}(\cdot)$ is a rate function. But since $I^{(\beta)}(\cdot)$ is the rate function for a weak LDP of an exponentially tight family, it is classical (see, e.g., [69, Lemma 1.2.18]) that $I^{(\beta)}$ is in fact good. □

Appendices

Appendix A: Extensions to products of determinants

In this section, we are interested in expectations of products of determinants like

$$\mathbb{E}\left[\prod_{i=1}^{\ell} |\det(H_N^{(i)})|\right],$$

where ℓ is independent of N. In the landscape complexity program, these asymptotics help understand the ℓth moment of the number of critical points of some high-dimensional random function. Everything essentially is the same as in the case $\ell = 1$, and we obtain leading-order determinant asymptotics consistent with

$$\mathbb{E}\left[\prod_{i=1}^{\ell} |\det(H_N^{(i)})|\right] \approx \prod_{i=1}^{\ell} \mathbb{E}[|\det(H_N^{(i)})|], \tag{A.1}$$

on exponential scale in N. This is true *no matter the correlation structure between the* $H_N^{(i)}$*'s*, which is perhaps surprising at first glance. However, note that (A.1) should hold at "both ends of the correlation spectrum," so to speak: On the one hand, it holds with exact equality if the $H_N^{(i)}$'s are independent; on the other hand, if we believe in concentration then (A.1) is very plausible when the $H_N^{(i)}$'s are the same as each other.

However, (A.1) does require higher moment assumptions: for example, it holds when the $H_N^{(i)}$ are Wigner matrices with $2\ell + \varepsilon$ moments, which is consistent with the case $\ell = 1$. This is almost

optimal, because if all the $H_N^{(i)}$'s are the same Wigner matrix, then the left-hand side of (A.1) is infinite unless the entries have at least 2ℓ moments; see the remark just before Corollary 2.1.3 and the proof thereof in Section 2.3.3, which generalize to $\ell > 1$.

These new moment assumptions are encapsulated in the following generalization of Assumption (C) (notice that (C^ℓ) with $\ell = 1$ is the same as (C)).

(C^ℓ) In addition to the Wegner assumption (2.1.5), we require

$$\lim_{N\to\infty} \frac{1}{N} \log \mathbb{E}\left[\prod_{i=1}^{N} (1 + |\lambda_i| \mathbb{1}_{|\lambda_i| > e^{N^\varepsilon}})^\ell\right] = 0 \tag{A.2}$$

for every $\varepsilon > 0$ and

$$\limsup_{N\to\infty} \frac{1}{N \log N} \log \mathbb{E}[|\det(H_N)|^{\ell(1+\delta)}] < \infty \quad \text{for each } i, \tag{A.3}$$

for all sufficiently small $\delta > 0$.

Here is the analogue of Theorem 2.1.1.

Theorem A.1. ***(Convexity-preserving functionals)*** *Fix $\ell \in \mathbb{N}$, and consider ℓ collections $(X^{(i)})_{i=1}^{\ell}$ each consisting of M arbitrary independent entries. The collections can have any correlation structure with respect to each other. Consider matrices $H_N^{(i)} = \Phi^{(i)}(X^{(i)})$ that each satisfy Assumptions (I), (M), (E), (C^ℓ), and (S) with reference measures $\mu_N^{(i)}$. Then*

$$\lim_{N\to\infty} \left(\frac{1}{N} \log \mathbb{E}\left[\prod_{i=1}^{\ell} |\det(H_N^{(i)})|\right] - \sum_{i=1}^{\ell} \int_{\mathbb{R}} \log|\lambda| \mu_N^{(i)}(\mathrm{d}\lambda) \right) = 0. \tag{A.4}$$

Proof. We refer freely to objects from the proof of Theorem 2.1.1, adding a parenthetical index (i) to indicate their corresponding matrix. For example,

$$\mathcal{E}_{\mathrm{ss}}^{(i)} = \{d_{\mathrm{KS}}(\hat{\mu}_{\Phi^{(i)}(X^{(i)})}, \hat{\mu}_{\Phi^{(i)}(X_{\mathrm{cut}}^{(i)})}) \leqslant N^{-\kappa}\}$$

and so on. The main estimate in the upper bound is

$$\begin{aligned}
&\frac{1}{N}\log\mathbb{E}\left[\prod_{i=1}^{\ell}|\det(H_N^{(i)})|\mathbb{1}_{\mathcal{E}_{\mathrm{ss}}^{(i)}}\mathbb{1}_{\mathcal{E}_{\mathrm{conc}}^{(i)}}\right]\\
&\leqslant \frac{1}{N}\log\mathbb{E}\left[e^{\sum_{i=1}^{\ell}N\int\log_\eta^K(\lambda)\hat{\mu}_{\Phi^{(i)}(X^{(i)})}(\mathrm{d}\lambda)}\prod_{i=1}^{\ell}\left(\prod_{j=1}^{N}(1+|\lambda_j^{(i)}|\mathbb{1}_{|\lambda_j^{(i)}|>K})\right)\mathbb{1}_{\mathcal{E}_{\mathrm{ss}}^{(i)}}\mathbb{1}_{\mathcal{E}_{\mathrm{conc}}^{(i)}}\right]\\
&\leqslant \ell(2\varepsilon_1(N)+t)+\sum_{i=1}^{\ell}\int_{\mathbb{R}}\log_\eta^K(\lambda)\mu_N^{(i)}(\mathrm{d}\lambda)+\sum_{i=1}^{\ell}\frac{1}{\ell N}\log\mathbb{E}\left[\prod_{j=1}^{N}(1+|\lambda_j^{(i)}|\mathbb{1}_{|\lambda_j^{(i)}|>K})^{\ell}\right]
\end{aligned}$$

where we use Hölder's inequality in the last line. Using the assumption (A.2) and arguments as in the one-determinant case, we use this to find

$$\limsup_{N\to\infty}\left(\frac{1}{N}\log\mathbb{E}\left[\prod_{i=1}^{\ell}|\det(H_N^{(i)})|\mathbb{1}_{\mathcal{E}_{\mathrm{ss}}^{(i)}}\mathbb{1}_{\mathcal{E}_{\mathrm{conc}}^{(i)}}\right]-\sum_{i=1}^{\ell}\int\log|\lambda|\mu_N^{(i)}(\mathrm{d}\lambda)\right)\leqslant 0.$$

To conclude the upper bound, write $\mathcal{E}^{(i)}=\mathcal{E}_{\mathrm{ss}}^{(i)}\cap\mathcal{E}_{\mathrm{conc}}^{(i)}$. We expand

$$\mathbb{E}\left[\prod_{i=1}^{\ell}|\det(H_N^{(i)})|(\mathbb{1}_{\mathcal{E}^{(i)}}+\mathbb{1}_{(\mathcal{E}^{(i)})^c}),\right]$$

as a sum over 2^ℓ terms, each of which has a product of ℓ determinants and a product of ℓ indicators. We just studied the term with every indicator on $\mathcal{E}^{(i)}$, and now claim that any term with at least one indicator on the complement of $\mathcal{E}^{(i)}$ does not contribute. Indeed, suppose for concreteness that the indicator $\mathbb{1}_{(\mathcal{E}^{(1)})^c}$ appears; then the term is bounded above by

$$\mathbb{E}\left[\prod_{i=1}^{\ell}|\det(H_N^{(i)})|\mathbb{1}_{(\mathcal{E}^{(1)})^c}\right]\leqslant\left(\prod_{i=1}^{\ell}\mathbb{E}[|\det(H_N^{(i)})|^{\ell(1+\delta)}]^{\frac{1}{\ell(1+\delta)}}\right)\mathbb{P}((\mathcal{E}^{(1)})^c)^{\frac{\delta}{1+\delta}}$$

according to Hölder's. Using the new assumption (A.3), we proceed as in the proof of Lemma 2.2.4 to complete the proof of the upper bound.

The lower bound is easier to generalize; by following the proof of Lemma 2.2.5, we find

$$\frac{1}{N}\log\mathbb{E}\left[\prod_{i=1}^{\ell} e^{N\int(\log|\lambda|-\log_\eta(\lambda))\hat{\mu}_{\Phi^{(i)}(X^{(i)})}(\mathrm{d}\lambda)}\mathbb{1}_{\mathcal{E}^{(i)}_{\mathrm{gap}}}\mathbb{1}_{\mathcal{E}^{(i)}_{\mathrm{ss}}}\mathbb{1}_{\mathcal{E}^{(i)}_{\mathrm{conc}}}\right] \geqslant -\widetilde{\varepsilon}_2(N)$$

with

$$\widetilde{\varepsilon}_2(N) = \ell\left(\frac{p_b}{2}\log(1+e^{2N^\varepsilon}\eta^2)+\frac{\eta^2}{2w_b^2}\right) - \frac{1}{N}\log\mathbb{P}\left(\bigcap_{i=1}^{\ell}\mathcal{E}^{(i)}_{\mathrm{gap}},\mathcal{E}^{(i)}_{\mathrm{ss}},\mathcal{E}^{(i)}_{\mathrm{conc}},\mathcal{E}^{(i)}_{b}\right),$$

which tends to zero since each of the events $\mathcal{E}^{(i)}_{\dots}$ has probability tending to one. □

Here is the analogue of Theorem 2.1.2.

Theorem A.2. ***(Concentrated inputs)*** *Fix $\ell \in \mathbb{N}$, and suppose that each of the matrices $(H_N^{(i)})_{i=1}^{\ell}$ satisfies the assumptions of the one-determinant Theorem 2.1.2 with measures $\mu_N^{(i)}$. Then (A.4) holds.*

Proof. For the upper bound, we mimic the proof of the one-determinant case, using Hölder's to obtain terms of the form $\mathbb{E}[e^{\ell N\int\log_\eta(\lambda)(\hat{\mu}_{H_N^{(i)}}-\mathbb{E}[\hat{\mu}_{H_N^{(i)}}])(\mathrm{d}\lambda)}]^{1/\ell}$; we simply absorb this ℓ into the Lipschitz constant of $\log_\eta$. The lower bound is generalized as in the convexity-preserving-functional case, Theorem A.1. □

We give two corollaries.

Corollary A.3. ***(Products of ℓ Wigner matrices with $2\ell+\varepsilon$ moments)*** *Fix some $\varepsilon > 0$, and let $(\mu^{(i)})_{i=1}^{\ell}$ be a collection of centered probability measures on $\mathbb{R}$ with $2\ell+\varepsilon$ finite moments and unit variance. Let $W_N^{(i)}$ be a real symmetric Wigner matrix corresponding to $\mu^{(i)}$. Then for every collection $(E^{(i)})_{i=1}^{\ell}$ we have*

$$\lim_{N\to\infty}\frac{1}{N}\log\mathbb{E}\left[\prod_{i=1}^{\ell}|\det(W_N^{(i)}-E^{(i)})|\right] = \sum_{i=1}^{\ell}\int_{\mathbb{R}}\log|\lambda-E^{(i)}|\rho_{sc}(\lambda)\,\mathrm{d}\lambda.$$

Proof. We use Theorem A.1, verifying its assumptions as in the case of one Wigner matrix. We need $2\ell+\varepsilon$ moments in the verification of (A.2) and (A.3) as follows: Dropping $(\cdots)^{(i)}$ from the

notation and arguing as in the one-point case, we find

$$\mathbb{E}\left[\prod_{j=1}^{N}(1+|\lambda_j|\mathbb{1}_{|\lambda_j|>e^{N^\varepsilon}})^\ell\right] \leqslant \mathbb{E}\left[\prod_{j=1}^{N}\left(1+10\sum_{k=1}^{N}|B_{jk}|\right)^\ell\right],$$

where B is the matrix of tails, which has independent entries up to symmetry. When we expand and factor the right-hand side, entries of B now appear with power at most 2ℓ (instead of 2 before). Similarly, to verify (A.3) we mimic the original notation and find

$$|\det(W_N+E)|^{\ell(1+\delta)} \leqslant (N!)^{\ell(1+\delta)}\frac{\sum_\sigma X_\sigma^{\ell(1+\delta)}}{N!}$$

and $\mathbb{E}[X_\sigma^{\ell(1+\delta)}]$ is finite because we have finite $2\ell(1+\delta)$ moments. □

Corollary A.4. ***(Products of ℓ non-invariant Gaussian matrices)*** *If $(H_N^{(i)})_{i=1}^\ell$ are Gaussian matrices with a (co)variance profile satisfying the requirements of Corollary 2.1.8.B, or block-diagonal Gaussian matrices satisfying the requirements of Corollary 2.1.9 – or a mixture of both – and $\mu_N^{(i)}$ are the corresponding MDE measures, then*

$$\lim_{N\to\infty}\left(\frac{1}{N}\log\mathbb{E}\left[\prod_{i=1}^{\ell}|\det(H_N^{(i)})|\right]-\sum_{i=1}^{\ell}\int_{\mathbb{R}}\log|\lambda|\mu_N^{(i)}(\lambda)\,\mathrm{d}\lambda\right)=0.$$

Appendix B: Edge behavior of general free convolutions with semicircle

Recall the notation of Section 3.5.2 for the free convolution of a measure μ_D with the semi-circular distribution of variance t, and for its left edge:

$$\begin{aligned}\mu_t &= \rho_{\mathrm{sc},t}\boxplus\mu_D,\\ \ell_t &= \mathtt{l}(\mu_t)\\ m_t(z) &= \int_{\mathbb{R}}\frac{\mu_t(\mathrm{d}\lambda)}{\lambda-z}.\end{aligned}$$

Recall also the notation $\mu_t(\cdot)$ for the density of μ_t.

The following result might be of independent interest.

Proposition B.1. *Any free convolution with semicircle decays at least as fast as a square root at the extremal edges, in the following sense: For any compactly supported measure μ_D and any t, there exist $c, \varepsilon > 0$ such that*

$$\mu_t(x) \leqslant c\sqrt{x - \ell_t} \qquad \text{for } x \in [\ell_t, \ell_t + \varepsilon].$$

On the one hand, square-root decay is of course achieved if $\mu_D = \delta_0$ (so that the free convolution is semicircle). On the other hand, Lee and Schnelli have presented a family of examples where decay at the edge is strictly faster than square root [114, Lemma 2.7]. Thus the power in this result cannot be improved. We also mention works providing sufficient conditions on μ_D for a matching lower bound, i.e., to ensure that extremal-edge decay is *exactly* square root, such as [21, Theorem 2.2] (which actually considers free convolution between two Jacobi measures, not our special case when one of them is semicircular).

This result also complements [49, Corollary 5] of Biane, which shows that decay near *any* edge is at least as fast a *cube* root. As Biane shows, this is in fact the correct power at a cusp when two connected components of the support merge. Thus the "extremal" restriction in our result is necessary.

Proof. We adapt arguments of Biane [49] as follows. Biane considers the function $v_t(u) : \mathbb{R} \to [0, \infty)$ defined by

$$v_t(u) = \inf\left\{v \geqslant 0 : \int_{\mathbb{R}} \frac{\mu_D(\mathrm{d}x)}{(u-x)^2 + v^2} \leqslant \frac{1}{t}\right\}$$

and the open set $U_t = \{u \in \mathbb{R} : v_t(u) > 0\}$, then defines a certain homeomorphism $\psi_t : \mathbb{R} \to \mathbb{R}$ (whose exact form is not important to us now) and proves that

$$\mu_t(\psi_t(u)) = \frac{v_t(u)}{\pi t}$$

for all $u \in \mathbb{R}$.

On the one hand, by [49, Corollary 3], we have

$$u_t := \psi_t^{-1}(\ell_t) = \ell_t + t m_t(\ell_t).$$

This is at most $\mathtt{l}(\mu_D)$ by (3.5.35), and in fact the inequality is strict since $m_t(\ell_t) > 0$:

$$u_t < \mathtt{l}(\mu_D). \tag{B.1}$$

On the other hand, let x be such that $\mu_t(x) > 0$. Then $x = \psi_t(u)$ for some $u \in U_t$, and adapting the proofs of [49, Proposition 4, Lemma 5] we obtain

$$|\mu_t(x)\mu_t'(x)| \leqslant \frac{|v_t(u)v_t'(u)|}{\pi^2 t^2 \psi_t'(u)} \leqslant \frac{|v_t'(u)|}{2\pi^2 t |v_t(u)|(1+v_t'(u)^2)} \leqslant \frac{1}{2\pi^2 t} \cdot \frac{1}{|v_t(u)v_t'(u)|}. \tag{B.2}$$

But the proof of [49, Lemma 5] shows that

$$v_t(u)v_t'(u) = \frac{\int_{\mathbb{R}} \frac{(x-u)}{((u-x)^2+v_t(u)^2)} \mu_D(\mathrm{d}x)}{\int_{\mathbb{R}} \frac{1}{((u-x)^2+v_t(u)^2)} \mu_D(\mathrm{d}x)} \geqslant \mathtt{l}(\mu_D) - u.$$

For u in some $[u_t, u_t + \varepsilon]$ (corresponding via ψ_t to x in some $[\ell_t, \ell_t + \varepsilon']$), this lower bound is strictly positive by (B.1). By (B.2), this suffices. □

Appendix C: Computational details in large-deviations examples

We give the computational details for the example in Section 5.2.4. Using the equivalent [101, Theorem 6] formula

$$J^{(\beta)}(\nu, \theta, \mathscr{M}) = \theta R_\nu\left(\frac{2}{\beta}\theta\right) - \frac{\beta}{2}\int \log\left(1 + \frac{2}{\beta}\theta R_\nu\left(\frac{2}{\beta}\theta\right) - \frac{2}{\beta}\theta y\right)\nu(\mathrm{d}y),$$

valid if $0 \leqslant \frac{2}{\beta}\theta \leqslant G_\nu(\mathscr{M})$, and the constrained equation (5.2.6) implicitly defining $\theta_x^{(\beta)}$, one can see that

$$\begin{aligned} I^{(\beta)}(x) &= \frac{(\theta_x^{(\beta)})^2}{\beta} + \frac{\beta}{2}\int \log\left(x - \frac{2}{\beta}\theta_x^{(\beta)} - y\right)\mu_D(\mathrm{d}y) - \frac{\beta}{2}\int \log(x-y)(\rho_{\mathrm{sc}} \boxplus \mu_D)(\mathrm{d}y) \\ &= \frac{\beta}{4}\left(\frac{2}{\beta}\theta_x^{(\beta)}\right)^2 + \frac{\beta}{4}\log\left[\left(x - \frac{2}{\beta}\theta_x^{(\beta)}\right)^2 - a^2\right] - \frac{\beta}{2}\int \log(x-y)(\rho_{\mathrm{sc}} \boxplus \mu_D)(\mathrm{d}y) \end{aligned} \tag{C.1}$$

for $x \geqslant \mathtt{r}(\rho_{\mathrm{sc}} \boxplus \mu_D)$. We invert $K_{\rho_{\mathrm{sc}}\boxplus\mu_D}(y) = \frac{\sqrt{1+4a^2y^2}+2y^2+1}{2y}$ to obtain $G_{\rho_{\mathrm{sc}}\boxplus\mu_D}(y)$ for $y > \mathtt{r}(a)$, choosing branches according to the requirement that $G_{\rho_{\mathrm{sc}}\boxplus\mu_D}(y)$ be decreasing on $(\mathtt{r}(\rho_{\mathrm{sc}} \boxplus \mu_D), \infty)$; if $y > \mathtt{r}(a)$, this yields

$$G_{\rho_{\mathrm{sc}}\boxplus\mu_D}(y) = \frac{2}{3}\left[y - \sqrt{-3+3a^2+y^2}\sin\left(\frac{\pi}{3} - \frac{1}{3}\arctan(\mathtt{d}(y))\right)\right]$$

with $\mathtt{d}$ as in (5.2.9). In the limit $y \downarrow \mathtt{r}(a)$ we obtain

$$G_{\rho_{\mathrm{sc}}\boxplus\mu_D}(\mathtt{r}(a)) = \mathtt{c}(a)$$

with $\mathtt{c}$ as in (5.2.8). This gives us the bounds on the constrained problem (5.2.6); since $K_{\mu_D}(y) = \frac{\sqrt{1+4a^2y^2}+1}{2y}$, this has the solution

$$\frac{2}{\beta}\theta_x^{(\beta)} = \frac{2}{3}\left[x - \sqrt{-3+3a^2+x^2}\sin\left(\frac{1}{3}\arctan(\mathtt{d}(x))\right)\right].$$

if $x > \mathtt{r}(a)$. On the other hand, since $\rho_{\mathrm{sc}} \boxplus \mu_D$ decays at most like a cube root near its edges [49, Corollary 5], we can differentiate under the integral sign to obtain

$$\int \log(x-y)(\rho_{\mathrm{sc}} \boxplus \mu_D)(\mathrm{d}y) = \int_{\mathtt{r}(\rho_{\mathrm{sc}}\boxplus\mu_D)}^{x} G_{\rho_{\mathrm{sc}}\boxplus\mu_D}(t)\,\mathrm{d}t + \int \log(\mathtt{r}(\rho_{\mathrm{sc}} \boxplus \mu_D) - y)(\rho_{\mathrm{sc}} \boxplus \mu_D)(\mathrm{d}y).$$

We compute the second term on the right-hand side by setting $x = \mathtt{r}(\rho_{\mathrm{sc}} \boxplus \mu_D)$ in (C.1), since then $I^{(\beta)}(x) = 0$ and $\frac{2}{\beta}\theta_x^{(\beta)} = G_{\rho_{\mathrm{sc}} \boxplus \mu_D}(\mathtt{r}(\rho_{\mathrm{sc}} \boxplus \mu_D)) = \mathtt{c}(a)$; this yields

$$\int \log(\mathtt{r}(\rho_{\mathrm{sc}} \boxplus \mu_D) - y)(\rho_{\mathrm{sc}} \boxplus \mu_D)(\mathrm{d}y) = \frac{1}{2}(\mathtt{c}(a)^2 + \log((\mathtt{r}(a) - \mathtt{c}(a))^2 - a^2)).$$

This gives the stated formula for $I^{(\beta)}(x)$.

Bibliography

[1] A. Adhikari and J. Huang. Dyson Brownian motion for general β and potential at the edge. *Probab. Theory Related Fields*, 178(3-4):893–950, 2020.

[2] R. J. Adler and J. E. Taylor. *Random fields and geometry*. Springer Monographs in Mathematics. Springer, New York, 2007.

[3] E. Agliari, A. Barra, S. Bartolucci, A. Galluzzi, F. Guerra, and F. Moauro. Parallel processing in immune networks. *Phys. Rev. E*, 87:042701, Apr 2013.

[4] E. Agliari, A. Barra, A. Galluzzi, F. Guerra, and F. Moauro. Multitasking associative networks. *Phys. Rev. Lett.*, 109:268101, Dec 2012.

[5] O. H. Ajanki, L. Erdős, and T. Krüger. Stability of the matrix Dyson equation and random matrices with correlations. *Probab. Theory Related Fields*, 173(1-2):293–373, 2019.

[6] J. Alt, L. Erdős, T. Krüger, and Y. Nemish. Location of the spectrum of Kronecker random matrices. *Ann. Inst. Henri Poincaré Probab. Stat.*, 55(2):661–696, 2019.

[7] G. W. Anderson, A. Guionnet, and O. Zeitouni. *An introduction to random matrices*, volume 118 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2010.

[8] A. I. Aptekarev, P. M. Bleher, and A. B. J. Kuijlaars. Large n limit of Gaussian random matrices with external source. II. *Comm. Math. Phys.*, 259(2):367–389, 2005.

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

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

[11] A. Auffinger and W.-K. Chen. Free energy and complexity of spherical bipartite models. *J. Stat. Phys.*, 157(1):40–59, 2014.

[12] A. Auffinger and J. Gold. The number of saddles of the spherical p-spin model, 2020. arXiv:2007.09269v1.

[13] A. Auffinger and Q. Zeng. Complexity of high dimensional Gaussian random fields with isotropic increments, 2020. arXiv:2007.07668v1.

[14] F. Augeri. Large deviations principle for the largest eigenvalue of Wigner matrices without Gaussian tails. *Electron. J. Probab.*, 21:Paper No. 32, 49, 2016.

[15] F. Augeri, A. Guionnet, and J. Husson. Large Deviations for the Largest Eigenvalue of Sub-Gaussian Matrices. *Comm. Math. Phys.*, 383(2):997–1050, 2021.

[16] J.-M. Azaïs and C. Delmas. Mean number and correlation function of critical points of isotropic Gaussian fields and some results on GOE random matrices, 2019. arXiv:1911.02300v3.

[17] J.-M. Azaïs and M. Wschebor. *Level sets and extrema of random processes and fields*. John Wiley & Sons, Inc., Hoboken, NJ, 2009.

[18] Z. Bai and J. W. Silverstein. *Spectral analysis of large dimensional random matrices*. Springer Series in Statistics. Springer, New York, second edition, 2010.

[19] Z. D. Bai. Convergence rate of expected spectral distributions of large random matrices. I. Wigner matrices. *Ann. Probab.*, 21(2):625–648, 1993.

[20] J. Baik and J. O. Lee. Free energy of bipartite spherical Sherrington-Kirkpatrick model. *Ann. Inst. Henri Poincaré Probab. Stat.*, 56(4):2897–2934, 2020.

[21] Z. Bao, L. Erdős, and K. Schnelli. On the support of the free additive convolution. *J. Anal. Math.*, 142(1):323–348, 2020.

[22] Z. Bao, L. Erdős, and K. Schnelli. Spectral rigidity for addition of random matrices at the regular edge. *J. Funct. Anal.*, 279(7):108639, 94, 2020.

[23] Z. Bao, G. Pan, and W. Zhou. The logarithmic law of random determinant. *Bernoulli*, 21(3):1600–1628, 2015.

[24] A. Barra, P. Contucci, E. Mingione, and D. Tantari. Multi-species mean field spin glasses. Rigorous results. *Ann. Henri Poincaré*, 16(3):691–708, 2015.

[25] A. Barra, A. Galluzzi, F. Guerra, A. Pizzoferrato, and D. Tantari. Mean field bipartite spin models treated with mechanical techniques. *Eur. Phys. J. B*, 87(3):Art. 74, 13, 2014.

[26] A. Barra, G. Genovese, and F. Guerra. The replica symmetric approximation of the analogical neural network. *J. Stat. Phys.*, 140(4):784–796, 2010.

[27] A. Barra, G. Genovese, and F. Guerra. Equilibrium statistical mechanics of bipartite spin systems. *J. Phys. A*, 44(24):245002, 22, 2011.

[28] A. Barra, G. Genovese, P. Sollich, and D. Tantari. Phase diagram of restricted boltzmann machines and generalized hopfield networks with arbitrary priors. *Phys. Rev. E*, 97:022310, Feb 2018.

[29] N. P. Baskerville, J. P. Keating, F. Mezzadri, and J. Najnudel. The loss surfaces of neural networks with general activation functions, 2020. arXiv:2004.03959v2.

[30] N. P. Baskerville, J. P. Keating, F. Mezzadri, and J. Najnudel. A spin-glass model for the loss surfaces of generative adversarial networks, 2021. arXiv:2101.02524v1.

[31] M. Bauer and O. Golinelli. Random incidence matrices: moments of the spectral density. *J. Statist. Phys.*, 103(1-2):301–337, 2001.

[32] S. Belinschi, A. Guionnet, and J. Huang. Large deviation principles via spherical integrals, 2020. arXiv:2004.07117v2.

[33] S. T. Belinschi and M. Capitaine. Spectral properties of polynomials in independent Wigner and deterministic matrices. *J. Funct. Anal.*, 273(12):3901–3963, 2017.

[34] D. Belius, J. Černý, S. Nakajima, and M. Schmidt. Triviality of the geometry of mixed p-spin spherical hamiltonians with external field, 2021. arXiv:2104.06345v1.

[35] G. Ben Arous, P. Bourgade, and B. McKenna. Exponential growth of random determinants beyond invariance, 2021.

[36] G. Ben Arous, P. Bourgade, and B. McKenna. Landscape complexity beyond invariance and the elastic manifold, 2021.

[37] G. Ben Arous, A. Dembo, and A. Guionnet. Aging of spherical spin glasses. *Probab. Theory Related Fields*, 120(1):1–67, 2001.

[38] G. Ben Arous, Y. Fyodorov, and B. Khoruzhenko. Counting equilibria of large complex systems by instability index, 2020. To appear in *Proc. Natl. Acad. Sci. U.S.A.* arXiv:2008.00690v1.

[39] G. Ben Arous, R. Gheissari, and A. Jagannath. Algorithmic thresholds for tensor PCA. *Ann. Probab.*, 48(4):2052–2087, 2020.

[40] G. Ben Arous, R. Gheissari, and A. Jagannath. Bounding flows for spherical spin glass dynamics. *Comm. Math. Phys.*, 373(3):1011–1048, 2020.

[41] G. Ben Arous, R. Gheissari, and A. Jagannath. Online stochastic gradient descent on non-convex losses from high-dimensional inference, 2020. arXiv:2003.10409v3.

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

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

[44] G. Ben Arous, E. Subag, and O. Zeitouni. Geometry and temperature chaos in mixed spherical spin glasses at low temperature: the perturbative regime. *Comm. Pure Appl. Math.*, 73(8):1732–1828, 2020.

[45] F. Benaych-Georges, A. Guionnet, and M. Maida. Large deviations of the extreme eigenvalues of random deformations of matrices. *Probab. Theory Related Fields*, 154(3-4):703–751, 2012.

[46] F. Benaych-Georges and S. Péché. Largest eigenvalues and eigenvectors of band or sparse random matrices. *Electron. Commun. Probab.*, 19:no. 4, 9, 2014.

[47] G. Bennett. Upper bounds on the moments and probability inequalities for the sum of independent, bounded random variables. *Biometrika*, 52:559–569, 1965.

[48] H. Bercovici and D. Voiculescu. Free convolution of measures with unbounded support. *Indiana Univ. Math. J.*, 42(3):733–773, 1993.

[49] P. Biane. On the free convolution with a semi-circular distribution. *Indiana Univ. Math. J.*, 46(3):705–718, 1997.

[50] G. Biroli and A. Guionnet. Large deviations for the largest eigenvalues and eigenvectors of spiked Gaussian random matrices. *Electron. Commun. Probab.*, 25:Paper No. 70, 13, 2020.

[51] P. Bleher and A. B. J. Kuijlaars. Large n limit of Gaussian random matrices with external source. I. *Comm. Math. Phys.*, 252(1-3):43–76, 2004.

[52] P. M. Bleher and A. B. J. Kuijlaars. Large n limit of Gaussian random matrices with external source. III. Double scaling limit. *Comm. Math. Phys.*, 270(2):481–517, 2007.

[53] A. Bloemendal, L. Erdős, A. Knowles, H.-T. Yau, and J. Yin. Isotropic local laws for sample covariance and generalized Wigner matrices. *Electron. J. Probab.*, 19:no. 33, 53, 2014.

[54] C. Bordenave and P. Caputo. A large deviation principle for Wigner matrices without Gaussian tails. *Ann. Probab.*, 42(6):2454–2496, 2014.

[55] C. Bordenave, P. Caputo, and D. Chafaï. Spectrum of non-Hermitian heavy tailed random matrices. *Comm. Math. Phys.*, 307(2):513–560, 2011.

[56] F. Bornemann and M. La Croix. The singular values of the GOE. *Random Matrices Theory Appl.*, 4(2):1550009, 32, 2015.

[57] P. Bourgade. Random band matrices. In *Proceedings of the International Congress of Mathematicians—Rio de Janeiro 2018. Vol. IV. Invited lectures*, pages 2759–2784. World Sci. Publ., Hackensack, NJ, 2018.

[58] P. Bourgade, L. Erdős, H.-T. Yau, and J. Yin. Fixed energy universality for generalized Wigner matrices. *Comm. Pure Appl. Math.*, 69(10):1815–1881, 2016.

[59] P. Bourgade and K. Mody. Gaussian fluctuations of the determinant of Wigner matrices. *Electron. J. Probab.*, 24:Paper No. 96, 28, 2019.

[60] P. Bourgade, K. Mody, and M. Pain. Optimal local law and central limit theorem for β-ensembles, 2021. arXiv:2103.06841v2.

[61] J. Bourgain, V. H. Vu, and P. M. Wood. On the singularity probability of discrete random matrices. *J. Funct. Anal.*, 258(2):559–603, 2010.

[62] A. J. Bray and D. S. Dean. Statistics of critical points of gaussian fields on large-dimensional spaces. *Phys. Rev. Lett.*, 98:150201, Apr 2007.

[63] T. T. Cai, T. Liang, and H. H. Zhou. Law of log determinant of sample covariance matrix and optimal estimation of differential entropy for high-dimensional Gaussian distributions. *J. Multivariate Anal.*, 137:161–172, 2015.

[64] M. Capitaine, C. Donati-Martin, D. Féral, and M. Février. Free convolution with a semicircular distribution and eigenvalues of spiked deformations of Wigner matrices. *Electron. J. Probab.*, 16:no. 64, 1750–1792, 2011.

[65] M. Capitaine and S. Péché. Fluctuations at the edges of the spectrum of the full rank deformed GUE. *Probab. Theory Related Fields*, 165(1-2):117–161, 2016.

[66] Z. Che and B. Landon. Local spectral statistics of the addition of random matrices. *Probab. Theory Related Fields*, 175(1-2):579–654, 2019.

[67] R. Delannay and G. Le Caër. Distribution of the determinant of a random real-symmetric matrix from the Gaussian orthogonal ensemble. *Phys. Rev. E (3)*, 62(2, part A):1526–1536, 2000.

[68] A. Dembo. On random determinants. *Quart. Appl. Math.*, 47(2):185–195, 1989.

[69] A. Dembo and O. Zeitouni. *Large deviations techniques and applications*, volume 38 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, 2010. Corrected reprint of the second (1998) edition.

[70] C. Donati-Martin and M. Maïda. Large deviations for the largest eigenvalue of an Hermitian Brownian motion. *ALEA Lat. Am. J. Probab. Math. Stat.*, 9(2):501–530, 2012.

[71] R. M. Dudley. *Real analysis and probability*, volume 74 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2002. Revised reprint of the 1989 original.

[72] A. Edelman and M. La Croix. The singular values of the GUE (less is more). *Random Matrices Theory Appl.*, 4(4):1550021, 37, 2015.

[73] A. Engel. Replica symmetry breaking in zero dimension. *Nucl. Phys. B*, 410(3):617–646, 1993.

[74] L. Erdős, A. Knowles, H.-T. Yau, and J. Yin. Spectral statistics of Erdős-Rényi graphs I: Local semicircle law. *Ann. Probab.*, 41(3B):2279–2375, 2013.

[75] L. Erdős, T. Krüger, and D. Schröder. Random matrices with slow correlation decay. *Forum Math. Sigma*, 7:e8, 89, 2019.

[76] L. Erdős, B. Schlein, and H.-T. Yau. Wegner estimate and level repulsion for Wigner random matrices. *Int. Math. Res. Not. IMRN*, 2010(3):436–479, 2010.

[77] L. Erdős, H.-T. Yau, and J. Yin. Bulk universality for generalized Wigner matrices. *Probab. Theory Related Fields*, 154(1-2):341–407, 2012.

[78] Z. Fan, S. Mei, and A. Montanari. TAP free energy, spin glasses and variational inference. *Ann. Probab.*, 49(1):1–45, 2021.

[79] A. Fey, R. van der Hofstad, and M. J. Klok. Large deviations for eigenvalues of sample covariance matrices, with applications to mobile communication systems. *Adv. in Appl. Probab.*, 40(4):1048–1071, 2008.

[80] D. S. Fisher. Interface fluctuations in disordered systems: $5 - \varepsilon$ expansion and failure of dimensional reduction. *Phys. Rev. Lett.*, 56:1964–1967, May 1986.

[81] G. E. Forsythe and J. W. Tukey. The extent of n random unit vectors. *Bull. Amer. Math. Soc.*, 58:502, 1952.

[82] R. Fortet. Random determinants. *J. Research Nat. Bur. Standards*, 47:465–470, 1951.

[83] Z. Füredi and J. Komlós. The eigenvalues of random symmetric matrices. *Combinatorica*, 1(3):233–241, 1981.

[84] Y. V. Fyodorov. Complexity of random energy landscapes, glass transition, and absolute value of the spectral determinant of random matrices. *Phys. Rev. Lett.*, 92(24):240601, 4, 2004.

[85] Y. V. Fyodorov and J.-P. Bouchaud. Statistical mechanics of a single particle in a multiscale random potential: Parisi landscapes in finite-dimensional Euclidean spaces. *J. Phys. A*, 41(32):324009, 25, 2008.

[86] Y. V. Fyodorov, I. Y. Korenblit, and E. F. Shender. Antiferromagnetic ising spin glass. *J. Phys. C*, 20(12):1835–1839, Apr 1987.

[87] Y. V. Fyodorov, I. Y. Korenblit, and E. F. Shender. Phase transitions in frustrated metamagnets. *Europhysics Letters (EPL)*, 4(7):827–832, Oct 1987.

[88] Y. V. Fyodorov and P. Le Doussal. Manifolds in a high-dimensional random landscape: Complexity of stationary points and depinning. *Phys. Rev. E*, 101:020101, Feb 2020.

[89] Y. V. Fyodorov and P. Le Doussal. Manifolds pinned by a high-dimensional random landscape: Hessian at the global energy minimum. *J. Stat. Phys.*, 179(1):176–215, 2020.

[90] Y. V. Fyodorov, P. Le Doussal, A. Rosso, and C. Texier. Exponential number of equilibria and depinning threshold for a directed polymer in a random potential. *Ann. Physics*, 397:1–64, 2018.

[91] Y. V. Fyodorov and C. Nadal. Critical behavior of the number of minima of a random landscape at the glass transition point and the tracy-widom distribution. *Phys. Rev. Lett.*, 109:167203, Oct 2012.

[92] Y. V. Fyodorov and H.-J. Sommers. Classical particle in a box with random potential: Exploiting rotational symmetry of replicated hamiltonian. *Nucl. Phys. B*, 764:128–167, March 2007.

[93] Y. V. Fyodorov, H.-J. Sommers, and I. Williams. Density of stationary points in a high dimensional random energy landscape and the onset of glassy behavior. *JETP Lett.*, 85:261–266, May 2007.

[94] Y. V. Fyodorov and I. Williams. Replica symmetry breaking condition exposed by random matrix calculation of landscape complexity. *J. Stat. Phys.*, 129(5-6):1081–1116, 2007.

[95] T. Giamarchi. Disordered elastic media. In R. A. Meyers, editor, *Encyclopedia of Complexity and Systems Science*, pages 2019–2038. Springer New York, New York, NY, 2009.

[96] T. Giamarchi and P. Le Doussal. Statics and dynamics of disordered elastic systems. In A. P. Young, editor, *Spin Glasses and Random Fields*, pages 321–356. World Scientific, 1997.

[97] V. L. Girko. *Theory of random determinants*, volume 45 of *Mathematics and its Applications (Soviet Series)*. Kluwer Academic Publishers Group, Dordrecht, 1990. Translated from the Russian.

[98] N. R. Goodman. The distribution of the determinant of a complex Wishart distributed matrix. *Ann. Math. Statist.*, 34:178–180, 1963.

[99] A. Guionnet and J. Husson. Large deviations for the largest eigenvalue of Rademacher matrices. *Ann. Probab.*, 48(3):1436–1465, 2020.

[100] A. Guionnet and J. Husson. Asymptotics of k dimensional spherical integrals and applications, 2021. arXiv:2101.01983.

[101] A. Guionnet and M. Maïda. A Fourier view on the R-transform and related asymptotics of spherical integrals. *J. Funct. Anal.*, 222(2):435–490, 2005.

[102] A. Guionnet and M. Maïda. Large deviations for the largest eigenvalue of the sum of two random matrices. *Electron. J. Probab.*, 25:Paper No. 14, 24, 2020.

[103] A. Guionnet and O. Zeitouni. Concentration of the spectral measure for large matrices. *Electron. Comm. Probab.*, 5:119–136, 2000.

[104] A. Guionnet and O. Zeitouni. Large deviations asymptotics for spherical integrals. *J. Funct. Anal.*, 188(2):461–515, 2002.

[105] G. S. Hartnett, E. Parker, and E. Geist. Replica symmetry breaking in bipartite spin glasses and neural networks. *Phys. Rev. E*, 98:022116, Aug 2018.

[106] J. W. Helton, R. R. Far, and R. Speicher. Operator-valued semicircular elements: Solving a quadratic matrix equation with positivity constraints. *International Mathematics Research Notices*, 2007, 01 2007. rnm086.

[107] M. Kac. On the average number of real roots of a random algebraic equation. *Bull. Amer. Math. Soc.*, 49(4):314–320, 04 1943.

[108] J. Kahn, J. Komlós, and E. Szemerédi. On the probability that a random ±1-matrix is singular. *J. Amer. Math. Soc.*, 8(1):223–240, 1995.

[109] J. M. Kincaid and E. G. D. Cohen. Phase diagrams of liquid helium mixtures and metamagnets: Experiment and mean field theory. *Phys. Lett. C*, 22(2):57–143, 1975.

[110] J. Komlós. On the determinant of (0, 1) matrices. *Studia Sci. Math. Hungar.*, 2:7–21, 1967.

[111] J. Komlós. On the determinant of random matrices. *Studia Sci. Math. Hungar.*, 3:387–399, 1968.

[112] I. Y. Korenblit and E. F. Shender. Spin glass in an Ising two-sublattice magnet. *Zh. Eksp. Teor. Fiz.*, 89:1785—1795, 1985.

[113] B. Landon, P. Sosoe, and H.-T. Yau. Fixed energy universality of Dyson Brownian motion. *Adv. Math.*, 346:1137–1332, 2019.

[114] J. O. Lee and K. Schnelli. Local deformed semicircle law and complete delocalization for Wigner matrices with random potential. *J. Math. Phys.*, 54(10):103504, 62, 2013.

[115] J. O. Lee and K. Schnelli. Edge universality for deformed Wigner matrices. *Rev. Math. Phys.*, 27(8):1550018, 94, 2015.

[116] J. O. Lee, K. Schnelli, B. Stetler, and H.-T. Yau. Bulk universality for deformed Wigner matrices. *Ann. Probab.*, 44(3):2349–2425, 2016.

[117] M. S. Longuet-Higgins. The statistical analysis of a random, moving surface. *Philos. Trans. Roy. Soc. London Ser. A*, 249:321–387, 1957.

[118] M. Maïda. Large deviations for the largest eigenvalue of rank one deformations of Gaussian ensembles. *Electron. J. Probab.*, 12:1131–1150, 2007.

[119] A. Matytsin. On the large-N limit of the Itzykson-Zuber integral. *Nuclear Phys. B*, 411(2-3):805–820, 1994.

[120] B. McKenna. Complexity of bipartite spherical spin glasses, 2021.

[121] B. McKenna. Large deviations for extreme eigenvalues of deformed Wigner random matrices. *Electronic Journal of Probability*, 26(paper no. 34):1 – 37, 2021.

[122] M. Mézard and G. Parisi. Replica field theory for random manifolds. *J. Phys. I France*, 1(6):809–836, 1991.

[123] M. Mézard and G. Parisi. Manifolds in random media: two extreme cases. *J. Phys. I France*, 2(12):2231–2242, 1992.

[124] J.-C. Mourrat. Nonconvex interactions in mean-field spin glasses, 2020. arXiv:2004.01679v6.

[125] H. H. Nguyen. On the least singular value of random symmetric matrices. *Electron. J. Probab.*, 17:no. 53, 19, 2012.

[126] H. H. Nguyen and V. Vu. Random matrices: law of the determinant. *Ann. Probab.*, 42(1):146–167, 2014.

[127] J. Novak. On the complex asymptotics of the HCIZ and BGW integrals, 2020. arXiv:2006.04304v1.

[128] H. Nyquist, S. O. Rice, and J. Riordan. The distribution of random determinants. *Quart. Appl. Math.*, 12:97–104, 1954.

[129] D. Panchenko. The free energy in a multi-species Sherrington-Kirkpatrick model. *Ann. Probab.*, 43(6):3494–3513, 2015.

[130] L. Pastur and M. Shcherbina. *Eigenvalue distribution of large random matrices*, volume 171 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2011.

[131] L. A. Pastur. The spectrum of random matrices. *Teoret. Mat. Fiz.*, 10(1):102–112, 1972.

[132] A. Prékopa. On random determinants. I. *Studia Sci. Math. Hungar.*, 2:125–132, 1967.

[133] S. O. Rice. Mathematical analysis of random noise. *Bell System Tech. J.*, 23:282–332, 1944.

[134] S. O. Rice. Mathematical analysis of random noise. *Bell System Tech. J.*, 24:46–156, 1945.

[135] V. Ros. Distribution of rare saddles in the p-spin energy landscape. *J. Phys. A*, 53(12):125002, 53, 2020.

[136] V. Ros, G. Ben Arous, G. Biroli, and C. Cammarota. Complex energy landscapes in spiked-tensor and simple glassy models: Ruggedness, arrangements of local minima, and phase transitions. *Phys. Rev. X*, 9:011003, Jan 2019.

[137] V. Ros, G. Biroli, and C. Cammarota. Complexity of energy barriers in mean-field glassy systems. *EPL (Europhysics Letters)*, 126(2):20003, may 2019.

[138] A. Rouault. Asymptotic behavior of random determinants in the Laguerre, Gram and Jacobi ensembles. *ALEA Lat. Am. J. Probab. Math. Stat.*, 3:181–230, 2007.

[139] I. J. Schoenberg. Metric spaces and completely monotone functions. *Ann. of Math. (2)*, 39(4):811–841, 1938.

[140] T. Shcherbina. On universality of local edge regime for the deformed Gaussian unitary ensemble. *J. Stat. Phys.*, 143(3):455–481, 2011.

[141] E. Subag. The complexity of spherical p-spin models—a second moment approach. *Ann. Probab.*, 45(5):3385–3450, 2017.

[142] G. Szekeres and P. Turán. On an extremal problem in the theory of determinants. *Math. Naturwiss. Am . Ungar. Akad. Wiss*, 56:796–806, 1937.

[143] M. Talagrand. A new look at independence. *Ann. Probab.*, 24(1):1–34, 1996.

[144] T. Tao and V. Vu. On random ± 1 matrices: singularity and determinant. *Random Structures Algorithms*, 28(1):1–23, 2006.

[145] T. Tao and V. Vu. On the singularity probability of random Bernoulli matrices. *J. Amer. Math. Soc.*, 20(3):603–628, 2007.

[146] T. Tao and V. Vu. A central limit theorem for the determinant of a Wigner matrix. *Adv. Math.*, 231(1):74–101, 2012.

[147] A. N. Tikhomirov. On the rate of convergence of the expected spectral distribution function of a Wigner matrix to the semi-circular law. *Siberian Adv. Math.*, 19(3):211–223, 2009.

[148] A. N. Tikhomirov. The rate of convergence of the expected spectral distribution function of a sample covariance matrix to the Marchenko-Pastur distribution. *Siberian Adv. Math.*, 19(4):277–286, 2009.

[149] K. Tikhomirov. Singularity of random Bernoulli matrices. *Ann. of Math. (2)*, 191(2):593–634, 2020.

[150] L. V. Tran, V. H. Vu, and K. Wang. Sparse random graphs: eigenvalues and eigenvectors. *Random Structures Algorithms*, 42(1):110–134, 2013.

[151] P. Turán. On a problem in the theory of determinants. *Acta Math. Sinica*, 5:411–423, 1955.

[152] D. Voiculescu. Addition of certain noncommuting random variables. *J. Funct. Anal.*, 66(3):323–346, 1986.

[153] D. Voiculescu. Limit laws for random matrices and free products. *Invent. Math.*, 104(1):201–220, 1991.

[154] V. H. Vu. Spectral norm of random matrices. *Combinatorica*, 27(6):721–736, 2007.

[155] K. W. Wachter. The strong limits of random matrix spectra for sample matrices of independent elements. *Ann. Probability*, 6(1):1–18, 1978.

[156] A. M. Yaglom. Certain types of random fields in n-dimensional space similar to stationary stochastic processes. *Teor. Veroyatnost. i Primenen*, 2:292–338, 1957.

[157] A. Zee. Law of addition in random matrix theory. *Nuclear Phys. B*, 474(3):726–744, 1996.