

Overlaps of a spherical spin glass model with microscopic external field*

Elizabeth Collins-Woodfin[†]

Abstract

We examine the behavior of the 2-spin spherical Sherrington-Kirkpatrick model with an external field by analyzing the overlap of a spin with the external field. Previous research has noted that, at low temperature, this overlap exhibits dramatically different behavior in the presence of an external field as compared to the model with no external field. The transition between those two settings was examined in a recent physics paper by Baik, Collins-Woodfin, Le Doussal, and Wu as well as a recent math paper by Landon and Sosoe. Both papers focus on the setting in which the external field strength, h , approaches zero as the dimension, N , approaches infinity. In particular, the paper of Baik et al studies the overlap with a microscopic external field ($h \sim N^{-1/2}$) but without a rigorous proof. This paper aims to give a proof of that result. The proof involves representing the generating function of the overlap as a ratio of contour integrals and then analyzing the asymptotics of those contour integrals using results from random matrix theory.

Keywords: Sherrington-Kirkpatrick; spin glass.

MSC2020 subject classifications: Primary 60, Secondary 82; 15.

Submitted to EJP on January 29, 2021, final version accepted on November 3, 2021.

1 Introduction

1.1 Model set-up and notations

The 2-spin spherical Sherrington-Kirkpatrick (SSK) spin glass model involves a spin variable $\sigma = (\sigma_1, \dots, \sigma_N)$ in S_{N-1} , the sphere of radius \sqrt{N} in \mathbb{R}^N :

$$S_{N-1} = \{\sigma \in \mathbb{R}^N : \|\sigma\| = \sqrt{N}\}.$$

The SSK model with external field is defined by the Hamiltonian

$$\mathcal{H}(\sigma) = -\frac{1}{2} \sum_{i,j=1}^N M_{ij} \sigma_i \sigma_j - h \sum_{i=1}^N g_i \sigma_i = -\frac{1}{2} \sigma \cdot M \sigma - h \mathbf{g} \cdot \sigma \quad (1.1)$$

*Supported in part by NSF grants DMS-1701577 and DMS-1954790.

[†]University of Michigan, United States of America. E-mail: elicolli@umich.edu

where \mathbf{g} is a standard Gaussian random vector and M is an $N \times N$ random symmetric matrix from the Gaussian orthogonal ensemble (GOE). More specifically, for $i \leq j$, the variables M_{ij} are independent centered Gaussian random variables with variance $\frac{1}{N}$ for $i < j$ and $\frac{2}{N}$ for $i = j$. By the symmetry condition, $M_{ij} = M_{ji}$ for $i > j$.

The Gibbs measure for this model is

$$p(\boldsymbol{\sigma}) = \frac{1}{\mathcal{Z}_N} e^{-\beta \mathcal{H}(\boldsymbol{\sigma})} \quad \text{for } \boldsymbol{\sigma} \in S_{N-1} \quad (1.2)$$

where the parameter $\beta > 0$ denotes the inverse temperature and is also written $\beta = \frac{1}{T}$. The partition function \mathcal{Z}_N is given by

$$\mathcal{Z}_N = \int_{S_{N-1}} e^{-\beta \mathcal{H}(\boldsymbol{\sigma})} d\omega_N(\boldsymbol{\sigma}) \quad (1.3)$$

where ω_N is the normalized uniform measure on S_{N-1} . Since M and \mathbf{g} are random, the Gibbs measure is a random measure. We use the notation

$$\langle A \rangle = \int A(\boldsymbol{\sigma}) dp(\boldsymbol{\sigma}) \quad (1.4)$$

to denote the expectation of A with respect to the Gibbs measure where A depends on $\boldsymbol{\sigma}$. Since the Gibbs measure depends on M and \mathbf{g} , the Gibbs expectation $\langle A \rangle$ is a function of M and \mathbf{g} .

Overlaps are of particular interest in the study of SSK and other spin glass models. In this paper we focus on the overlap of a spin with the external field, denoted by \mathfrak{M} , and briefly discuss the overlap with a replica, denoted by \mathfrak{R} . These are defined as

$$\mathfrak{M} = \frac{\mathbf{g} \cdot \boldsymbol{\sigma}}{N} \quad \text{and} \quad \mathfrak{R} = \frac{\boldsymbol{\sigma}^{(1)} \cdot \boldsymbol{\sigma}^{(2)}}{N} \quad (1.5)$$

where $\boldsymbol{\sigma}$ is chosen randomly according to the Gibbs measure and $\boldsymbol{\sigma}^{(1)}, \boldsymbol{\sigma}^{(2)}$ are two independent copies (or replicas) of $\boldsymbol{\sigma}$, chosen from the Gibbs measure with the same values for M and \mathbf{g} .

We denote the eigenvalues of M and their corresponding unit eigenvectors by

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N \quad \text{and} \quad \mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_N.$$

Many calculations will involve the inner product of the external field with an eigenvector. For this purpose, we introduce the notation

$$n_i = \mathbf{g} \cdot \mathbf{u}_i. \quad (1.6)$$

1.2 Background

The SSK model was first studied by Kosterlitz, Thouless and Jones [15] as a continuous analog of the model with Ising spins ($\boldsymbol{\sigma} \in \{-1, +1\}^N$) that was introduced by Sherrington and Kirkpatrick [22]. Some of the most influential early work on these models focused on calculating the limiting free energy (as $N \rightarrow \infty$) for the SK model, SSK model and the more general p -spin models. A formula for the limiting free energy of the SK model (as well as the more general p -spin case) was first established by Parisi [21], and analogous formulas for spherical models were developed by Kosterlitz et al [15] for the SSK model and by Crisanti and Sommers [8] for p -spin spherical models. These formulas were later proven rigorously by Talagrand [26, 25].

The SSK model allows for certain the types of analysis that are not possible with the SK model. In particular, the partition function for SSK has an equivalent formulation

as a contour integral, which we describe in more detail in Section 3.1. This contour integral representation of \mathcal{Z}_N was first observed by [15] and was later used by Baik and Lee [3] to analyze the free energy of SSK with zero external field up to fluctuations for both the high temperature ($T > 1$) and low temperature ($T < 1$) cases. They also used this method to analyze the free energy of several related models [4, 5, 6].

Overlaps are useful in terms of studying the distribution of spins (see, for example, [24] for a discussion of the geometry of the Gibbs measure in p -spin spherical models without external field). In the case of SSK with $h = 0$, the overlap with a replica concentrates around the values $\pm(1 - T)_+$ [20]. The contour integral representation of the partition function can be used to study the fluctuations of this overlap. In particular, Nguyen and Sosoe [19] find that, at high temperature and zero external field, the overlap with a replica converges to a centered Gaussian distribution. Landon and Sosoe [16] extend this analysis to the low temperature case and find that the overlap with a replica is no longer normally distributed but instead converges to a mean zero, bimodal random variable whose distribution can be expressed explicitly as a function of the GOE Airy point process.

The aforementioned results all indicate that, in the $h = 0$ setting, the SSK model displays a phase transition at temperature $T = 1$. However, this transition does not occur in the case where we have fixed positive external field strength $h > 0$ [15, 8, 25, 7]. In that case, the overlap concentrates around a positive value, $q(h)$, for all temperatures [20].

Since the low temperature case reveals a striking contrast between the SSK model with $h = 0$ and the one with $h > 0$, it is interesting to consider the behavior of the low temperature model when $h \rightarrow 0$ as $N \rightarrow \infty$. This question has been considered in the physics context by Fyodorov and le Doussal [13] who examine the free energy of SSK with external field and find transitional regimes at scalings $h \sim N^{-1/6}$ and $h \sim N^{-1/2}$ in the zero temperature case (see also [10, 14]). We refer to these scalings as mesoscopic and microscopic respectively.

Recent papers by Landon and Sosoe [17] and by Baik, Collins-Woodfin, le Doussal, and Wu [2] have computed the distribution of the overlap with the external field as well as the overlap with a replica at the transitional scalings $h \sim N^{-1/6}$ and $h \sim N^{-1/2}$ for $T < 1$. (Both papers also analyze free energy, but that will not be discussed in the current paper). At the microscopic scaling, both papers find that the overlap with a replica concentrates around the values $\pm(1 - T)$, as it does in the $h = 0$ case. However, the mean is no longer zero, but positive, and the mean increases as $hN^{1/2}$ increases. We note that both papers use a similar method of representing the overlap in terms of contour integrals. The difference is that, while [17] is mathematically rigorous, [2] does not provide all the details of the argument but instead focuses on the physical implications.

In addition to the overlap with a replica, it is also interesting to study the overlap with the external field. This quantity is used by physicists to study magnetism and susceptibility (for example [15, 9]), which will be discussed further in Section 2.1. Both [17] and [2] analyze the overlap with the external field at the macroscopic ($h = O(1)$) and mesoscopic scalings. The paper [2] also analyzes this overlap at the microscopic scaling by providing a non-rigorous computation of the moment generating function for the overlap. Their computation suggests that, in its leading order, the overlap is distributed like the sum of two independent random variables, one of which is Bernoulli and one of which is Gaussian, and both are of order $N^{-1/2}$. Providing a rigorous proof of this result will be the focus of the current paper.

1.3 Scope and organization of this paper

As mentioned above, the goal of the current paper is to provide a rigorous proof of the formula conjectured in [2] for the distribution of the overlap with the external field at the microscopic scaling. Although some intermediate steps of [2] are rigorous, many details are omitted and, most notably, they do not provide rigorous proofs for the asymptotics of the integrals. We follow similar steps to those used in [2], but fill in the missing details and supply the rigorous asymptotic analysis of the integrals. Furthermore, we provide more specificity regarding the probability with which the results hold. The authors of [2] conjecture that the formula holds with a probability that tends to 1 as $N \rightarrow \infty$. We show that it holds with probability at least $1 - N^{-\varepsilon}$ for any sufficiently small $\varepsilon > 0$.

The proof in this paper utilizes the contour integral representation of the partition function as well as many results from random matrix theory. The asymptotic evaluation of the contour integral associated with the overlap requires a choice of the contour. The papers [17] and [2] use the steepest-descent contour. We found it simpler to use an explicit contour that agrees with the true steepest-descent contour only locally.

Finally, we note that our method can also be used to compute the moment generating function for the overlap with a replica, as conjectured in [2]. However, we do not provide the details of that proof, since the result is proved via a different method in [17].

Section 2 summarizes the main theorem to be proved in this paper and its implications. In Section 3 we provide the notations and prerequisites that we will use throughout the paper including some lemmas that were implied but not rigorously proven in [2]. Section 4 provides the proof of Theorems 2.1 following similar steps as in [2] but filling in some details and providing more specificity about the probability with which results hold. Section 5 provides the detailed computations for the decay of the contour integrals outside a certain neighborhood of the critical point. This is the most technical part of the paper and supplies key computations that were not addressed in [2]. The lemma in this section is used in the proof in Section 4. Finally, Section 6 provides a brief description of how this approach can be applied to the analysis of the overlap of two replicas.

2 Summary of main theorem

This paper focuses on the proof of Theorem 2.1, which provides the moment generating function for the overlap \mathfrak{M} . It is important to note that this overlap involve two types of randomness. First, we have randomness from the choice of M and \mathbf{g} , which we refer to jointly as the “disorder sample.” Second, we have randomness from the choice of spin variable. For the results in this paper, we fix an arbitrary disorder sample so that \mathfrak{M} is a random variable depending on a fixed disorder sample and random spin variable. The moment generating function in Theorem 2.1 provides the distribution of \mathfrak{M} as a function of the fixed disorder sample.

This result is valid for an arbitrary disorder sample, subject to certain constraints that hold with high probability. In particular, for any sufficiently small $\varepsilon > 0$, the event \mathcal{E}_ε (defined in Section 3.3) provides a set of conditions on M and \mathbf{g} that are sufficient for Theorem 2.1 to hold. Section 3.3 provides a detailed description of the event \mathcal{E}_ε along with a proof that

$$\mathbb{P}(\mathcal{E}_\varepsilon) \geq 1 - N^{-\varepsilon/10} \quad \text{for all sufficiently small } \varepsilon > 0 \text{ and all sufficiently large } N. \quad (2.1)$$

Theorem 2.1. *Given $T < 1$ with $h = HN^{-1/2}$ for some fixed $H \geq 0$ and n_1 as defined in (1.6), we have the following asymptotic formula for the moment generating function of \mathfrak{M} , the overlap with the external field. This formula holds on the event \mathcal{E}_ε*

(which has probability at least $1 - N^{-\varepsilon/10}$) for any sufficiently small $\varepsilon > 0$ and $\xi = O(1)$.

$$\langle e^{\xi\sqrt{N}\mathfrak{M}} \rangle = e^{H\xi + \frac{T\xi^2}{2}} \frac{\cosh\left((H + T\xi)|n_1|\frac{\sqrt{1-T}}{T}\right)}{\cosh\left(H|n_1|\frac{\sqrt{1-T}}{T}\right)} \left(1 + O(N^{-\frac{1}{21} + \frac{\varepsilon}{7}})\right). \quad (2.2)$$

Note that the leading term on the right-hand side is the product of two terms implying that it is the moment generating function of a sum of two independent random variables. The exponential term is the moment generating function of a Gaussian random variable. For the ratio of the cosh functions, we note that the moment generating function of a shifted Bernoulli random variable that takes values 1 and -1 with probabilities P and $1 - P$ respectively is $Pe^t + (1 - P)e^{-t}$. The ratio of cosh functions in Theorem 2.1 is of this form with $t = \xi|n_1|\sqrt{1 - T}$ and

$$P = \frac{e^{\frac{H}{T}|n_1|\sqrt{1-T}}}{e^{\frac{H}{T}|n_1|\sqrt{1-T}} + e^{-\frac{H}{T}|n_1|\sqrt{1-T}}}. \quad (2.3)$$

Hence, for any large N , we can conclude that, on the event \mathcal{E}_ε , the scaled overlap $\sqrt{N}\mathfrak{M}$ behaves in its leading order like the independent sum of a Gaussian random variable (with mean H and variance T) and a shifted Bernoulli random variable (which takes values $|n_1|\sqrt{1 - T}$ and $-|n_1|\sqrt{1 - T}$ with probability P and $1 - P$ respectively for the value of P stated above).

We can use Theorem 2.1 to obtain various information about the overlaps, including formulas for all moments of \mathfrak{M} . Of particular interest are the first moment (Gibbs expectation) and the variance with respect to the Gibbs measure. Since \mathfrak{M} is of order $N^{-1/2}$ in the case of a microscopic external field, we examine the scaled overlap $\mathfrak{M}\sqrt{N}$. For the expectation, we get

$$\langle \mathfrak{M}\sqrt{N} \rangle = H + |n_1|\sqrt{1 - T} \tanh\left(H|n_1|\frac{\sqrt{1 - T}}{T}\right) + O\left(N^{-\frac{1}{21} + \frac{\varepsilon}{3}}\right) \quad (2.4)$$

and for the variance with respect to the Gibbs measure, we get

$$\text{Var}(\mathfrak{M}\sqrt{N}) = T + n_1^2(1 - T) \left(1 - \tanh^2\left(H|n_1|\frac{\sqrt{1 - T}}{T}\right)\right) + O\left(N^{-\frac{1}{21} + \frac{\varepsilon}{3}}\right), \quad (2.5)$$

where both of these formulas hold on the event \mathcal{E}_ε .

2.1 Application to magnetization and susceptibility

One important application of Theorem 2.1 is that it confirms the conjectures of [2] regarding magnetization and susceptibility. Magnetization is defined to be $\langle \mathfrak{M} \rangle$, the Gibbs average of the overlap with the external field. Susceptibility is the magnetization per unit external field strength, given by

$$\mathcal{X} = \frac{\langle \mathfrak{M} \rangle}{h}. \quad (2.6)$$

It follows from Theorem 2.1 that, on the event \mathcal{E}_ε , when $T < 1$ and $h = HN^{-1/2}$ for H constant, the susceptibility is

$$\mathcal{X} = 1 + \frac{|n_1|\sqrt{1 - T}}{H} \tanh\left(H|n_1|\frac{\sqrt{1 - T}}{T}\right) + O(N^{-\frac{1}{21} + \frac{\varepsilon}{3}}). \quad (2.7)$$

Of particular interest in the physics literature is the zero external field limit of the susceptibility. Cugliandolo, Dean, and Yoshino [9] discuss two ways to taking this limit,

namely $\lim_{h \rightarrow 0} \lim_{N \rightarrow \infty} \mathcal{X}$ and $\lim_{N \rightarrow \infty} \lim_{h \rightarrow 0} \mathcal{X}$ (the first of these was also considered by [15]). Our results for the microscopic external field give a different way of calculating the second of these limits. In particular, since $n_1 \sim \mathcal{N}(0, 1)$ for all N , we can compute

$$\lim_{H \rightarrow 0} \lim_{\substack{N \rightarrow \infty \\ h = HN^{-1/2}}} \mathcal{X} \stackrel{\mathcal{D}}{=} 1 + \frac{\nu^2(1-T)}{T} \quad \text{for } T < 1 \text{ and } \nu \sim \mathcal{N}(0, 1) \quad (2.8)$$

where $\stackrel{\mathcal{D}}{=}$ denotes equality in distribution. This confirms the conjecture of [2]. It is also consistent with [9], which found that, after imposing the constraint $|n_1| = 1$,

$$\lim_{N \rightarrow \infty} \lim_{h \rightarrow 0} \mathcal{X} = \frac{1}{T} \quad \text{for } T < 1 \quad [9]. \quad (2.9)$$

By removing this constraint on $|n_1|$ and applying Theorem 2.1, we are able to show that the limiting susceptibility is not constant, as the result from [9] might suggest, but rather it is a random variable whose distribution is given explicitly in (2.8). The paper [2] contains some further conjectures about the zero external field limit of differential susceptibility, which can also be verified using Theorem 2.1.

2.2 Comparison with the results of [2] and [17]

Theorem 2.1 is similar to Result 8.6 from [2] but we provide a more precise statement of the result as well as a rigorous proof. In particular, we specify bounds for the order of the subleading term and the probability with which the result holds. The proof of Theorem 2.1 can be found in Section 4. An additional lemma needed in the proof is included in section 5.

A similar approach can also be used to obtain a moment generating function for the overlap of two replicas. In other words, we can adapt the methods from Sections 4 and 5 to prove a rigorous version of Result 10.6 from [2]. This will be discussed further in Section 6. We do not provide the details of that proof because a comparable result was obtained via a different method in [17] (see Theorem 2.14).

3 Preliminaries

3.1 Contour integral representations

Recall that the partition function \mathcal{Z}_N is defined by the surface integral

$$\mathcal{Z}_N = \int_{S_{N-1}} e^{-\beta \mathcal{H}(\boldsymbol{\sigma})} d\omega_N(\boldsymbol{\sigma}). \quad (3.1)$$

It was shown by Kostelitz, Thouless and Jones [15] that this surface integral can be rewritten as a contour integral of the form

$$\mathcal{Z}_N = C_N \int_{\gamma-i\infty}^{\gamma+i\infty} e^{\frac{N}{2} \mathcal{G}(z)} dz \quad \text{with} \quad C_N = \frac{\Gamma(N/2)}{2\pi i (N\beta/2)^{N/2-1}} \quad (3.2)$$

where

$$\mathcal{G}(z) = \beta z - \frac{1}{N} \sum_{i=1}^N \log(z - \lambda_i) + \frac{h^2 \beta}{N} \sum_{i=1}^N \frac{n_i^2}{z - \lambda_i} \quad (3.3)$$

and the contour is a vertical line intersecting the real axis at any $\gamma > \lambda_1$. This result can be extended in a straightforward way to obtain contour integral representations for generating functions of overlaps [2, 17]. The results we use are given in the following lemma and a proof can be found in [2].

Lemma 3.1 ([2]). *For real parameter η , the moment generation function of the overlaps is*

$$\langle e^{\beta\eta\mathfrak{M}} \rangle = \frac{\int e^{\frac{N}{2}\mathcal{G}_{\mathfrak{M}}(z)} dz}{\int e^{\frac{N}{2}\mathcal{G}(z)} dz} \quad (3.4)$$

where each contour is a vertical line passing to the right of all singularities and the function $\mathcal{G}_{\mathfrak{M}}(z)$ is defined as follows:

$$\mathcal{G}_{\mathfrak{M}}(z) := \beta z - \frac{1}{N} \sum_{i=1}^N \log(z - \lambda_i) + \frac{(h + \frac{\eta}{N})^2 \beta}{N} \sum_{i=1}^N \frac{n_i^2}{z - \lambda_i} \quad (3.5)$$

Note that $\mathcal{G}_{\mathfrak{M}}(z)$ is $\mathcal{G}(z)$ with h replaced by $h + \eta N^{-1}$.

3.2 Preliminaries from random matrix theory and probability

In this section we present a few classical results from random matrix theory as well as some specific convergence results for certain functions of eigenvalues that we will use throughout the paper.

Semicircle Law

For every bounded, continuous function $f(x)$, we have the following convergence of the empirical distribution of eigenvalues of M [18]:

$$\frac{1}{N} \sum_{i=1}^N f(\lambda_i) \rightarrow \int f(x) d\sigma_{scl}(x) \quad \text{where} \quad d\sigma_{scl}(x) = \frac{\sqrt{4-x^2}}{2\pi} \mathbb{1}_{x \in [-2,2]} dx \quad (3.6)$$

with probability 1 as $N \rightarrow \infty$.

Eigenvalue Rigidity

For $i = 1, 2, \dots, N$, we let $\hat{\lambda}_i$ denote the classical location of the i th eigenvalue, defined by

$$\int_{\hat{\lambda}_i}^2 d\sigma_{scl}(x) = \frac{i}{N}. \quad (3.7)$$

We set $\hat{\lambda}_0 = 2$. The rigidity result [12, 11] states that, for any $\delta > 0$ and $D > 0$ and sufficiently large N ,

$$\mathbb{P} \left(\bigcap_{i=1}^N \left\{ |\lambda_i - \hat{\lambda}_i| \leq N^{-\frac{2}{3}+\delta} (\min\{i, N+1-i\})^{-1/3} \right\} \right) \geq 1 - N^{-D} \quad (3.8)$$

Airy Point Process

Define the rescaled eigenvalues

$$a_i := N^{2/3}(\lambda_i - 2). \quad (3.9)$$

As $N \rightarrow \infty$, the rescaled eigenvalues converge in distribution to the GOE Airy point process [27, 23]. We denote this as $\{\alpha_i\}_{i=1}^\infty$ satisfying

$$\{a_i\} \Rightarrow \{\alpha_i\}. \quad (3.10)$$

Heuristically, we expect that, for $1 \ll i \ll N$,

$$a_i \approx \alpha_i \approx - \left(\frac{3\pi i}{2} \right)^{2/3} \quad (3.11)$$

since the semicircle law is asymptotic to $\frac{\sqrt{2-x}}{\pi}dx$ as $x \rightarrow 2$. The above approximation and the rigidity property suggest that,

$$a_i \asymp -i^{2/3} \quad \text{as } i, N \rightarrow \infty \text{ satisfying } i \leq N. \quad (3.12)$$

For proofs throughout this paper, we need a more rigorous version of the approximation above, which we obtain in the following lemma.

Lemma 3.2. (adapted from [16]) *There exist some integer K and some $c > 0$, which do not depend on N such that, for all $k > K$, we have*

$$\mathbb{P} \left(\bigcup_{N^{2/5} \geq j \geq k} \{a_1 - a_j \geq cj^{2/3}\} \right) \geq 1 - \frac{2}{k^{1/2}}. \quad (3.13)$$

Proof. In line (6.33) of [16], Landon and Sosoë obtain the result that there exists some K_1 (not depending on N) such that, for all $k > K_1$,

$$\mathbb{P} \left(\bigcap_{N^{2/5} \geq j \geq k} \left\{ N^{2/3}(\lambda_j - 2) \leq -\left(\frac{3\pi j}{2}\right)^{2/3} + \frac{1}{10}j^{2/3} \right\} \right) \geq 1 - \frac{1}{k^{1/2}}. \quad (3.14)$$

(Note that the original statement of this inequality in the arxiv version of [16] contains a typo, but the result above is what follows from the preceding lines of [16] and we confirmed this with the authors.) Next, we observe that there exists some K' such that, for all $k > K'$, we have

$$\mathbb{P} \left(N^{2/3}(2 - \lambda_1) \leq \frac{1}{10}k^{2/3} \right) \geq 1 - \frac{1}{k^{1/2}} \quad (3.15)$$

for N sufficiently large. This comes from the fact that the GOE Tracy-Widom distribution has sub-exponential tails. Neither K_1 nor K' depends on N , so we take K to be the maximum of these two values and, combining (3.14) and (3.15), we conclude the desired result. \square

Special sums

There are a few sums that will be particularly important throughout this paper. Below we present some convergence results for these sums, which depend upon the random matrix properties described above.

In particular, for $m = 1, 2, \dots$, we consider sums of the form

$$\frac{1}{N} \sum_{i=2}^N \frac{1}{(\lambda_1 - \lambda_i)^m}. \quad (3.16)$$

We need an asymptotic formula for $m = 1$ as $N \rightarrow \infty$. This was obtained recently in [16]. Landon and Sosoë proved that

$$\Xi_N := N^{1/3} \left(\frac{1}{N} \sum_{i=2}^N \frac{1}{\lambda_1 - \lambda_i} - 1 \right) \Rightarrow \Xi \quad (3.17)$$

for a random variable Ξ as $N \rightarrow \infty$. The limiting random variable Ξ can be expressed in terms of the GOE Airy kernel point process as

$$\Xi = \lim_{n \rightarrow \infty} \left(\sum_{i=2}^n \frac{1}{\alpha_1 - \alpha_i} - \frac{1}{\pi} \int_0^{(\frac{3\pi n}{2})^{2/3}} \frac{dx}{\sqrt{x}} \right) \quad (3.18)$$

where the limit exists almost surely.

We also need another version of the result (3.17) where the constant numerators are replaced n_i^2 :

$$N^{1/3} \left(\frac{1}{N} \sum_{i=2}^N \frac{n_i^2}{\lambda_1 - \lambda_i} - 1 \right) \Rightarrow \lim_{n \rightarrow \infty} \left(\sum_{i=2}^n \frac{\nu_i^2}{\alpha_1 - \alpha_i} - \frac{1}{\pi} \int_0^{(\frac{3\pi n}{2})^{2/3}} \frac{dx}{\sqrt{x}} \right) \quad (3.19)$$

where ν_i are i.i.d standard Gaussians, independent of the GOE Airy point process α_i . This follows from (3.17) and the fact that

$$\frac{1}{N^{2/3}} \sum_{i=2}^N \frac{n_i^2 - 1}{\lambda_1 - \lambda_i} \Rightarrow \sum_{i=2}^{\infty} \frac{\nu_i^2 - 1}{\alpha_1 - \alpha_i} \quad (3.20)$$

which is a convergent series due to Kolmogorov's three series theorem and Lemma 3.2.

Next, we have two lemmas concerning the convergence of special sums.

Lemma 3.3. For any $\delta > 0$,

$$\frac{1}{N} \sum_{i=2}^N \frac{1}{\lambda_1 - \lambda_i} = 1 + O(N^{-\frac{1}{3}+\delta}) \quad \text{and} \quad \frac{1}{N} \sum_{i=2}^N \frac{n_i^2}{\lambda_1 - \lambda_i} = 1 + O(N^{-\frac{1}{3}+\delta}) \quad (3.21)$$

with probability at least $1 - N^{-\delta/2}$. (This lemma is adapted from a similar result in [17]).

Proof. Define an event

$$E_\delta := \left\{ \lambda_1 - \lambda_2 \geq N^{-\frac{2}{3}(1+\delta)} \right\} \cap \left\{ \bigcap_{i=1}^N \left\{ |\lambda_i - \hat{\lambda}_i| \leq N^{-\frac{2}{3}+\delta} (\min\{i, N+1-i\})^{-1/3} \right\} \right\}. \quad (3.22)$$

The first equation in (3.21) holds on this event, which we can see by writing

$$\begin{aligned} \frac{1}{N} \sum_{i=2}^N \frac{1}{\lambda_1 - \lambda_i} &= \frac{1}{N} \sum_{i=2}^{N^{\delta/3}} \frac{1}{\lambda_1 - \lambda_i} + \frac{1}{N} \sum_{i=N^{\delta/3}+1}^N \frac{1}{\lambda_1 - \lambda_i} \\ &= O(N^{-\frac{1}{3}+\delta}) + \left(1 + O(N^{-\frac{1}{3}+\delta}) \right) \end{aligned} \quad (3.23)$$

where, for the first sum, we use $\lambda_1 - \lambda_i \geq N^{-\frac{2}{3}(1+\delta)}$ and, for the second sum, we use eigenvalue rigidity and the semicircle law. The second equation in (3.21) also holds on E_δ using the same reasoning along with the fact that the sum in (3.20) is convergent. It remains only to show that

$$\mathbb{P}(E_\delta) \geq 1 - N^{-\delta/2}. \quad (3.24)$$

From Lemma 3.4 from [16], we have

$$\mathbb{P}(\lambda_1 - \lambda_2 \geq N^{-\frac{2}{3}(1+\delta)}) \geq 1 - N^{-\frac{2}{3}\delta+\delta'} \quad (3.25)$$

for any $\delta' > 0$. This, along with (3.8), implies the lemma. \square

We also consider a similar class of sums with a larger exponent in the denominator and get the following lemma.

Lemma 3.4. For any $\delta > 0$

$$\sum_{i=2}^N \frac{1}{(a_1 - a_i)^m} = O(N^\delta) \quad \text{and} \quad \sum_{i=2}^N \frac{n_i^2}{(a_1 - a_i)^m} = O(N^\delta), \quad m \geq 2, \quad (3.26)$$

with probability at least $1 - N^{-\frac{\delta}{3m}}$.

Proof. To prove the first of these inequalities we consider the event

$$\mathcal{F}_{\delta,m} = \left\{ a_1 - a_2 > N^{-\frac{\delta}{2m}} \right\} \cap \left\{ \bigcap_{i=1}^N \left\{ |\lambda_i - \hat{\lambda}_i| \leq N^{-\frac{2}{3}+\delta} (\min\{i, N+1-i\})^{-1/3} \right\} \right\} \quad (3.27)$$

The event $\left\{ a_1 - a_2 > N^{-\frac{\delta}{2m}} \right\}$ occurs with probability at least $1 - N^{-\frac{\delta}{2m}+e'}$ for any $\delta' > 0$ (see [16] Lemma 3.4). Using this fact along with the eigenvalue rigidity result (3.8), we can conclude that the event $\mathcal{F}_{\delta,m}$ occurs with probability at least $1 - N^{-\frac{\delta}{3m}}$. Now we show that the first inequality in (3.4) holds on the event $\mathcal{F}_{\delta,m}$. In particular, on that event, we have

$$\begin{aligned} \sum_{i=2}^N \frac{1}{(a_1 - a_i)^m} &= \sum_{i=2}^{\lfloor N^{\delta/2} \rfloor} \frac{1}{(a_1 - a_i)^m} + \sum_{i=\lfloor N^{\delta/2} \rfloor+1}^N \frac{1}{(a_1 - a_i)^m} \\ &< N^{\delta/2} \cdot \frac{1}{(N^{-\frac{\delta}{2m}})^m} + 2 \sum_{i=\lfloor N^{\delta/2} \rfloor+1}^N \frac{1}{(-a_i)^m} \\ &\leq N^{\delta} + 2 \sum_{i=\lfloor N^{\delta/2} \rfloor+1}^N \frac{1}{(-\hat{a}_i)^m} \left(1 + \frac{|a_i^m - \hat{a}_i^m|}{(-a_i)^m} \right) \\ &< N^{\delta} + 4 \sum_{i=\lfloor N^{\delta/2} \rfloor+1}^N \frac{1}{(-\hat{a}_i)^m} \end{aligned} \quad (3.28)$$

The summation in the last line is well approximated by the integral

$$N^{-\frac{2m}{3}+1} \int_{-2}^{\hat{\lambda}_{\lfloor N^{\delta/2} \rfloor+1}} \frac{1}{(2-x)^m} d\sigma_{sc}(x) < 4N^{-\frac{2m}{3}+1} \int_{-2}^{\hat{\lambda}_{\lfloor N^{\delta/2} \rfloor+1}} \frac{1}{(2-x)^{m-\frac{1}{2}}} dx \quad (3.29)$$

Using the approximation $2 - \hat{\lambda}_{\lfloor N^{\delta/2} \rfloor+1} \approx cN^{-\frac{2}{3}+\frac{\delta}{3}}$ from (3.11), we see that the right hand side of the inequality above is of order $N^{-m\delta/3}$. Thus $\sum_{i=2}^N \frac{1}{(a_1 - a_i)^m} = O(N^{\delta})$ on the event $\mathcal{F}_{\delta,m}$. Because n_i are standard Gaussians, the sum $\sum_{i=2}^N \frac{n_i^2}{(a_1 - a_i)^m}$ has the same order with comparable probability. \square

Chi-squared distributions

One quantity that we make use of throughout this paper is $n_1 = \mathbf{u}_1^T \mathbf{g}$. We note that n_1 has a standard normal distribution which means that n_1^2 has a chi-squared distribution with one degree of freedom. We prove many results that hold on the event where n_1^2 is roughly of order 1. More specifically, we have the following lemma

Lemma 3.5. For any sufficiently small $\delta > 0$,

$$\mathbb{P} \left(N^{-\delta} < n_1^2 < \delta \log N \right) \geq 1 - N^{-\delta/2} \quad (3.30)$$

The proof of this lemma is straightforward from the probability density function for chi-squared random variables. We note for the purpose of future results that n_1 is independent of the eigenvalues of M .

3.3 Defining the event on which our result holds

For any $\varepsilon > 0$, we define an event \mathcal{E}_ε as follows:

$$\begin{aligned} \mathcal{E}_\varepsilon := & \{N^{-\varepsilon} < n_1^2 < \varepsilon \log N\} \cap \\ & \left\{ \frac{1}{N} \sum_{i=2}^N \frac{1}{\lambda_1 - \lambda_i} = 1 + O(N^{-\frac{1}{3}+\varepsilon}) \text{ and } \frac{1}{N} \sum_{i=2}^N \frac{n_i^2}{\lambda_1 - \lambda_i} = 1 + O(N^{-\frac{1}{3}+\varepsilon}) \right\} \\ & \cap \left\{ \sum_{i=2}^N \frac{1}{(a_1 - a_i)^m} \leq N^\varepsilon \text{ and } \sum_{i=2}^N \frac{n_i^2}{(a_1 - a_i)^m} \leq N^\varepsilon \text{ for } m = 2, 3 \right\} \end{aligned} \quad (3.31)$$

Lemma 3.6. For $\varepsilon > 0$ sufficiently small and N sufficiently large,

$$\mathbb{P}(\mathcal{E}_\varepsilon) \geq 1 - N^{-\varepsilon/10} \quad (3.32)$$

Proof. The event \mathcal{E}_ε as defined above is the intersection of three events, each with probability close to 1. For sufficiently large N , we know from Lemma 3.5 that the first event in the intersection has probability at least $1 - N^{-\varepsilon/2}$ and, from Lemma 3.3, the second event in the intersection has probability at least $1 - N^{-\varepsilon/2}$. The third event in the intersection is actually composed of two events, the one for $m = 2$ and the one for $m = 3$. By Lemma 3.4, these hold with probability $1 - N^{-\varepsilon/6}$ and $1 - N^{-\varepsilon/9}$ respectively. Putting these together, we see that, even if the complements of all of these events are disjoint, we have $\mathbb{P}(\mathcal{E}_\varepsilon) \geq 1 - N^{-\varepsilon/10}$ for any sufficiently small $\varepsilon > 0$ and sufficiently large N . \square

Throughout the rest of this paper, we will prove various results assuming that we are on the event \mathcal{E}_ε . We can then conclude that those results hold with probability at least $1 - N^{-\varepsilon/10}$.

4 Proof of Theorem 2.1

In the proof of Theorem 2.1, we make use of Lemma 3.1, which can be restated as follows:

$$\langle e^{\beta \xi \sqrt{N} \mathfrak{M}} \rangle = e^{\frac{N}{2}(\mathcal{G}_{\mathfrak{M}}(\gamma_{\mathfrak{M}}) - \mathcal{G}(\gamma))} \frac{\int_{\gamma_{\mathfrak{M}} - i\infty}^{\gamma_{\mathfrak{M}} + i\infty} e^{\frac{N}{2}(\mathcal{G}_{\mathfrak{M}}(z) - \mathcal{G}_{\mathfrak{M}}(\gamma_{\mathfrak{M}}))} dz}{\int_{\gamma - i\infty}^{\gamma + i\infty} e^{\frac{N}{2}(\mathcal{G}(z) - \mathcal{G}(\gamma))} dz} \quad (4.1)$$

where

$$\mathcal{G}_{\mathfrak{M}}(z) = \beta z - \frac{1}{N} \sum_{i=1}^N \log(z - \lambda_i) + \frac{(h + \frac{\xi}{\sqrt{N}})^2 \beta}{N} \sum_{i=1}^N \frac{n_i^2}{z - \lambda_i} \quad (4.2)$$

and we use γ and $\gamma_{\mathfrak{M}}$ to denote critical points of $\mathcal{G}(z)$ and $\mathcal{G}_{\mathfrak{M}}(z)$ respectively, which satisfy $\gamma > \lambda_1$ and $\gamma_{\mathfrak{M}} > \lambda_1$. In the next two lemmas, we show that these critical points are unique and we compute upper and lower bounds for them. After accomplishing this, we turn to the more delicate task of computing the integrals in the formula for the generating function of \mathfrak{M} . This is more difficult for $h \sim N^{-1/2}$ than in the other scaling regimes because the critical point is very close to a branch point. Since a straightforward application of Taylor approximation and steepest descent analysis does not work in this case, we directly compute the integral in a neighborhood of the critical point and then show that the tails of the integral are of smaller order.

4.1 Critical point analysis

We begin by computing the critical point γ of $\mathcal{G}(z)$. In [2], the authors use the ansatz that $\gamma = \lambda_1 + pN^{-1}$ with $N^{-\delta} < p < N^\delta$ for any $\delta > 0$ and sufficiently large N on some

event whose probability tends to 1 as $N \rightarrow \infty$. Without making any assumption about the order of p , we set

$$\gamma = \lambda_1 + pN^{-1} \quad (4.3)$$

and then prove that the order of p indeed satisfies the ansatz of [2] (in fact we prove something more precise). In particular, we can define p via the formula for $\mathcal{G}'(z)$ and prove the following lemma.

Lemma 4.1. *There exists a unique $p > 0$ satisfying the equation*

$$\mathcal{G}'(\lambda_1 + pN^{-1}) = \beta - \frac{1}{N} \sum_{i=1}^N \frac{1}{\lambda_1 + pN^{-1} - \lambda_i} - \frac{H^2 \beta}{N^2} \sum_{i=1}^N \frac{n_i^2}{(\lambda_1 + pN^{-1} - \lambda_i)^2} = 0 \quad (4.4)$$

and, for any sufficiently small $\varepsilon > 0$ and sufficiently large N , we have $T < p < \varepsilon \log N$ on the event \mathcal{E}_ε , which occurs with probability at least $1 - N^{-\varepsilon/10}$.

Proof. The existence and uniqueness of p can be seen from the fact that, for $x \in (\lambda_1, \infty)$, the function $\mathcal{G}'(x)$ is monotonically increasing with $\mathcal{G}'(x) \rightarrow \beta$ as $x \rightarrow \infty$ and $\mathcal{G}'(x) \rightarrow -\infty$ as $x \rightarrow \lambda_1$. Having established that (4.4) has a unique solution $p > 0$, we turn to the task of bounding p . On the event \mathcal{E}_ε , the last sum in equation (4.4) is $O(N^{-\frac{2}{3}+\varepsilon})$ for any sufficiently small $\varepsilon > 0$. From this, we get

$$\beta - \frac{1}{N} \sum_{i=2}^N \frac{1}{\lambda_1 - \lambda_i} - \frac{1}{p} - \frac{H^2 \beta n_1^2}{p^2} + O(N^{-\frac{2}{3}+\varepsilon}) < 0 < \beta - \frac{1}{p} - \frac{H^2 \beta n_1^2}{p^2} \quad (4.5)$$

on \mathcal{E}_ε . Further applying definition of \mathcal{E}_ε to the sum on the left hand side and rearranging terms, we get

$$\beta - 1 + O(N^{-\frac{1}{3}+\varepsilon}) < \frac{1}{p} + \frac{H^2 \beta n_1^2}{p^2} < \beta. \quad (4.6)$$

Hence, on \mathcal{E}_ε , the expression $\frac{1}{p} + \frac{H^2 \beta n_1^2}{p^2}$ is bounded above and below by order 1 quantities. The upper bound ensures that $p > \frac{1}{\beta} = T$ (note this is not a sharp bound). The lower bound on $\frac{1}{p} + \frac{H^2 \beta n_1^2}{p^2}$ ensures that $p = O(\varepsilon \log N)$ provided that $|n_i| = O(\varepsilon \log N)$. Since $|n_i| < (\varepsilon \log N)^{1/2}$ for sufficiently large N on \mathcal{E}_ε , we can definitely ensure that $|n_i| < C\varepsilon \log N$ for any constant C and sufficiently large N . \square

Having proved the lemma, we apply the bounds on the order of p to equation (4.4) and conclude that p satisfies

$$\beta - 1 - \frac{1}{p} - \frac{H^2 \beta n_1^2}{p^2} + O(N^{-\frac{1}{3}+\varepsilon}) = 0 \quad (4.7)$$

with probability $1 - N^{-\varepsilon/10}$. We note that, when $h = HN^{-1/2}$, the equation for $\mathcal{G}_\mathfrak{M}$ is same as the one for \mathcal{G} with H replaced by $H + \xi$. Thus $\gamma_\mathfrak{M} = \lambda_1 + p_\mathfrak{M}N^{-1}$ where $p_\mathfrak{M} > 0$ solves the equation

$$\beta - 1 - \frac{1}{p_\mathfrak{M}} - \frac{(H + \xi)^2 \beta n_1^2}{p_\mathfrak{M}^2} + O(N^{-\frac{1}{3}+\varepsilon}) = 0, \quad (4.8)$$

and the lemma below follows by the same reasoning as in the lemma above.

Lemma 4.2. *There exists a unique $p_\mathfrak{M} > 0$ satisfying the equation*

$$\mathcal{G}'_\mathfrak{M}(\lambda_1 + p_\mathfrak{M}N^{-1}) = \beta - \frac{1}{N} \sum_{i=1}^N \frac{1}{\lambda_1 + p_\mathfrak{M}N^{-1} - \lambda_i} - \frac{(H + \xi)^2 \beta}{N^2} \sum_{i=1}^N \frac{n_i^2}{(\lambda_1 + p_\mathfrak{M}N^{-1} - \lambda_i)^2} = 0 \quad (4.9)$$

and, for any sufficiently small $\varepsilon > 0$ and sufficiently large N , we have $T < p_\mathfrak{M} < \varepsilon \log N$ on the event \mathcal{E}_ε , which occurs with probability at least $1 - N^{-\varepsilon/10}$.

4.2 Contour integral computation

We now consider the ratio of the integrals in the formula (3.4). For the integral in the numerator, we have the following lemma.

Lemma 4.3. For fixed $H > 0$ with $h = HN^{-1/2}$ and $T < 1$

$$\begin{aligned} & \int e^{\frac{N}{2}(\mathcal{G}_{\mathfrak{M}}(z) - \mathcal{G}_{\mathfrak{M}}(\gamma_{\mathfrak{M}}))} dz \\ &= \frac{2i\sqrt{2\pi p_{\mathfrak{M}}} e^{-(\beta-1)p_{\mathfrak{M}} + \frac{1}{2}}}{N\sqrt{\beta-1}} \cosh\left((H+\xi)|n_1|\sqrt{\beta(\beta-1)}\right) \left(1 + O(N^{-\frac{1}{21} + \frac{\varepsilon}{7}})\right) \end{aligned} \quad (4.10)$$

on the event $\mathcal{E}_{\varepsilon}$, which occurs with probability at least $1 - N^{-\varepsilon/10}$ for any sufficiently small $\varepsilon > 0$.

Proof. To compute this integral, we need a formula for $N(\mathcal{G}_{\mathfrak{M}}(z) - \mathcal{G}_{\mathfrak{M}}(\gamma_{\mathfrak{M}}))$ in terms of u where $z = \gamma_{\mathfrak{M}} + uN^{-1}$. We will begin by focusing on the central portion of the integral and then we will handle the tails separately. When we are on the event $\mathcal{E}_{\varepsilon}$ and $|u| = o(N^{\frac{1}{3}-\varepsilon})$, we get the following computation:

$$\begin{aligned} N(\mathcal{G}_{\mathfrak{M}}(z) - \mathcal{G}_{\mathfrak{M}}(\gamma_{\mathfrak{M}})) &= N(\mathcal{G}_{\mathfrak{M}}(z) - \mathcal{G}_{\mathfrak{M}}(\gamma_{\mathfrak{M}}) - \mathcal{G}'_{\mathfrak{M}}(\gamma_{\mathfrak{M}})uN^{-1}) \\ &= -\sum_{i=1}^N \left[\log\left(1 + \frac{uN^{-1}}{\gamma_{\mathfrak{M}} - \lambda_i}\right) - \frac{uN^{-1}}{\gamma_{\mathfrak{M}} - \lambda_i} \right] + \frac{(H+\xi)^2\beta}{N} \sum_{i=1}^N \frac{n_i^2 u^2 N^{-2}}{(\gamma_{\mathfrak{M}} + uN^{-1} - \lambda_i)(\gamma_{\mathfrak{M}} - \lambda_i)^2} \\ &= -\log\left(1 + \frac{u}{p_{\mathfrak{M}}}\right) + \frac{u}{p_{\mathfrak{M}}} + O\left(\sum_{j=2}^N \frac{|u|^2 N^{-2}}{(\gamma_{\mathfrak{M}} - \lambda_j)^2}\right) \\ &\quad + \frac{(H+\xi)^2\beta n_1^2 u^2}{(p_{\mathfrak{M}} + u)p_{\mathfrak{M}}^2} + O\left(\sum_{j=2}^N \frac{|u|^2 N^{-3}}{(\gamma_{\mathfrak{M}} - \lambda_j)^3}\right). \end{aligned} \quad (4.11)$$

Using properties of the event $\mathcal{E}_{\varepsilon}$, the quantity in the last line above can be simplified as follows:

$$\begin{aligned} &= -\log\left(1 + \frac{u}{p_{\mathfrak{M}}}\right) + \frac{u}{p_{\mathfrak{M}}} + \frac{(H+\xi)^2\beta n_1^2 u^2}{(p_{\mathfrak{M}} + u)p_{\mathfrak{M}}^2} + O(|u|^2 N^{-\frac{2}{3} + \varepsilon}) \\ &= -\log\left(1 + \frac{u}{p_{\mathfrak{M}}}\right) + (\beta - 1 + O(N^{-\frac{1}{3} + \varepsilon}))(u - p_{\mathfrak{M}}) + 1 + \frac{(H+\xi)^2\beta n_1^2}{(p_{\mathfrak{M}} + u)} + O(|u|^2 N^{-\frac{2}{3} + \varepsilon}) \\ &= -\log\left(1 + \frac{u}{p_{\mathfrak{M}}}\right) + (\beta - 1)(u - p_{\mathfrak{M}}) + 1 + \frac{(H+\xi)^2\beta n_1^2}{(p_{\mathfrak{M}} + u)} + O(|u| + 1)N^{-\frac{1}{3} + \varepsilon}. \end{aligned} \quad (4.12)$$

Now, set $g(u) = \frac{1}{2} \left(-\log\left(1 + \frac{u}{p_{\mathfrak{M}}}\right) + (\beta - 1)(u - p_{\mathfrak{M}}) + 1 + \frac{(H+\xi)^2\beta n_1^2}{(p_{\mathfrak{M}} + u)} \right)$ and let $0 < \delta < \frac{1}{6} - \frac{\varepsilon}{2}$. Then we see that, on the event $\mathcal{E}_{\varepsilon}$,

$$\begin{aligned} & \int_{\gamma_{\mathfrak{M}} - i\infty}^{\gamma_{\mathfrak{M}} + i\infty} \exp\left[\frac{N}{2}(\mathcal{G}_{\mathfrak{M}}(z) - \mathcal{G}_{\mathfrak{M}}(\gamma_{\mathfrak{M}}))\right] dz \\ &= \frac{1}{N} \left(\int_{-iN^{\delta}}^{iN^{\delta}} \exp\left(g(u) + O(|u| + 1)N^{-\frac{1}{3} + \varepsilon}\right) du + \text{integrals of tails} \right). \end{aligned} \quad (4.13)$$

For the purposes of computing this, it helps to deform the contour by shifting it leftward so that, instead of the vertical contour from $\gamma_{\mathfrak{M}} - i\infty$ to $\gamma_{\mathfrak{M}} + i\infty$, we consider a contour

from $\lambda_1 - i\infty$ to $\lambda_1 + i\infty$ which is a straight vertical line except near λ_1 where it passes to the right of λ_1 . The integral on this contour will be the same as on the original contour and we get

$$\begin{aligned} & \int_{\lambda_1 - i\infty}^{\lambda_1 + i\infty} \exp \left[\frac{N}{2} (\mathcal{G}_{\mathfrak{M}}(z) - \mathcal{G}_{\mathfrak{M}}(\gamma_{\mathfrak{M}})) \right] dz \\ &= \frac{1}{N} \left(\int_{-iN^\delta}^{iN^\delta} \exp \left(g(u - p_{\mathfrak{M}}) + O \left((|u| + 1)N^{-1/3} \right) \right) du + \text{integrals of tails} \right) \\ &= \frac{1}{N} \left(\int_{-iN^\delta}^{iN^\delta} \exp(g(u - p_{\mathfrak{M}})) \left(1 + O \left((|u| + 1)N^{-\frac{1}{3} + \varepsilon} \right) \right) du + \text{integrals of tails} \right). \end{aligned} \quad (4.14)$$

Note that we have implemented a leftward shift of $\gamma_{\mathfrak{M}} - \lambda_1$ with respect to z , which corresponds to a leftward shift of $p_{\mathfrak{M}}$ with respect to u . We make this shift in the integrand rather than in the contour bounds, so the contour with respect to u is a vertical line along the imaginary axis, but passing to the right of the origin. Next, we compute the integral on the portion of the contour from $-iN^\delta$ to iN^δ . Call this portion of the contour C . We define C more specifically (in terms of u) to be composed of three pieces:

- C_1 is the straight line from $-iN^\delta$ to $-ip_{\mathfrak{M}}$.
- C_2 is the semicircle given by $p_{\mathfrak{M}}e^{i\theta}$ with $\theta \in [-\frac{\pi}{2}, \frac{\pi}{2}]$.
- C_3 is the straight line from $ip_{\mathfrak{M}}$ to iN^δ .

We show that $\exp(g(u - p_{\mathfrak{M}}))$ is bounded on C_1, C_2, C_3 by bounding the real part of $g(u - p_{\mathfrak{M}})$. On C_1 and C_3 , we have

$$\operatorname{Re}(g(u - p_{\mathfrak{M}})) = -\frac{1}{2} \log \left(\frac{|u|}{p_{\mathfrak{M}}} \right) - 2p_{\mathfrak{M}}(\beta - 1) + 1 < 1. \quad (4.15)$$

On C_2 , we have

$$\begin{aligned} \operatorname{Re}(g(p_{\mathfrak{M}}e^{i\theta} - p_{\mathfrak{M}})) &= -\frac{1}{2} \log(|e^{i\theta}|) + (\beta - 1) \cdot \operatorname{Re}(p_{\mathfrak{M}}e^{i\theta} - 2p_{\mathfrak{M}}) + 1 + \operatorname{Re} \left(\frac{(H + \xi)^2 \beta n_1^2}{p_{\mathfrak{M}}e^{i\theta}} \right) \\ &< 1 + \frac{(H + \xi)^2 \beta n_1^2}{p_{\mathfrak{M}}}. \end{aligned} \quad (4.16)$$

Since the real part of $g(u - p_{\mathfrak{M}})$ is bounded, the magnitude of $\exp(g(u - p_{\mathfrak{M}}))$ is also bounded by some constant (call it c) so we have

$$\begin{aligned} & \int_{C_1} \exp(g(u - p_{\mathfrak{M}})) \left(1 + O \left((|u| + 1)N^{-\frac{1}{3} + \varepsilon} \right) \right) du \\ &= \int_{-iN^\delta}^{-ip_{\mathfrak{M}}} \exp(g(u - p_{\mathfrak{M}})) \left(1 + O \left((|u| + 1)N^{-\frac{1}{3} + \varepsilon} \right) \right) du \\ &= \int_{-iN^\delta}^{-ip_{\mathfrak{M}}} \exp(g(u - p_{\mathfrak{M}})) du + O(c \cdot 2N^{2\delta - \frac{1}{3} + \varepsilon}) \\ &= \int_{C_1} \exp(g(u - p_{\mathfrak{M}})) du + O(N^{2\delta - \frac{1}{3} + \varepsilon}). \end{aligned} \quad (4.17)$$

Similarly, we have

$$\int_{C_3} \exp(g(u - p_{\mathfrak{M}})) \left(1 + O \left((|u| + 1)N^{-\frac{1}{3} + \varepsilon} \right) \right) du = \int_{C_3} \exp(g(u - p_{\mathfrak{M}})) du + O(N^{2\delta - \frac{1}{3} + \varepsilon}). \quad (4.18)$$

Finally, for C_2 , we get

$$\begin{aligned}
 & \int_{C_2} \exp(g(u - p_{\mathfrak{M}})) \left(1 + O\left(|u| + 1\right) N^{-\frac{1}{3} + \varepsilon}\right) du \\
 &= \int_{-\pi/2}^{\pi/2} \exp(g(p_{\mathfrak{M}} e^{i\theta} - p_{\mathfrak{M}})) \left(1 + O\left(N^{-\frac{1}{3} + \varepsilon}\right)\right) p_{\mathfrak{M}} i e^{i\theta} d\theta \\
 &= \int_{-\pi/2}^{\pi/2} \exp(g(p_{\mathfrak{M}} e^{i\theta} - p_{\mathfrak{M}})) p_{\mathfrak{M}} i e^{i\theta} d\theta + O\left(p_{\mathfrak{M}} c \pi N^{-\frac{1}{3} + \varepsilon}\right) \\
 &= \int_{C_2} \exp(g(u - p_{\mathfrak{M}})) du + O(N^{-\frac{1}{3} + \varepsilon}).
 \end{aligned} \tag{4.19}$$

Thus, we conclude that, on the event \mathcal{E}_ε , for any $0 < \delta < \frac{1}{6} - \frac{\varepsilon}{2}$,

$$\begin{aligned}
 & \int_{-iN^\delta}^{iN^\delta} \exp(g(u - p_{\mathfrak{M}})) \left(1 + O\left(|u| + 1\right) N^{-\frac{1}{3} + \varepsilon}\right) du \\
 &= \int_{-iN^\delta}^{iN^\delta} \exp(g(u - p_{\mathfrak{M}})) du + O(N^{2\delta - \frac{1}{3} + \varepsilon}).
 \end{aligned} \tag{4.20}$$

We use lemma 5.1 to show that the integral of the tails has order $O(N^{-\delta/3})$. This has the same order as $O(N^{2\delta - \frac{1}{3} + \varepsilon})$ when $\delta = \frac{1}{7}(1 - 3\varepsilon)$, which is positive for any $0 < \varepsilon < \frac{1}{3}$. Since we are free to choose any $0 < \delta < \frac{1}{6} - \frac{\varepsilon}{2}$, we set $\delta = \frac{1}{7}(1 - 3\varepsilon)$ and get

$$\begin{aligned}
 & \int e^{\frac{N}{2}(\mathcal{G}_{\mathfrak{M}}(z) - \mathcal{G}_{\mathfrak{M}}(\gamma_{\mathfrak{M}}))} dz = \frac{1}{N} \left(\int_{-iN^\delta}^{iN^\delta} \exp(g(u - p_{\mathfrak{M}})) du + O\left(N^{-\frac{1}{21} + \frac{\varepsilon}{7}}\right) \right) \\
 &= \frac{1}{N} \left(\int_{-iN^\delta}^{iN^\delta} \sqrt{\frac{p_{\mathfrak{M}}}{u}} e^{\frac{(\beta-1)(u-2p_{\mathfrak{M}})}{2} + \frac{(H+\xi)^2 \beta n_1^2}{2u}} du + O\left(N^{-\frac{1}{21} + \frac{\varepsilon}{7}}\right) \right) \\
 &= \frac{p_{\mathfrak{M}}^{1/2} e^{-(\beta-1)p_{\mathfrak{M}} + \frac{1}{2}}}{N} \left(\int_{-iN^\delta}^{iN^\delta} \frac{1}{\sqrt{u}} e^{\frac{(\beta-1)u}{2} + \frac{(H+\xi)^2 \beta n_1^2}{2u}} du + O\left(N^{-\frac{1}{21} + \frac{\varepsilon}{7}}\right) \right).
 \end{aligned} \tag{4.21}$$

The integral $\int_{0_+ + i\mathbb{R}} \frac{1}{\sqrt{u}} \exp\left(\frac{(\beta-1)u}{2} + \frac{(H+\xi)^2 \beta n_1^2}{2u}\right) du$ can be evaluated using the contour integral formula for the modified Bessel function (see e.g. [1]):

$$\int_{0_+ + i\mathbb{R}} \frac{1}{\sqrt{w}} e^{aw + \frac{b}{w}} dw = 2\pi i \left(\frac{b}{a}\right)^{1/4} I_{-\frac{1}{2}}(2\sqrt{ab}) = \frac{2i\sqrt{\pi}}{\sqrt{a}} \cosh(2\sqrt{ab}). \tag{4.22}$$

Since this integral converges, the integral in the last line of equation (4.21) must converge to the same value. Furthermore, the tails of the integral in (4.22) beyond order N^δ only contribute $O(N^{-\delta/2})$ to the value of the integral. This can be seen by applying integration by parts to the integral $i \int_{N^\delta}^\infty \frac{1}{\sqrt{iy}} e^{i(ay - \frac{b}{y})} dy$ where the factor $\frac{1}{\sqrt{y}} e^{-\frac{ib}{y}}$ is differentiated and the factor e^{iay} is integrated. This gives us $O(N^{-\delta/2})$, which is less than $O(N^{-\frac{1}{21} + \frac{\varepsilon}{7}})$ since $\delta = \frac{1}{7}(1 - 3\varepsilon)$. Hence, we conclude that, on the event \mathcal{E}_ε ,

$$\begin{aligned}
 & \int e^{\frac{N}{2}(\mathcal{G}_{\mathfrak{M}}(z) - \mathcal{G}_{\mathfrak{M}}(\gamma_{\mathfrak{M}}))} dz \\
 &= \frac{2i\sqrt{2\pi p_{\mathfrak{M}}} e^{-(\beta-1)p_{\mathfrak{M}} + \frac{1}{2}}}{N\sqrt{\beta-1}} \cosh\left((H+\xi)|n_1| \sqrt{\beta(\beta-1)}\right) \left(1 + O(N^{-\frac{1}{21} + \frac{\varepsilon}{7}})\right). \quad \square
 \end{aligned} \tag{4.23}$$

We now return to the task of computing the moment generating function of \mathfrak{M} using the formula in line (4.1). The integral in the denominator of that formula can be viewed

as a special case of the numerator in which $\xi = 0$ and $p_{\mathfrak{M}}$ is replaced with p . Therefore, on the event \mathcal{E}_ε ,

$$\frac{\int e^{\frac{N}{2}(\mathcal{G}_{\mathfrak{M}}(z) - \mathcal{G}_{\mathfrak{M}}(\gamma_{\mathfrak{M}}))} dz}{\int e^{\frac{N}{2}(\mathcal{G}(z) - \mathcal{G}(\gamma))} dz} = \sqrt{\frac{p_{\mathfrak{M}}}{p}} e^{-(\beta-1)(p_{\mathfrak{M}}-p)} \frac{\cosh\left((H+\xi)|n_1|\sqrt{\beta(\beta-1)}\right)}{\cosh\left(H|n_1|\sqrt{\beta(\beta-1)}\right)} \left(1 + O(N^{-\frac{1}{21}+\frac{\varepsilon}{7}})\right). \quad (4.24)$$

To compute the moment generating function of \mathfrak{M} from the formula (4.1), it remains only to evaluate the factor $e^{\frac{N}{2}(\mathcal{G}_{\mathfrak{M}}(\gamma_{\mathfrak{M}}) - \mathcal{G}(\gamma))}$, which is also computed in [2] (see section 8.6 of that paper). Although we require more precision for the order of the sub-leading term than is provided in [2], this can easily be achieved by repeating their steps using the assumptions that hold on the event \mathcal{E}_ε and carefully tracking the order of each term. This yields the result

$$N(\mathcal{G}_{\mathfrak{M}}(\gamma_{\mathfrak{M}}) - \mathcal{G}(\gamma)) = -\log\left(\frac{p_{\mathfrak{M}}}{p}\right) + 2(\beta-1)(p_{\mathfrak{M}}-p) + (2H\xi + \xi^2)\beta + O(N^{-\frac{1}{3}+\varepsilon}). \quad (4.25)$$

Thus we can conclude that

$$\langle e^{\beta\xi\sqrt{N}\mathfrak{M}} \rangle = e^{\frac{(2H\xi+\xi^2)\beta}{2}} \frac{\cosh\left((H+\xi)|n_1|\sqrt{\beta(\beta-1)}\right)}{\cosh\left(H|n_1|\sqrt{\beta(\beta-1)}\right)} \left(1 + O(N^{-\frac{1}{21}+\frac{\varepsilon}{7}})\right). \quad (4.26)$$

Replacing $\beta\xi$ by ξ and using $T = 1/\beta$, we obtain the result stated in Theorem 2.1.

5 Integral Approximation Lemmas

The proof of Theorems 2.1 in the preceding section required us to compute a contour integral. In that computation, we relied on the fact that the dominant contribution to the integral comes from a neighborhood of the critical point. In this section, we prove that fact by providing an upper bound for the value of the integral outside of a neighborhood of the critical point and showing that the upper bound shrinks to zero as $N \rightarrow \infty$. This is the most technical part of the contour integral computations.

Lemma 5.1. *Tail approximation for overlap with external field when $h = H^{-1/2}$ and $T < 1$: For any $\delta > 0$,*

$$\int_{iN^\delta}^{i\infty} \exp\left[\frac{N}{2}(\mathcal{G}_{\mathfrak{M}}(\lambda_1 + uN^{-1}) - \mathcal{G}_{\mathfrak{M}}(\gamma_{\mathfrak{M}}))\right] du = O(N^{-\delta/3}) \quad \text{as } N \rightarrow \infty \quad (5.1)$$

on the event \mathcal{E}_ε .

Proof. First, observe that we are using a vertical contour with real part equal to λ_1 as opposed to the original contour, which had real part equal to γ . This is due to a contour deformation that we did when computing the integral on the central portion of the contour. To show that the integrals of the tails tend to zero, we deform the contour yet again. Instead of the vertical line contour given by $\lambda_1 + i(N^\delta + t)N^{-1}$ with $t \in [0, \infty)$, we consider the contour C_4 given by $\lambda_1 - f(t)N^{-1} + i(N^\delta + t)N^{-1}$ where $f(t) = (t+1)^\Delta - 1$ for some $0 < \Delta < \frac{1}{3}$. We briefly comment on this choice of contour: In order to show decay of the tails of the integral, we want the real part of $z(t)$ to approach $-\infty$ as $t \rightarrow \pm\infty$. However, in order to control the integrand, we want the real part of $|z(t) - \lambda_1|$ to be smaller than $\lambda_1 - \lambda_2$ when $|t| < N$. We choose $f(t) = (t+1)^\Delta - 1$ because we also need $f(0) = 0$ and $f'(t)$ bounded.

To bound $\int_{iN^\delta}^{i\infty} \exp[\frac{N}{2}(\mathcal{G}_{\mathfrak{M}}(\lambda_1 + uN^{-1}) - \mathcal{G}_{\mathfrak{M}}(\gamma_{\mathfrak{M}}))]du$ we observe that

$$\begin{aligned} & \left| \int_{iN^\delta}^{i\infty} \exp \left[\frac{N}{2} (\mathcal{G}_{\mathfrak{M}}(\lambda_1 + uN^{-1}) - \mathcal{G}_{\mathfrak{M}}(\gamma_{\mathfrak{M}})) \right] du \right| \\ &= \int_0^\infty \left| (f'(t) + i) \exp \left[\frac{N}{2} (\mathcal{G}_{\mathfrak{M}}(\lambda_1 + (-f(t) + i(N^\delta + t))N^{-1}) - \mathcal{G}_{\mathfrak{M}}(\gamma_{\mathfrak{M}})) \right] \right| dt \\ &\leq \int_0^\infty \left| -\frac{\Delta}{(t+1)^{1-\Delta}} + i \right| \cdot \left| \exp \left[\frac{N}{2} (\mathcal{G}_{\mathfrak{M}}(\lambda_1 + (-f(t) + i(N^\delta + t))N^{-1}) - \mathcal{G}_{\mathfrak{M}}(\gamma_{\mathfrak{M}})) \right] \right| dt \\ &\leq \int_0^\infty \sqrt{2} \left| \exp \left[\frac{N}{2} (\mathcal{G}_{\mathfrak{M}}(\lambda_1 + (-f(t) + i(N^\delta + t))N^{-1}) - \mathcal{G}_{\mathfrak{M}}(\gamma_{\mathfrak{M}})) \right] \right| dt \end{aligned} \quad (5.2)$$

Thus, it suffices to show $\int_0^\infty \left| \exp \left[\frac{N}{2} (\mathcal{G}_{\mathfrak{M}}(\lambda_1 + (-f(t) + i(N^\delta + t))N^{-1}) - \mathcal{G}_{\mathfrak{M}}(\gamma_{\mathfrak{M}})) \right] \right| du = O(N^{-\delta/3})$. We use the notation $\mathcal{G}_{\mathfrak{M}}(z) = A(z) + B(z)$ where

$$A(z) = \beta z - \frac{1}{N} \sum_{j=1}^N \log(z - \lambda_j) \quad B(z) = \frac{(H + \xi)^2 \beta}{N^2} \sum_{j=1}^N \frac{n_j^2}{z - \lambda_j} \quad (5.3)$$

We begin by noting that

$$\begin{aligned} & \int_0^\infty \left| \exp \left[\frac{N}{2} (\mathcal{G}_{\mathfrak{M}}(z(t)) - \mathcal{G}_{\mathfrak{M}}(\gamma_{\mathfrak{M}})) \right] \right| dt \\ &\leq \int_0^\infty \left| \exp \left[\frac{N}{2} (A(z(t)) - A(\gamma_{\mathfrak{M}})) \right] \right| \cdot \left| \exp \left[\frac{N}{2} (B(z(t)) - B(\gamma_{\mathfrak{M}})) \right] \right| dt \end{aligned} \quad (5.4)$$

Therefore, in order to show that the integral on the tail has order $O(N^{-\delta/3})$, it is enough to prove the following two things:

- $\int_0^\infty \left| \exp \left(\frac{N}{2} (A(\lambda_1 - f(t)N^{-1} + i(N^\delta + t)N^{-1}) - A(\gamma_{\mathfrak{M}})) \right) \right| dt = O(N^{-\delta/3})$ and
- $\left| \exp \left(\frac{N}{2} (B(\lambda_1 - f(t)N^{-1} + i(N^\delta + t)N^{-1}) - B(\gamma_{\mathfrak{M}})) \right) \right|$ is bounded for $t > 0$.

The integral in the first bullet point can be rewritten as

$$\begin{aligned} & \int_0^\infty \left| \exp \left(\frac{N}{2} (A(\lambda_1 - f(t)N^{-1} + i(N^\delta + t)N^{-1}) - A(\gamma_{\mathfrak{M}})) \right) \right| dt \\ &= \int_0^\infty \exp \left(\frac{N\beta}{2} (\lambda_1 - f(t)N^{-1} - \gamma_{\mathfrak{M}}) \right) \\ &\quad \cdot \left| \exp \left[-\frac{1}{2} \sum_{j=1}^N \log \left(\frac{\lambda_1 - f(t)N^{-1} + i(N^\delta + t)N^{-1} - \lambda_j}{\gamma_{\mathfrak{M}} - \lambda_j} \right) \right] \right| dt \end{aligned} \quad (5.5)$$

and this can be further simplified as

$$\begin{aligned} &= \int_0^\infty \exp \left(-\frac{\beta(p_{\mathfrak{M}} + f(t))}{2} \right) \\ &\quad \cdot \left| \exp \left[-\frac{1}{2} \sum_{j=1}^N \log \left(\frac{\lambda_1 - f(t)N^{-1} + i(N^\delta + t)N^{-1} - \lambda_j}{\gamma_{\mathfrak{M}} - \lambda_j} \right) \right] \right| dt \\ &= \int_0^\infty \exp \left(-\frac{\beta(p_{\mathfrak{M}} + f(t))}{2} \right) \cdot \exp \left[-\frac{1}{2} \sum_{j=1}^N \log \left| \frac{\lambda_1 - f(t)N^{-1} + i(N^\delta + t)N^{-1} - \lambda_j}{\gamma_{\mathfrak{M}} - \lambda_j} \right| \right] dt \end{aligned} \quad (5.6)$$

We begin by showing that this integral restricted to the interval $[2(\gamma_{\mathfrak{M}} - \lambda_N)N, \infty)$ is of order $O(e^{-N/2})$. If we integrate over just the first factor in the expression above, we get

$$\begin{aligned} & \int_{2(\gamma_{\mathfrak{M}} - \lambda_N)N}^{\infty} \exp\left(-\frac{\beta(p_{\mathfrak{M}} + f(t))}{2}\right) dt \\ &= \exp\left(-\frac{\beta(p_{\mathfrak{M}} - 1)}{2}\right) \int_{2(\gamma_{\mathfrak{M}} - \lambda_N)N}^{\infty} \exp\left(-\frac{(t+1)^{\Delta}}{2}\right) dt \\ &= O(\exp(-N^{\Delta})) \end{aligned} \quad (5.7)$$

In the last equality above, we used the fact (see Lemma 4.2) that $p_{\mathfrak{M}} > T$ on the event $\mathcal{E}_{\varepsilon}$. Since this integral converges, it suffices to show that

$\exp\left[-\frac{1}{2} \sum_{j=1}^N \log\left|\frac{\lambda_1 - f(t)N^{-1} + i(N^{\delta} + t)N^{-1} - \lambda_j}{\gamma_{\mathfrak{M}} - \lambda_j}\right|\right]$ is of order $O(e^{-N/2})$ for $t \geq 2(\gamma_{\mathfrak{M}} - \lambda_N)N$.

$$\begin{aligned} & \exp\left[-\frac{1}{2} \sum_{j=1}^N \log\left|\frac{\lambda_1 - f(t)N^{-1} + i(N^{\delta} + t)N^{-1} - \lambda_j}{\gamma_{\mathfrak{M}} - \lambda_j}\right|\right] \\ & \leq \exp\left[-\frac{1}{2} \sum_{j=1}^N \log\left|\frac{(N^{\delta} + t)N^{-1}}{\gamma_{\mathfrak{M}} - \lambda_j}\right|\right] \leq \exp\left[-\frac{1}{2} \sum_{j=1}^N \log\left|\frac{2(\gamma_{\mathfrak{M}} - \lambda_N)}{\gamma_{\mathfrak{M}} - \lambda_j}\right|\right] \\ & \leq \exp\left[-\frac{N}{2} \log(2)\right] < e^{-N/2} \end{aligned} \quad (5.8)$$

Next, we show that the integral is of order $O(N^{-\delta/3})$ on the interval $[0, 2(\gamma_{\mathfrak{M}} - \lambda_N)N]$. For t in this interval we have

$$\begin{aligned} & \exp\left[-\frac{1}{2} \sum_{j=1}^N \log\left|\frac{\lambda_1 - f(t)N^{-1} + i(N^{\delta} + t)N^{-1} - \lambda_j}{\gamma_{\mathfrak{M}} - \lambda_j}\right|\right] \\ &= \exp\left[-\frac{1}{2} \sum_{j=1}^N \log\left|\frac{a_1 - a_j - f(t)N^{-1/3} + i(N^{\delta} + t)N^{-1/3}}{p_{\mathfrak{M}}N^{-1/3} + a_1 - a_j}\right|\right]. \end{aligned} \quad (5.9)$$

To obtain an upper bound for this quantity, we begin with the $j = 1$ term and observe that

$$\exp\left[-\frac{1}{2} \log\left|\frac{-f(t) + i(N^{\delta} + t)}{p_{\mathfrak{M}}}\right|\right] \leq \exp\left[-\frac{1}{2} \log\left|\frac{N^{\delta}}{p_{\mathfrak{M}}}\right|\right] \leq \left(\frac{p_{\mathfrak{M}}}{N^{\delta}}\right)^{1/2}. \quad (5.10)$$

For the summation of the $j \geq 2$ terms, we get

$$\begin{aligned} & \exp\left[-\frac{1}{2} \sum_{j=2}^N \log\left|\frac{a_1 - a_j - f(t)N^{-1/3} + i(N^{\delta} + t)N^{-1/3}}{p_{\mathfrak{M}}N^{-1/3} + a_1 - a_j}\right|\right] \\ & \leq \exp\left[-\frac{1}{4} \sum_{j=2}^N \log\left|\frac{(a_1 - a_j - f(t)N^{-1/3})^2 + (N^{\delta} + t)^2N^{-2/3}}{(p_{\mathfrak{M}}N^{-1/3} + a_1 - a_j)^2}\right|\right] \\ & \leq \exp\left[-\frac{1}{4} \sum_{j=2}^N \log\left|1 + \frac{-2(f(t) + p_{\mathfrak{M}})(a_1 - a_j)N^{-1/3} - p_{\mathfrak{M}}^2N^{-2/3} + (N^{\delta} + t)^2N^{-2/3}}{(p_{\mathfrak{M}}N^{-1/3} + a_1 - a_j)^2}\right|\right]. \end{aligned} \quad (5.11)$$

Because $-p_{\mathfrak{M}}^2N^{-2/3} + (N^{\delta} + t)^2N^{-2/3} > 0$, the last line of this inequality is bounded

above by

$$\begin{aligned}
 & \exp \left[-\frac{1}{4} \sum_{j=2}^N \log \left| 1 - \frac{2(f(t) + p_{\mathfrak{M}})(a_1 - a_j)N^{-1/3}}{(p_{\mathfrak{M}}N^{-1/3} + a_1 - a_j)^2} \right| \right] \\
 & \leq \exp \left[-\frac{1}{4} \sum_{j=2}^N \log \left| 1 - \frac{2(f(t) + p_{\mathfrak{M}})N^{-1/3}}{a_1 - a_j} \right| \right] \\
 & \leq \exp \left[\frac{1}{4} \sum_{j=2}^N \left(\frac{2(f(t) + p_{\mathfrak{M}})N^{-1/3}}{a_1 - a_j} + \left(\frac{2(f(t) + p_{\mathfrak{M}})N^{-1/3}}{a_1 - a_j} \right)^2 \right) \right] \\
 & = \exp \left[\frac{f(t) + p_{\mathfrak{M}}}{2} \sum_{j=2}^N \frac{N^{-1/3}}{a_1 - a_j} + ((f(t) + p_{\mathfrak{M}})N^{-1/3})^2 \sum_{j=2}^N \frac{1}{(a_1 - a_j)^2} \right].
 \end{aligned} \tag{5.12}$$

Next, using the properties of the event \mathcal{E}_ε , we see that the last line above has upper bound

$$\begin{aligned}
 & \exp \left[\frac{f(t) + p_{\mathfrak{M}}}{2} \left((1 + O(N^{-\frac{1}{3}+\varepsilon})) + O(N^{-\frac{2}{3}+\Delta+\varepsilon}) \right) \right] \\
 & = \exp \left[\frac{f(t) + p_{\mathfrak{M}}}{2} \left(1 + O(N^{-\frac{1}{3}+\varepsilon}) \right) \right].
 \end{aligned} \tag{5.13}$$

Combining this with the upper bound from the $j = 1$ term in (5.10), we conclude that

$$\begin{aligned}
 & \exp \left[-\frac{1}{2} \sum_{j=1}^N \log \left| \frac{\lambda_1 - f(t)N^{-1} + i(N^\delta + t)N^{-1} - \lambda_j}{\gamma_{\mathfrak{M}} - \lambda_j} \right| \right] \\
 & \leq \left(\frac{p_{\mathfrak{M}}}{N^\delta} \right)^{1/2} \exp \left[\frac{f(t) + p_{\mathfrak{M}}}{2} \left(1 + O(N^{-\frac{1}{3}+\varepsilon}) \right) \right].
 \end{aligned} \tag{5.14}$$

Finally, plugging this back into the original integral, we get

$$\begin{aligned}
 & \int_0^{2(\gamma_{\mathfrak{M}} - \lambda_N)N} \left| \exp \left(\frac{N}{2} (A(\lambda_1 - f(t)N^{-1} + i(N^\delta + t)N^{-1}) - A(\gamma_{\mathfrak{M}})) \right) \right| dt \\
 & \leq \int_0^{2(\gamma_{\mathfrak{M}} - \lambda_N)N} \exp \left(-\frac{\beta(p_{\mathfrak{M}} + f(t))}{2} \right) \cdot \left(\frac{p_{\mathfrak{M}}}{N^\delta} \right)^{1/2} \exp \left[\frac{f(t) + p_{\mathfrak{M}}}{2} (1 + O(N^{-\frac{1}{3}+\varepsilon})) \right] dt \\
 & = \left(\frac{p_{\mathfrak{M}}}{N^\delta} \right)^{1/2} \int_0^{2(\gamma_{\mathfrak{M}} - \lambda_N)N} \exp \left[-\frac{\beta - 1 + O(N^{-\frac{1}{3}+\varepsilon})}{2} \cdot (f(t) + p_{\mathfrak{M}}) \right] dt
 \end{aligned} \tag{5.15}$$

Since $\beta > 1$, there exists some $C'' > 0$ such that the integral is bounded above by

$$\begin{aligned}
 & \left(\frac{p_{\mathfrak{M}}}{N^\delta} \right)^{1/2} \int_0^{2(\gamma_{\mathfrak{M}} - \lambda_N)N} \exp [-C''((t+1)^\Delta - 1)] dt \\
 & = O \left(\left(\frac{p_{\mathfrak{M}}}{N^\delta} \right)^{1/2} \right) = O \left(\left(\frac{\varepsilon \log N}{N^\delta} \right)^{1/2} \right) = O(N^{-\delta/3})
 \end{aligned} \tag{5.16}$$

Lastly, it remains to show that $\left| \exp \left(\frac{N}{2} (B(\lambda_1 - f(t)N^{-1} + i(N^\delta + t)N^{-1}) - B(\gamma_{\mathfrak{M}})) \right) \right|$ is bounded and it suffices to show that $\operatorname{Re} \left(\frac{N}{2} (B(\lambda_1 - f(t)N^{-1} + i(N^\delta + t)N^{-1}) - B(\gamma_{\mathfrak{M}})) \right)$

is bounded above.

$$\begin{aligned} & \operatorname{Re} \left[\frac{N}{2} (B(\lambda_1 - f(t)N^{-1} + i(N^\delta + t)N^{-1}) - B(\gamma_{\mathfrak{M}})) \right] \\ &= \operatorname{Re} \left[\frac{N}{2} \cdot \frac{(H + \xi)^2 \beta}{N^2} \sum_{j=1}^N \left(\frac{n_j^2}{\lambda_1 - f(t)N^{-1} + i(N^\delta + t)N^{-1} - \lambda_j} - \frac{n_j^2}{\gamma_{\mathfrak{M}} - \lambda_j} \right) \right] \end{aligned} \quad (5.17)$$

We observe that the real part of the $j = 1$ term in the summation is negative and, furthermore, $\sum_{j=2}^N (-\frac{n_j^2}{\gamma_{\mathfrak{M}} - \lambda_j})$ is negative. Removing these terms, we see that the quantity above has upper bound

$$\begin{aligned} & \operatorname{Re} \left[\frac{(H + \xi)^2 \beta}{2N} \sum_{j=2}^N \frac{n_j^2}{\lambda_1 - f(t)N^{-1} + i(N^\delta + t)N^{-1} - \lambda_j} \right] \\ &= \frac{(H + \xi)^2 \beta}{2N} \sum_{j=2}^N \frac{n_j^2 (\lambda_1 - f(t)N^{-1} - \lambda_j)}{(\lambda_1 - f(t)N^{-1} - \lambda_j)^2 + (N^\delta + t)^2 N^{-2}}. \end{aligned} \quad (5.18)$$

Now consider two cases. For $t < N$, the expression in the last line is bounded above by

$$\frac{(H + \xi)^2 \beta}{2N} \sum_{j=2}^N \frac{n_j^2}{\lambda_1 - f(t)N^{-1} - \lambda_j} = \frac{(H + \xi)^2 \beta}{2} \sum_{j=2}^N \frac{n_j^2 N^{-1/3}}{a_1 - a_j - f(t)N^{-1/3}} \quad (5.19)$$

This will be $O(1)$ because $\sum_{j=2}^N \frac{n_j^2 N^{-1/3}}{a_1 - a_j} = 1 + O(N^{-\frac{1}{3} + \varepsilon})$ on the event \mathcal{E}_ε and, for sufficiently small ε , we have $f(t)N^{-1/3} < \frac{1}{2}(a_1 - a_2)$ since $f(t)N^{-1/3} = O(N^{\Delta - \frac{1}{3}})$ where $\Delta < \frac{1}{3}$ and $a_1 - a_2 > N^{-\varepsilon/3}$ on \mathcal{E}_ε .

In the case where $t \geq N$, we instead use the upper bound

$$\frac{(H + \xi)^2 \beta}{2N} \sum_{j=2}^N \frac{n_j^2 (\lambda_1 - \lambda_j)}{(N^\delta + t)^2 N^{-2}} \leq \frac{(H + \xi)^2 \beta}{2N} \sum_{j=2}^N 4n_j^2 \quad (5.20)$$

Since n_j^2 are i.i.d. chi-squared random variables, this sum is $O(1)$ with overwhelming probability. \square

6 Applying this method to the overlap of two replicas

Using a method similar to the proofs in Sections 4 and 5, we can prove Theorem 6.1 for \mathfrak{R} , the overlap of two replicas (this is a rigorous re-formulation of Result 10.6 from [2]). The generating function for the overlap with a replica involves a double integral rather than a single integral, but we can use the same contour as in Sections 4 and 5 for both integrals and then transform to polar coordinates in order to prove the desired decay properties outside a neighborhood of the critical point. While our method works to prove this theorem, we do not provide the details here because it also follows from Theorem 2.14 of [17], as we explain below.

Theorem 6.1. *Given $T < 1$ and $h = HN^{-1/2}$ for some fixed $H \geq 0$, we have the following asymptotic formula for the moment generating function of \mathfrak{R} , the overlap with a replica. This formula holds on the event \mathcal{E}_ε (which has probability at least $1 - N^{-\varepsilon/10}$) for any sufficiently small $\varepsilon > 0$ and $\xi = O(1)$.*

$$\langle e^{\xi \frac{\mathfrak{R}}{1-T}} \rangle = \frac{\cosh \left(\frac{2\sqrt{1-T}H|n_1|}{T} \right) e^\xi + e^{-\xi}}{\cosh \left(\frac{2\sqrt{1-T}H|n_1|}{T} \right) + 1} + O(N^{-\frac{1}{21} + \frac{\varepsilon}{7}}) \quad (6.1)$$

Note that the leading order term on the right hand side is the moment generating function of a shifted Bernoulli random variable that takes values 1 and -1 with probability P and $1 - P$ respectively, where

$$P = \frac{\cosh\left(\frac{2\sqrt{1-T}H|n_1|}{T}\right)}{\cosh\left(\frac{2\sqrt{1-T}H|n_1|}{T}\right) + 1}. \quad (6.2)$$

Thus, for large N , we can conclude that, on the event \mathcal{E}_ε , the overlap \mathfrak{R} behaves in its leading order like a shifted Bernoulli random variable. This conclusion also follows from Theorem 2.14 of [17], which states that, for sufficiently small $\varepsilon > 0$, there exists $\varepsilon_1 > 0$ such that with probability at least $1 - N^{-\varepsilon_1}$ and all $t > 0$,

$$\begin{aligned} & \langle \mathbf{1}_{\{|N^{-1}\sigma^{(1)} \cdot \sigma^{(2)} \mp (1-\beta^{-1})| \leq t\}} \rangle \\ &= \frac{1}{2} \pm \frac{1}{2} \tanh^2\left(\sqrt{v_1^2\theta(\beta-1)}\right) + N^\varepsilon O\left(t + N^{-2/3+\varepsilon}t^{-2} + N^{-1/3}\right). \end{aligned} \quad (6.3)$$

While their theorem is formulated and proved in a different manner than Theorem 6.1, their result implies ours.

References

- [1] Milton Abramowitz and Irene A. Stegun (eds.), *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, Dover Publications, Inc., New York, 1992, Reprint of the 1972 edition. MR0415956
- [2] Jinho Baik, Elizabeth Collins-Woodfin, Pierre Le Doussal, and Hao Wu, *Spherical spin glass model with external field*, J. Stat. Phys. **183** (2021), no. 2. MR4257832
- [3] Jinho Baik and Ji Oon Lee, *Fluctuations of the free energy of the spherical Sherrington-Kirkpatrick model*, J. Stat. Phys. **165** (2016), no. 2, 185–224. MR3554380
- [4] Jinho Baik and Ji Oon Lee, *Fluctuations of the free energy of the spherical Sherrington-Kirkpatrick model with ferromagnetic interaction*, Ann. Henri Poincaré **18** (2017), no. 6, 1867–1917. MR3649446
- [5] Jinho Baik and Ji Oon Lee, *Free energy of bipartite spherical Sherrington-Kirkpatrick model*, Ann. Inst. Henri Poincaré Prob. Stat. **56** (2020), no. 4, 2897–2934. MR4164860
- [6] Jinho Baik, Ji Oon Lee, and Hao Wu, *Ferromagnetic to paramagnetic transition in spherical spin glass*, J. Stat. Phys. **173** (2018), no. 5, 1484–1522. MR3878351
- [7] Wei-Kuo Chen and Arnab Sen, *Parisi formula, disorder chaos and fluctuation for the ground state energy in the spherical mixed p -spin models*, Communications in Mathematical Physics **350** (2016), no. 1, 129–173. MR3606472
- [8] Andrea Crisanti and H-J Sommers, *The spherical p -spin interaction spin glass model: The statics*, Zeitschrift für Physik B Condensed Matter **87** (1992), no. 3, 341–354.
- [9] Leticia F Cugliandolo, David S Dean, and Hajime Yoshino, *Nonlinear susceptibilities of spherical models*, J. Phys. A: Math. Gen. **40** (2007), no. 16, 4285. MR2316706
- [10] Amir Dembo and Ofer Zeitouni, *Matrix optimization under random external fields*, J. Stat. Phys. **159** (2015), no. 6, 1306–1326. MR3350373
- [11] László Erdős, Antti Knowles, Horng-Tzer Yau, and Jun Yin, *Spectral statistics of Erdős-Rényi graphs I: Local semicircle law*, Ann. Probab. **41** (2013), no. 3B, 2279–2375. MR3098073
- [12] László Erdős, Horng-Tzer Yau, and Jun Yin, *Rigidity of eigenvalues of generalized Wigner matrices*, Advances in Mathematics (New York, 1965) **229** (2012), no. 3, 1435–1515. MR2871147
- [13] Yan V. Fyodorov and Pierre le Doussal, *Topology trivialization and large deviations for the minimum in the simplest random optimization*, J. Stat. Phys. **154** (2014), no. 1-2, 466–490. MR3162549

- [14] Pax Kivimae, *Critical fluctuations for the spherical Sherrington-Kirkpatrick model in an external field*, arXiv preprint, arXiv:1908.07512 (2019). MR4283044
- [15] JM Kosterlitz, DJ Thouless, and Raymund C Jones, *Spherical model of a spin-glass*, Phys. Rev. Lett. **36** (1976), no. 20, 1217.
- [16] Benjamin Landon and Philippe Sosoë, *Fluctuations of the overlap at low temperature in the 2-spin spherical SK model*, arXiv preprint arXiv:1905.03317 (2019).
- [17] Benjamin Landon and Philippe Sosoë, *Fluctuations of the 2-spin SSK model with magnetic field*, arXiv preprint arXiv:2009.12514 (2020).
- [18] Madan Lal Mehta, *Random matrices*, third ed., Pure and Applied Mathematics (Amsterdam), vol. 142, Elsevier/Academic Press, Amsterdam, 2004. MR2129906
- [19] Vu Lan Nguyen and Philippe Sosoë, *Central limit theorem near the critical temperature for the overlap in the 2-spin spherical SK model*, J. Math. Phys. **60** (2019), no. 10, 103302, 13. MR4016884
- [20] Dmitry Panchenko and Michel Talagrand, *On the overlap in the multiple spherical SK models*, Ann. Probab. **35** (2007), no. 6, 2321–2355. MR2353390
- [21] Giorgio Parisi, *A sequence of approximated solutions to the SK model for spin glasses*, J. Phys. A: Math. Gen. **13** (1980), no. 4, L115. MR0873974
- [22] D Sherrington and S Kirkpatrick, *Solvable model of a spin-glass*, Phys. Rev. Lett. **35** (1975), no. 26, 1792–1796.
- [23] Alexander Soshnikov, *Universality at the edge of the spectrum in Wigner random matrices*, Comm. Math. Phys. **207** (1999), no. 3, 697–733. MR1727234
- [24] Eliran Subag, *The geometry of the Gibbs measure of pure spherical spin glasses*, Invent. Math. **210** (2017), no. 1, 135–209. MR3698341
- [25] Michel Talagrand, *Free energy of the spherical mean field model*, Probab. Theory Related Fields **134** (2006), no. 3, 339–382. MR2226885
- [26] Michel Talagrand, *The Parisi formula*, Ann. of Math. (2) **163** (2006), no. 1, 221–263. MR2195134
- [27] Craig A. Tracy and Harold Widom, *Level-spacing distributions and the Airy kernel*, Comm. Math. Phys. **159** (1994), no. 1, 151–174. MR1257246

Acknowledgments. The author would like to thank Jinho Baik for all his helpful advice during the preparation of this paper. The author would also like to thank Benjamin Landon for answering questions about his paper and Jeffrey Lagarias for providing support through his NSF grant.