# Maximum of the characteristic polynomial of random unitary matrices


Louis-Pierre Arguin
*Department of Mathematics, Baruch College*
*Graduate Center, City University of New York*
*louis-pierre.arguin@baruch.cuny.edu*

David Belius
*Courant Institute,*
*New York University*
*david.belius@cantab.net*

Paul Bourgade
*Courant Institute,*
*New York University*
*bourgade@cims.nyu.edu*



It was recently conjectured by Fyodorov, Hiary and Keating that the maximum of the characteristic polynomial on the unit circle of a $N\times N$ random unitary matrix sampled from the Haar measure grows like $CN/(\log N)^{3/4}$ for some random variable $C$. In this paper, we verify the leading order of this conjecture, that is, we prove that with high probability the maximum lies in the range $[N^{1-\varepsilon},N^{1+\varepsilon}]$, for arbitrarily small $\varepsilon$.

The method is based on identifying an approximate branching random walk in the Fourier decomposition of the characteristic polynomial, and uses techniques developed to describe the extremes of branching random walks and of other log-correlated random fields. A key technical input is the asymptotic analysis of Toeplitz determinants with dimension-dependent symbols.

The original argument for these asymptotics followed the general idea that the statistical mechanics of $1/f$-noise random energy models is governed by a freezing transition. We also prove the conjectured freezing of the free energy for random unitary matrices.




## 1 Introduction

For $N\in\mathbb{N}$, consider a random matrix $\mathrm{U}_N$ sampled from the group of $N\times N$ unitary matrices with Haar measure. This distribution is also known as the *Circular Unitary Ensemble* (CUE). This paper studies the extreme values of the characteristic polynomial $\mathrm{P}_N$ of $\mathrm{U}_N$, on the unit circle, as $N\to\infty$. The main result concerns the asymptotics of

$$\max_{h\in[0,2\pi]}|\mathrm{P}_N(e^{\mathrm{i}h})|=\max_{h\in[0,2\pi]}|\det(e^{\mathrm{i}h}-\mathrm{U}_N)|.$$

It was shown by Keating and Snaith [44] that for a fixed $h$, $\log|\mathrm{P}_N(e^{\mathrm{i}h})|$, converges to a standard Gaussian variable when normalized by $(\frac{1}{2}\log N)^{1/2}$. A recent conjecture of Fyodorov, Hiary and Keating makes a precise prediction for the large values of the logarithm of the characteristic polynomial.

**Conjecture 1.1** (Fyodorov-Hiary-Keating [35, 36]). *For $N \in \mathbb{N}$, let $U_N$ be a random matrix sampled uniformly from the group of $N \times N$ unitary matrices. Write $\mathrm{P}_N(z)$, $z \in \mathbb{C}$, for its characteristic polynomial. Then*

$$\max_{h\in[0,2\pi]} \log |\mathrm{P}_N(e^{\mathrm{i}h})| = \log N - \frac{3}{4}\log\log N + \mathcal{M}_N, \tag{1.1}$$

*where $(\mathcal{M}_N, N \in \mathbb{N})$ is a sequence of random variables that converges in distribution.*

The main result of this paper is a rigorous verification of the prediction for the leading order.

**Theorem 1.2.** *For $N \in \mathbb{N}$, let $\mathrm{U}_N$ be a random matrix sampled uniformly from the group of $N \times N$ unitary matrices. Write $\mathrm{P}_N(z)$, $z \in \mathbb{C}$, for its characteristic polynomial. Then*

$$\lim_{N\to\infty} \frac{\max_{h\in[0,2\pi]} \log |\mathrm{P}_N(e^{\mathrm{i}h})|}{\log N} = 1 \qquad \textit{in probability.} \tag{1.2}$$

It is known that the random field $\big((\log N)^{-1/2} \log |\mathrm{P}_N(e^{\mathrm{i}h})|, h \in [0, 2\pi]\big)$ converges in the sense of finite-dimensional distribution to a Gaussian field, with independent values at macroscopically separated evaluation points [43]. On mesoscopic scales, the covariance between two points $h_1$ and $h_2$ at distance $\Delta = |e^{\mathrm{i}h_1} - e^{\mathrm{i}h_2}|$ behaves like $\frac{\log(1/\Delta)}{\log N}$ when $\Delta$ is at least $1/N$, and approaches 1 for smaller distances [14]. This kind of decay of correlations is the defining characteristic of a log*-correlated random field*. The extrema of such fields have recently attracted much attention, cf. Section 1.1.

The almost perfect correlations below scale $1/N$ suggest that, to first approximation, one can think of the maximum over $[0, 2\pi]$ as a maximum over $N$ random variables with strong correlations on mesoscopic scales. Strikingly, the leading order prediction of Conjecture 1.1 is that the maximum is close to that of $N$ centered independent Gaussian random variables of variance $\frac{1}{2}\log N$, which would lie around $\log N - \frac{1}{4}\log\log N$. In other words, despite strong correlations between the values of $\log |\mathrm{P}_N(e^{\mathrm{i}h})|$ for different $h$, an analogy with independent Gaussian random variables correctly predicts the leading order of the maximum. The constant in front of the subleading correction $\log\log N$, however, differs. But, as we will explain below, it is exactly the constant expected for a log-correlated Gaussian field.

The conjecture was derived from precise computations of the moments of a suitable partition function and of the measure of high points, using statistical mechanics techniques developed for describing the extreme value statistics of disordered systems [34, 38, 39]. It is also supported by strong numerical evidence. A precise form for the distribution of the limiting fluctuations, which is consistent with those of log-correlated fields, is also predicted. We point out that $\log |\mathrm{P}_N(e^{\mathrm{i}h})|$ is believed to be a good model for the local behavior of the Riemann zeta function on the critical line. In particular, the authors conjecture a similar behavior for the extremes of the Riemann zeta function on an interval $[T, T + 2\pi]$ of the critical line with $N$ replaced by $\log T$, see [35, 36] for details and [4, 41] for rigorous proofs for a different random model of the zeta function.

The proof of Theorem 1.2 is outlined in Section 1.2 below. The key conceptual idea is the identification of an approximate branching random walk, or hierarchical field, in the Fourier decomposition of the characteristic polynomial. This is inspired by a branching structure in the Euler product of the Riemann Zeta function employed in [4]. In Section 1.2, it is explained how branching random walk heuristics provide an alternative justification of Conjecture 1.1. Furthermore, these heuristics can be made rigorous for the leading order, thanks to a robust approach introduced by Kistler in [45]. Technical difficulties remain to rigorously verify the finer predictions of the conjecture.

It is straightforward to adapt the approach to get information about the measure of *high points* for $\gamma \in (0, 1)$:

$$\mathcal{L}_N(\gamma) = \{h \in [0, 2\pi] : \log |\mathrm{P}_N(e^{\mathrm{i}h})| \geq \gamma \log N\} \ . \tag{1.3}$$

We show that with high probability the Lebesgue measure of $\mathcal{L}_N(\gamma)$ is close to $N^{-\gamma^2}$:

**Theorem 1.3.** *For $\gamma \in (0, 1)$,*

$$\lim_{N\to\infty} \frac{\log \mathrm{Leb}(\mathcal{L}_N(\gamma))}{\log N} = -\gamma^2, \textit{ in probability,} \tag{1.4}$$

*where* $\mathrm{Leb}(\cdot)$ *denotes the Lebesgue measure on* $[0, 2\pi]$.

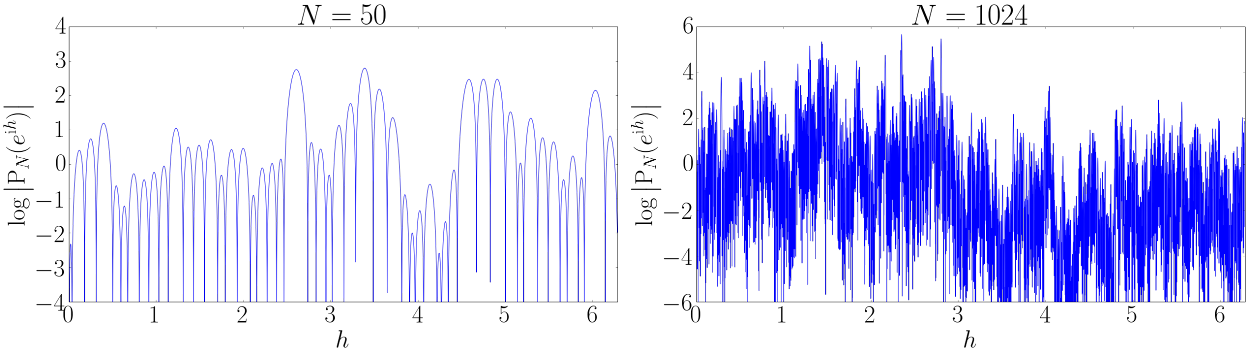


Figure 1: Realizations of $\log|\mathrm{P}_N(e^{\mathrm{i}h})|$, $0 \leq h < 2\pi$, for $N = 50$ and $N = 1024$. At microscopic scales, the field is smooth away from the eigenvalues, in contrast with the rugged landscape at mesoscopic and macroscopic scales.

This was conjectured by Fyodorov & Keating, see Section 2.4 in [36]. In fact, a more precise expression for the measure of high points was instrumental for their prediction of the subleading order in Conjecture 1.1, following the ideas of [39]. The theorem can be used to obtain the limit of the *free energy*

$$\frac{1}{\log N} \log\left(\frac{N}{2\pi}\int_0^{2\pi} |\mathrm{P}_N(h)|^\beta \mathrm{d}h\right) \tag{1.5}$$

of the random field $\log|\mathrm{P}_N(e^{\mathrm{i}h})|$. In particular, it is proposed in Section 2.2 of [36] that the free energy exhibits *freezing*, i.e. that above a critical temperature $\beta_c$, the free energy (1.5) divided by the inverse temperature $\beta$ becomes constant in the limit. The following, which is essentially an immediate consequence of Theorem 1.3, proves the conjecture.

**Corollary 1.4.** *For $\beta \geq 0$,*

$$\lim_{N\to\infty} \frac{1}{\log N} \log\left(\frac{N}{2\pi}\int_0^{2\pi} |\mathrm{P}_N(h)|^\beta \mathrm{d}h\right) = \begin{cases} 1 + \frac{\beta^2}{4} & \text{if } \beta < 2, \\ \beta & \text{if } \beta \geq 2, \end{cases} \quad \text{in probability}\ . \tag{1.6}$$

The work [36] contains other interesting conjectures on statistics of characteristic polynomials. One of them, a transition for the second moment of the partition function, was proved in [22].

*1.1 Relations to Previous Works.* This paper is part of the current research effort to develop a theory of extreme value statistics of log-correlated fields. There have been many rigorous works on the subject in recent years, and we give here a non-exhaustive list. In the physics literature, most predictions on the extreme value statistics of log-correlated fields can be found in [21]. In mathematics, the leading order of the two-dimensional Gaussian Free Field, was determined in [12]. In a series of impressive work, the form of the subleading correction as well as convergence of the fluctuations have been obtained [11, 16, 19, 32]. The approach (with the exception of [11]) follows closely the one used for branching random walks. This started with the seminal work of Bramson [15] for branching Brownian motion and was later extended to general branching random walks [2, 3, 7, 17, 18]. Log-correlated models are closely related to Gaussian Multiplicative chaos, see [48] for a review. In particular, convergence of the maximum of a related model of log-correlated Gaussian field was proved in [46]. We also refer to [52] for connections between the characteristic polynomial of unitary matrices and Gaussian Multiplicative chaos. From the perspective of spin glasses, Corollary 1.4 suggests that the model exihibits a *one-step replica symmetry breaking*. This was proved for Gaussian log-correlated fields in [5, 6, 13, 28]. A general theorem for the convergence of the maximum of log-correlated Gaussian fields was proved in [31]. A unifying point of view including non-Gaussian log-correlated fields and their hierarchical structure is developed in [45]. Important non-Gaussian examples include cover times of the two-dimensional random walk on the torus whose leading order was determined in [27] and subleading order in [9]. Also, the leading and subleading order of the Fyodorov-Hiary-Keating Conjecture are known for a random model of the Riemann zeta function other than CUE [4,41]. Finally, the analogue conjecture is expected to hold for other random matrix ensembles such as the *Gaussian Unitary Ensemble* [40].

*Notation.* Throughout this paper, we use the notation $\mathrm{O}(1)$ (resp. $\mathrm{o}(1)$) for a quantity uniformly bounded in $N$ (resp. going to 0 with $N$). The constants $c$ and $C$ denote universal constants varrying from line to line. The notation $a_N \lesssim b_N$ means that $a_N \leq C b_N$ for some $C$ independent of $N$.

*1.2 Outline of the Proof: Connection to Branching Random Walk.* Let $e^{\mathrm{i}\theta_1}, \ldots, e^{\mathrm{i}\theta_N}$ be the eigenvalues of $\mathrm{U}_N$. We are interested in

$$\log |\mathrm{P}_N(e^{\mathrm{i}h})| = \sum_{k=1}^{N} \log |1 - e^{\mathrm{i}(\theta_k - h)}|.$$

Recall that an integrable $2\pi$-periodic function has a Fourier series which converges pointwise wherever the function is differentiable (see e.g. [50, Theorem 2.1]). Since the $2\pi$-periodic integrable function $h \mapsto \mathrm{Re} \log(1 - e^{\mathrm{i}h})$ has Fourier series $-\sum_{j=1}^{\infty} \frac{\mathrm{Re}\, e^{-\mathrm{i}jh}}{j}$ we have

$$\log |\mathrm{P}_N(e^{-\mathrm{i}h})| = \sum_{k=1}^{N} \sum_{j=1}^{\infty} -\frac{\mathrm{Re}(e^{\mathrm{i}j(\theta_k - h)})}{j} = \sum_{j=1}^{\infty} -\frac{\mathrm{Re}(e^{-\mathrm{i}jh}\mathrm{Tr}\mathrm{U}_N^j)}{j}, \quad h \in \mathbb{R}, \tag{1.7}$$

where Tr stands for the trace and both right and left-hand sides are interpreted as $-\infty$ if $h$ equals an eigenangle. The starting point of the approach is to treat the above expansion as a *multiscale decomposition* for the process.

Though the traces of powers of $\mathrm{U}_N$ are not independent, it was shown in [29,30] that they are *uncorrelated*

$$\mathbb{E}\left(\mathrm{Tr}\mathrm{U}_N^j \ \overline{\mathrm{Tr}\mathrm{U}_N^k}\right) = \delta_{kj} \min(k, N), \tag{1.8}$$

where $\mathbb{E}$ is the expectation under the Haar measure $\mathbb{P}$ (by rotational invariance also $\mathbb{E}(\mathrm{Tr}\mathrm{U}_N^j \mathrm{Tr}\mathrm{U}_N^k) = 0$). At a heuristic level, the covariance structure of the traces explains the asymptotic Gaussianity of $\log |\mathrm{P}_N(e^{\mathrm{i}h})|$ as well as the correlation structure for different angles $h_1, h_2$, see (1.12) below. It is also the starting point of the connection to branching random walk.

Because of (1.8), one expects that the contribution to $\log |\mathrm{P}_N(e^{\mathrm{i}h})|$ of traces of powers $N$ or greater should be of order 1 since $\sum_{j \geq N} \frac{N}{j^2} = \mathrm{O}(1)$. Moreover, the variance of the powers less than $N$ becomes

$$\mathbb{E}\left(\left(\sum_{j=1}^{N-1} -\frac{\mathrm{Re}(e^{-\mathrm{i}jh}\mathrm{Tr}\mathrm{U}_N^j)}{j}\right)^2\right) = \frac{1}{2} \sum_{j<N} \frac{1}{j} = \frac{1}{2} \log N + \mathrm{O}(1).$$

The key idea is to divide the truncated sum into increments

$$W_\ell(h) = \sum_{e^{\ell-1} \leq j < e^\ell} -\frac{\mathrm{Re}(e^{-\mathrm{i}jh}\mathrm{Tr}\mathrm{U}_N^j)}{j} \quad \text{for } \ell = 1, \ldots, \log N, \tag{1.9}$$

which, thanks to (1.8), are uncorrelated and have variance $\frac{1}{2}\sum_{e^{\ell-1} \leq j < e^\ell} \frac{1}{j} \approx \frac{1}{2}$. Thus, at a heuristic level, one may think of the partial sums

$$X_\ell(h) = \sum_{\ell'=1}^{\ell} W_{\ell'}(h), \ \ \ell = 0, \ldots, \log N, \tag{1.10}$$

as a Gaussian random walk with increments of variance $\frac{1}{2}$, for any fixed $h$. Furthermore, for each $\ell$, the collection $(X_\ell(h), h \in [0, 2\pi])$, defines a random field on the unit circle which can be thought of as a sequence

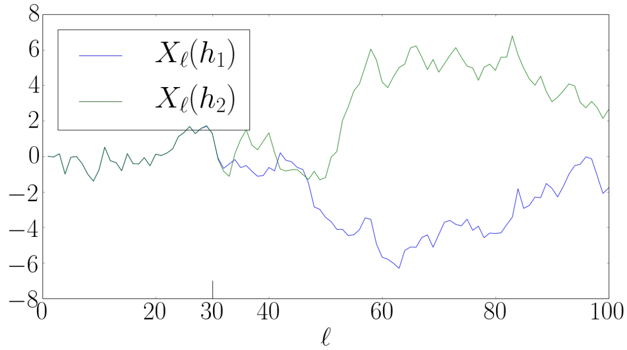


Figure 2: Illustration of the processes $X_\ell(h_1)$ and $X_\ell(h_2)$ for $h_1$ and $h_2$ at distance roughly $e^{-30}$, for a random matrix $\mathrm{U}_N$ with side-length $N = e^{100}$. The processes decorrelate at roughly scale $\ell = 30$.

of regularizations of $\log |\mathrm{P}_N(e^{ih})|$. It turns out that for $h$ and $h'$ in $[0, 2\pi]$ the corresponding partial sums $X_\ell(h)$ and $X_\ell(h')$ exhibit an approximate *branching structure*. To see this, define the *branching scale* of $h$ and $h'$ as

$$h \wedge h' = -\log \|h - h'\| \ , \tag{1.11}$$

where $\|h - h'\| = \min\{|h - h'|, (2\pi - |h - h'|)\}$ is the distance on the circle. Equation (1.8) implies

$$\mathbb{E}(W_\ell(h) W_\ell(h')) = \sum_{e^{\ell-1} \leq j < e^\ell} \frac{\cos(j\|h - h'\|)}{2j}. \tag{1.12}$$

For $j$ where $j\|h - h'\|$ is small the cosine is essentially 1, and for $j$ such that $j\|h - h'\|$ is large it oscillates, causing cancellation. In fact, by expanding the cosine in the first case and by using summation by parts in the other, it is not hard to see that

$$\mathbb{E}(W_\ell(h) W_\ell(h')) = \begin{cases} \frac{1}{2} + \mathrm{O}(e^{\ell - h\wedge h'}) & \text{if } \ell \leq h \wedge h', \\ \mathrm{O}(e^{-2(\ell - h \wedge h')}) & \text{if } \ell > h \wedge h'. \end{cases} \tag{1.13}$$

In other words, the increments $W_\ell(h)$ and $W_\ell(h')$ are almost perfectly correlated for $\ell$ before the branching scale and almost perfectly uncorrelated after the branching scales. We conclude from (1.13) that, if we restrict the field to the discrete set of $N$ points

$$\mathcal{H}_N = \left\{0, \frac{2\pi}{N}, 2\frac{2\pi}{N}, \dots, (N-1)\frac{2\pi}{N}\right\} \ , \tag{1.14}$$

the process $\big((X_\ell(h), l \leq \log N), h \in \mathcal{H}_N\big)$ defined by (1.10) is an *approximate branching random walk*. Namely, the random walks $X_\ell(h)$ and $X_\ell(h')$ are almost identical before the branching scale $\ell = h \wedge h'$, and continue almost independently after that scale, akin to a particle of a branching random walk that splits into two independent walks at time $h \wedge h'$, see Figure 2. Moreover, if $\|h - h'\| \leq e^{-\ell}$, the variance of the difference $X_\ell(h) - X_\ell(h')$ is of order one. Thus, the "variation" of $(X_\ell(h), h \in [0, 2\pi])$ is effectively captured by the values at $e^\ell$ equally spaced points for each $\ell$. This is reminiscent of a branching random walk where the mean number of offspring of a particle is $e$ and the average number of particles at time $\ell$ is $e^\ell$, see Figure 3.

Keeping the connection with branching random walk in mind, the proof of Theorem 1.2 is carried out in two steps. First, we obtain upper and lower bounds for truncated sums restricted to the discrete set $\mathcal{H}_N$. For this, we follow a multiscale refinement of the second method proposed by Kistler [45]. The second step is to derive from these upper and lower bounds for the entire sum (including large powers of the matrix) over the whole continuous interval $[0, 2\pi]$.

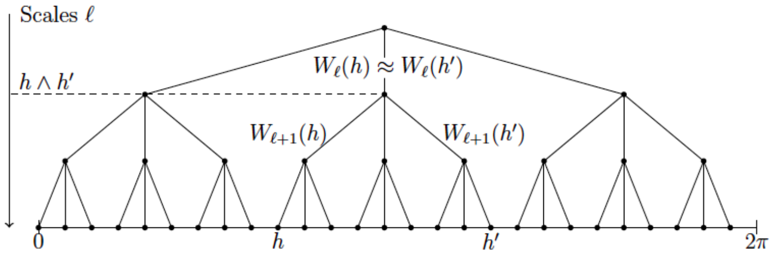


Figure 3: Illustration of the approximate branching random walk structure of the multiscale decomposition with $\approx e^\ell$ distinct values at each scale $\ell$.

*First step: the truncated sum on a discrete set.* The first result is an upper and lower bound for the maximum of the truncated sum for powers slightly smaller than $N$:

$$X_{(1-\delta)\log N}(h) = \sum_{1\leq \ell<(1-\delta)\log N} W_\ell(h) = \sum_{j<N^{1-\delta}} -j^{-1}\mathrm{Re}(e^{-\mathrm{i}jh}\mathrm{Tr}\mathrm{U}_\mathrm{N}^\mathrm{j}),\ h\in\mathcal{H}_N\ ,$$

where the parameter $\delta$ will be taken small enough in terms of $\varepsilon>0$. For a given $\varepsilon$, we prove

$$\lim_{N\to\infty}\mathbb{P}\left((1-\varepsilon)\log N\leq \max_{h\in\mathcal{H}_N}\sum_{1\leq\ell<(1-\delta)\log N} W_\ell(h)\leq \log N\right)=1\ . \tag{1.15}$$

The point of restricting to powers smaller than $N$ is that for such sums we can obtain sharp large deviation estimates.

The upper estimate (1.15) follows by a union bound

$$\mathbb{P}\left(\max_{h\in\mathcal{H}_N} X_{(1-\delta)\log N}(h)>\log N\right)\leq N\ \mathbb{P}\left(X_{(1-\delta)\log N}(h)>\log N\right)\leq \mathbb{E}(\mathcal{Z}(0))\ , \tag{1.16}$$

where $\mathcal{Z}(\varepsilon)$ is the number of exceedances

$$\mathcal{Z}(\varepsilon)=\#\{h\in\mathcal{H}_N : X_{(1-\delta)\log N}(h)>(1-\varepsilon)\log N\}\ . \tag{1.17}$$

The expectation $\mathbb{E}(\mathcal{Z}(0))$ goes to zero if $X_{(1-\delta)\log N}(h)$, whose variance is approximatively $(1-\delta)\log N$, admits Gaussian large deviations. This is proved by computing the exponential moments of $X_\ell(h)$ using a Riemann-Hilbert approach, see Proposition 5.11. The Riemann-Hilbert approach to compute the Fourier transform of linear statistics is an idea from [23].

The lower bound in (1.15) would follow from Chebyshev's inequality (or Paley-Zygmund inequality) if one could show $\mathbb{E}(\mathcal{Z}(\varepsilon)^2)=(1+\mathrm{o}(1))\mathbb{E}(\mathcal{Z}(\varepsilon))^2$. However, as for branching random walk, the second moment $\mathbb{E}(\mathcal{Z}(\varepsilon)^2)$ is in fact exponentially larger than $\mathbb{E}(\mathcal{Z}(\varepsilon))^2$ , due to rare events. A way around this is to modify the count by introducing a condition which takes into account the branching structure. At the level of the leading order, this can be achieved by a *K-level coarse graining* as explained in [45]. More precisely, for $K\in\mathbb{N}$ and $\delta=K^{-1}$, consider $K$ large increments of the "random walk" $X_\ell(h)$:

$$Y_m(h)=\sum_{\frac{m-1}{K}\log N<\ell\leq\frac{m}{K}\log N} W_\ell(h)=\sum_{N^{\frac{m-1}{K}}<j\leq N^{\frac{m}{K}}} -\frac{\mathrm{Re}(e^{-\mathrm{i}jh}\mathrm{Tr}(\mathrm{U}_N^j))}{j}\ , \tag{1.18}$$

for $m=1,\dots,K$, so that $X_{(1-\delta)\log N}(h)=Y_1(h)+\cdots+Y_{K-1}(h)$. It follows from (1.13) that each $Y_m(h)$ has variance roughly $\frac{1}{2K}\log N$. Moreover, $Y_m(h)$ and $Y_m(h')$ decorrelate for $\|h-h'\|\gg N^{-(m-1)/K}$. Because of the self-similar nature of branching random walk, the leading order should be linear in the scales $\ell$. In particular, for $h$ such that $X_{(1-\delta)\log N}(h)$ is at least $(1-\delta)\log N$, one expects each $Y_m(h)$ to contribute $\frac{1-\delta}{K}\log N$ to the maximum. Following [45], this leads us to consider the modified count

$$\tilde{\mathcal{Z}}(\varepsilon) = \#\{h \in \mathcal{H}_N : Y_m(h) \geq \frac{(1-\varepsilon)}{K}\log N, m = 2, \dots, K-1\} \ . \tag{1.19}$$

Note that the first increment $Y_1(h)$ is omitted. This is crucial since for

$$\mathbb{E}(\tilde{\mathcal{Z}}(\varepsilon)^2) = (1+\mathrm{o}(1))\mathbb{E}(\tilde{\mathcal{Z}}(\varepsilon))^2 \ , \tag{1.20}$$

to hold a sufficiently large proportion of all pairs $h_1$ and $h_2$ must have decorrelated increments, and the first increments $Y_1(h_1)$ and $Y_1(h_2)$ only decorrelate for macroscopically separated pairs, cf. (1.13). Dropping the first increments ensures enough independence between most pairs $h_1, h_2$ for the moments to match. Moreover, a separate first moment argument shows that this comes at negligible cost if $K$ is chosen large enough depending on $\varepsilon$. To prove the estimate (1.20), one also needs precise large deviation estimates for the vector $(Y_2(h), \dots, Y_{K-1}(h))$ for one $h$ and jointly for two different $h$'s. Again, such precise estimates on the exponential moments are obtained through the Riemann-Hilbert approach, see Proposition 5.11.

At this point, a few words on the prediction for the subleading correction are in order. At that level of precision, the expected number of points exceeding the predicted level $m_N = \log N - 3/4 \log\log N$ diverges in contrast to (1.16). However, thanks to the seminal work of Bramson on branching Brownian motion [15], it is known that the first moment is inflated by rare events: with large probability, the random walks $X_\ell(h)$ must remain below a linear barrier for the terminal value to be maximal. To capture this, one is led to consider exceedances with a barrier along the scales

$$\tilde{\mathcal{Z}}_B = \#\{h \in \mathcal{H}_N : X_{\log N}(h) \geq m_N, \ X_\ell(h) \leq \ell + B \text{ for } \ell = 1, \dots, \log N\}, \tag{1.21}$$

for some suitable $B$. If $(X_\ell(h), \ell \leq \log N)$ were an exact random walk, then by the *ballot theorem* (see e.g. [1]), the additional barrier requirement would balance out the divergence so that $\mathbb{E}(\tilde{\mathcal{Z}}_B) = \mathrm{O}(1)$. In fact, the restriction on the increments in (1.19) is a weak version of this barrier condition: if a random walk $X_\ell(h)$ goes above the linear barrier for small $\ell$, its "descendants" at subsequent times cannot take advantage of this by rising at a slower rate, since to be counted the increments $Y_2, \dots, Y_{K-1}$ must all make equally large jumps.

*Second step: extension to the full sum and the continuous interval.* Once the result (1.15) is established for the truncated sum $X_{(1-\delta)\log N}(h)$ on the discrete set $\mathcal{H}_N$, it remains to include traces of high powers to the sum and to consider the full interval $h \in [0, 2\pi]$.

For the lower bound, the extension to the full interval is direct. Controlling the contribution of high powers is harder: it requires large deviation bounds for $-\mathrm{Re}\sum_{j\geq N^{1-\delta}} \frac{\mathrm{Tr}(e^{-\mathrm{i}jh}U_N^j)}{j}$. These come from exponential moment estimates

$$\mathbb{E}\left(\exp\left(-\alpha \mathrm{Re} \sum_{j\geq N^{1-\delta}} \frac{\mathrm{Tr}(e^{-\mathrm{i}jh}U_N^j)}{j}\right)\right) = \mathbb{E}\left(\exp\left(\sum_{k=1}^N v(e^{\mathrm{i}\theta_k})\right)\right), \text{ where } v(e^{\mathrm{i}\theta}) = -\alpha \mathrm{Re} \sum_{j\geq N^{1-\delta}} \frac{e^{\mathrm{i}j(\theta-h)}}{j}.$$

Obtaining a precise approximation of the above expectation is equivalent to deriving asymptotics of Toeplitz determinants with a singularity of Fisher-Hartwig type. Indeed, by Heine's formula [42] we have

$$\mathbb{E}\left(\prod_{k=1}^N f(e^{\mathrm{i}\theta_k})\right) = \det((\hat{f}_{i-j})_{0\leq i,j\leq N}) \text{ with } f(z) = e^{V(z)}|z-1|^\alpha, \ V(z) = \alpha \mathrm{Re} \sum_{1\leq k<N^{1-\delta}} \frac{z^k}{k}. \tag{1.22}$$

In other words, we study the distribution of high powers of unitary matrices by superposition of a logarithmic singularity with smooth linear statistics. For a fixed analytic external potential $V$ the Riemann-Hilbert technique from [24, 25] is a robust method to estimate asymptotics of the above Toeplitz determinant. We follow the same approach with an additional technicality: in our regime of interest, $V$ depends on $N$

with an emerging singularity *of arbitrarily small mesoscopic scale* $N^{-1+\delta}$ (we note that the Riemann-Hilbert method was also used in [37] to evaluate determinants associated with merging singularities, in the context of the Gaussian Unitary Ensemble). The resulting Laplace transform of traces of high powers of unitary matrices, Proposition 3.2, may be of independent interest.

For the upper bound, the extension to the full sum and interval is obtained as follows. First, we obtain an upper bound on the truncated sum over all of $[0, 2\pi]$ using a dyadic chaining argument. This relies on a large deviation control on the difference of values of the truncated sum between two close points (at distance less than $N^{-1}$) obtained from exponential moments via Proposition 5.11. Second, we control the large values of the traces of high powers on a dense discrete set in the same way as for the lower bound (using the aforementioned Proposition 3.2). Combining these, we obtain an upper bound for the full sum on the dense discrete set. In the last step a bound on small gaps between eigenvalues [10] is used, which gives an appropriate control of the derivative of $\log |\mathrm{P}_N(e^{\mathrm{i}h})|$ on the circle, and allows us to conclude that the maximum over $[0, 2\pi]$ is close to the maximum over the dense discrete set.

*1.3 The imaginary part.* Theorem 1.2 has a natural analogue concerning the imaginary part of the logarithm of the characteristic polynomial. More precisely, if we choose the principal branch so that $-\pi/2 < \mathrm{Im} \log(1-e^{\mathrm{i}\theta}) < \pi/2$, then

$$\lim_{N\to\infty} \frac{\max_{h\in[0,2\pi]} \mathrm{Im} \sum_{k=1}^N \log(1-e^{\mathrm{i}(\theta_k-h)})}{\log N} = 1 \qquad \text{in probability.} \tag{1.23}$$

To illustrate the meaning of this result, let $\mathscr{N}_h(\mathrm{U}_N) = \sum_{k=1}^N \mathbb{1}_{(0,h)}(\theta_k)$ denote the number of eigenangles in the range $(0, h)$. Following [43] for any $-\pi < s < t < \pi$ we have the identity $\mathbb{1}_{(s,t)}(\theta) = \frac{t-s}{2\pi} + \frac{1}{\pi}\mathrm{Im} \log(1 - e^{\mathrm{i}(\theta-t)}) - \frac{1}{\pi}\mathrm{Im} \log(1 - e^{\mathrm{i}(\theta-s)})$, hence (1.23) yields

$$\lim_{N\to\infty} \frac{\max_{h\in[0,2\pi]} (\mathscr{N}_h(\mathrm{U}_N) - \frac{Nh}{2\pi})}{\log N} = \frac{1}{\pi} \qquad \text{in probability.}$$

This easily implies the following optimal rigidity bound for all eigenangles. Note that, in the context of Wigner matrices, the best known rigidity estimates (for all eigenvalues simultaneously) in the bulk of the spectrum are of type $(\log N)^C/N$ for some non optimal constant $C$ (see e.g. [33], [20]). The following gives both the optimal logarithmic exponent and constant for the CUE.

**Theorem 1.5.** *We label the eigenangles of a Haar-distributed* $\mathrm{U}_N$ *so that* $0 \leq \theta_1 \leq \cdots \leq \theta_N < 2\pi$. *Then, for any* $\varepsilon > 0$*, we have*

$$\mathbb{P}\left((2-\varepsilon)\frac{\log N}{N} < \sup_{1\leq k\leq N} \left|\theta_k - \frac{2\pi k}{N}\right| < (2+\varepsilon)\frac{\log N}{N}\right) \underset{N\to\infty}{\to} 1.$$

For the proof of (1.23), note that the above choice of principal branch for the logarithm coincides with the Fourier expansion: $\mathrm{Im} \log(1 - e^{\mathrm{i}\theta}) = -\mathrm{Im} \sum_{k\geq 1} \frac{e^{\mathrm{i}k\theta}}{k}$, so that our branching technique applies. All elements of the proof for the real part only require notational changes when considering the imaginary part, except Proposition 5.1 which bounds the tail of the Fourier series. We omit the straightforward proof adjustments, and we give in Section 5 the proof of the tail estimate, for the imaginary part of the series, i.e. the counterpart of Proposition 5.1 (see Proposition 5.12).

**Acknowledgements.** The authors thank Nicola Kistler for insightful discussions on the subject of log-correlated fields. L.-P. A. is partially supported by NSF grant DMS-1513441, PSC-CUNY Research Award 68784-00 46 and a Eugene M. Lang Junior Faculty Research Fellowship. P. B. is partially supported by NSF grants DMS-1208859 and DMS-1513587.

## 2 Maximum of the truncated sum on a discrete set

As mentioned in the introduction, we first study the maximum over the set $\mathcal{H}_N$ (see (1.14)) of cardinality $N$. Precise large deviation bounds are a crucial input for the method. We have good control on the large deviations of the (untruncated) sum

$$\sum_{j=1}^{\infty} -\frac{\mathrm{Re}(e^{-\mathrm{i}jh}\mathrm{Tr}\mathrm{U}_N^j)}{j}, \text{ for } h \in \mathbb{R}, \tag{2.1}$$

since its characteristic function can be explicitly computed using the Selberg integral (cf. Lemma 5.10). However, the proof of the lower bound, being a second moment calculation, requires precise joint large deviation bounds for sums at two different $h$'s. Computing exponential moments of two such sums, which include traces of powers close to or larger than $N$, is difficult due to the singularity of the logarithm. If the sum only contains traces with powers less than $N$, the Riemann-Hilbert techniques of Section 5 give the needed characteristic function (and exponential moment) bounds. The upper and lower bounds in this section are therefore given for the maximum of the truncated sum

$$\sum_{\ell=1}^{(1-\delta)\log N} W_\ell(h) = \sum_{1\leq j<N^{1-\delta}} -\frac{\mathrm{Re}(e^{-\mathrm{i}jh}\mathrm{Tr}\mathrm{U}_N^j)}{j}, \tag{2.2}$$

for small $\delta > 0$, where the "random walk" increments $W_\ell(h)$ are defined in (1.9). In Section 3 we extend these to bounds where $\mathcal{H}_N$ is replaced with the full interval $[0, 2\pi]$, and the sum (2.2) is replaced by the full sum.

*2.1 Upper bound.* An upper bound for the maximum of the truncated sum (2.2) on $\mathcal{H}_N$ is obtained by a straightforward union bound. We use the following "Gaussian" exponential moment bound for the sum (2.2), which is proved in Section 5 using Riemann-Hilbert techniques.

**Lemma 2.1.** *For any $\delta \in (0, 1)$, there is a constant $C$ such that for $h, \xi \in \mathbb{R}$ with $|\xi| \leq N^{\delta/10}$, then*

$$\mathbb{E}\left(\exp\left(\xi \sum_{\ell=1}^{(1-\delta)\log N} W_\ell(h)\right)\right) \leq C \exp\left(\frac{1}{2}\xi^2\sigma^2\right), \quad \textit{where } \sigma^2 = \frac{1}{2}\sum_{j=1}^{N^{1-\delta}} \frac{1}{j}. \tag{2.3}$$

This implies a "Gaussian" tail bound using the exponential Chebyshev inequality with $\xi = \pm x/\sigma^2$.

$$\mathbb{P}\left(\Big|\sum_{\ell=1}^{(1-\delta)\log N} W_\ell(h)\Big| \geq x\right) \leq C \exp\left(-\frac{x^2}{2\sigma^2}\right), \text{ for all } x \leq C\sigma^2. \tag{2.4}$$

We thus arrive at the following:

**Proposition 2.2** (Upper bound for truncated sum over discrete set)**.** *For any $\delta > 0$,*

$$\lim_{N\to\infty} \mathbb{P}\left(\max_{h\in\mathcal{H}_N} \sum_{\ell=1}^{(1-\delta)\log N} W_\ell(h) \leq \log N\right) = 1. \tag{2.5}$$

*Proof.* Since $\sigma^2 = (1-\delta)\frac{1}{2}\log N + \mathrm{O}(1)$, a union bound using (2.4) gives

$$\begin{aligned}\mathbb{P}\left(\max_{h\in\mathcal{H}_N} \sum_{\ell=1}^{(1-\delta)\log N} W_\ell(h) \geq \log N\right) &\leq N\,\mathbb{P}\left(\sum_{\ell=1}^{(1-\delta)\log N} W_\ell(h) \geq \log N\right) \\ &\leq N \exp\left(-(1+\mathrm{o}(1))\frac{(\log N)^2}{(1-\delta)\log N}\right) \leq N^{-\delta+\mathrm{o}(1)} = \mathrm{o}(1)\ .\end{aligned} \tag{2.6}$$

□

*2.2 Lower bound.* As described in the introduction, we use a truncated second moment argument to prove th following result.

**Proposition 2.3** (Lower bound for truncated sum over discrete set)**.** *For any $\varepsilon > 0$, there exists $\delta = \delta(\varepsilon) > 0$ such that*

$$\lim_{N\to\infty} \mathbb{P}\left( \max_{h\in\mathcal{H}_N} \sum_{\ell=1}^{(1-\delta)\log N} W_\ell(h) \geq (1-\varepsilon)\log N \right) = 1 \ . \tag{2.7}$$

To formulate the truncation, recall the coarse increments $Y_1(h), \ldots, Y_K(h)$, $K \in \mathbb{N}$, defined in (1.18). Note that by (1.8) these increments have variance

$$\sigma_m^2 = \mathbb{E}(Y_m(h)^2) = \frac{1}{2} \sum_{N^{(m-1)/K} < j \leq N^{m/K}} \frac{1}{j} = \frac{1}{2}\frac{\log N}{K} + \mathrm{O}(N^{-(m-1)/K}), \qquad \forall h, \tag{2.8}$$

and, more generally, covariance

$$\rho_m(h_1,h_2) = \mathbb{E}(Y_m(h_1)Y_m(h_2)) = \frac{1}{2} \sum_{N^{(m-1)/K} < j \leq N^{m/K}} \frac{\cos(j\|h_1-h_2\|)}{j}, \qquad \forall h_1, h_2 \ . \tag{2.9}$$

Expanding $e^{-\mathrm{i}jh}$ for large $h_1 \wedge h_2$ and summing by parts for small $h_1 \wedge h_2$, one arrives at the estimate

$$\rho_m(h_1,h_2) = \begin{cases} \mathrm{O}(N^{-(m-1)/K} e^{h_1\wedge h_2}) & \text{if } h_1\wedge h_2 \leq (m-1)\frac{\log N}{K}, \\ \sigma_m^2 + \mathrm{O}\left( \left( N^{m/K} e^{-h_1\wedge h_2}\right)^2 \right) & \text{if } h_1 \wedge h_2 \geq m\frac{\log N}{K} \ . \end{cases} \tag{2.10}$$

Therefore, unless $(m-1)\frac{\log N}{K} \leq h_1 \wedge h_2 \leq m\frac{\log N}{K}$, the increments $Y(h_1), Y(h_2)$ are almost completely correlated or completely decorrelated.

The second moment method is applied to the counting random variable

$$Z = \sum_{h\in\mathcal{H}_N} \mathbb{1}_{J_x(h)}, \qquad J_x(h) = \{Y_m(h) \geq x, m = 2,\ldots,K-1\} \ . \tag{2.11}$$

The level $x$ needs to be picked appropriately. For a given $\varepsilon > 0$, take

$$x = (1-\varepsilon/2)\frac{1}{K}\log N \ . \tag{2.12}$$

Proposition 2.3 is a simple consequence of the following, which will be proved in the remainder of the section.

**Proposition 2.4.** *For any $K \geq 3$ and $\varepsilon > 0$,*

$$\mathbb{E}\left(Z^2\right) = (1+\mathrm{o}_K(1))\mathbb{E}(Z)^2, \ \textit{as } N \to \infty. \tag{2.13}$$

Here and throughout this section a subscript $K$ on a $\mathrm{o}_K(\cdot)$ or $\mathrm{O}_K(\cdot)$ term denotes that all constants inside those terms may depend on $K$.

*Proof of Proposition 2.3.* Take $\delta = K^{-1}$. On the event $\{Z \geq 1\}$, there is an $h \in \mathcal{H}_N$ such that

$$\sum_{\log N/K < \ell \leq (1-1/K)\log N} W_\ell(h) \geq (1-\varepsilon/2)(1-2/K)\log N. \tag{2.14}$$

Together with the Paley-Zygmund inequality

$$\mathbb{P}(Z \geq 1) \geq \frac{\mathbb{E}(Z)^2}{\mathbb{E}(Z^2)} \ , \tag{2.15}$$

and Proposition 2.4, this implies that for any $\varepsilon > 0$ and $K \geq 3$,

$$\lim_{N\to\infty} \mathbb{P}\left(\max_{h\in\mathcal{H}_N} \sum_{\log N/K<\ell\leq(1-1/K)\log N} W_\ell(h) \geq (1-\varepsilon/2)(1-2/K)\log N\right) = 1.$$

A union bound as in (2.6) for $\ell$ up to $K^{-1}\log N$ implies that for $K$ large enough

$$\lim_{N\to\infty} \mathbb{P}\left(\sum_{\ell=1}^{\log N/K} W_\ell(h) > -\frac{\varepsilon}{3}\log N \text{ for all } h\in\mathcal{H}_N\right) = 1\ .$$

The result follows by taking $K$ large enough in terms of $\varepsilon$. □

The rest of the section is devoted to proving (2.13). The first and second moments can be written as

$$\mathbb{E}(Z) = N\mathbb{P}(J_x(0)) \text{ and } \mathbb{E}(Z^2) = \sum_{h_1,h_2\in\mathcal{H}_N} \mathbb{P}(J_x(h_1)\cap J_x(h_2)). \tag{2.16}$$

It suffices to find a lower bound on $\mathbb{P}(J_x(0))$ and an upper bound on $\mathbb{P}(J_x(h_1)\cap J_x(h_2))$. Exponential moments of the increments $Y_m(h)$ and the exponential Chebyshev inequality do not yield bounds precise enough to match $\mathbb{E}(Z^2)$ to $\mathbb{E}(Z)^2$ up to a multiplicative constant tending to one. In fact, it is necessary to go beyond the level of precision of large deviations at least for the pairs $h_1$, $h_2$ that contribute the most to $E(Z^2)$, namely points that are close to being macroscopically separated. This is done using characteristic function bounds together with Fourier inversion. The Riemann-Hilbert techniques of Section 5 can be used to obtain the following bounds on the characteristic function.

**Lemma 2.5.** *Let $K \geq 1$, $h_1, h_2 \in \mathbb{R}$ and write $\boldsymbol{Y}_m = (Y_m(h_1), Y_m(h_2))$ for $m = 2,\dots,K-1$. For all $\boldsymbol{\xi}_m \in \mathbb{C}^2, m = 2,\dots,K-1$, with $\|\boldsymbol{\xi}\| \leq N^{1/(10K)}$ we have*

$$\mathbb{E}\left(\exp\left(\sum_{m=2}^{K-1}\boldsymbol{\xi}_m\cdot\boldsymbol{Y}_m\right)\right) = \left(1+\mathrm{O}(e^{-N^{1/(10K)}})\right)\exp\left(\frac{1}{2}\sum_{m=2}^{K}\boldsymbol{\xi}_m\cdot\boldsymbol{\Sigma}_m\boldsymbol{\xi}_m\right)\ ,$$

*where*

$$\boldsymbol{\Sigma}_m = \begin{pmatrix} \sigma_m^2 & \rho_m \\ \rho_m & \sigma_m^2 \end{pmatrix}\ , m = 2,\dots,K-1, \tag{2.17}$$

*for $\sigma_m^2$ and $\rho_m$ defined in* (2.8)-(2.9).

To quantitatively invert the Fourier transform, we use the following crude bound.

**Lemma 2.6.** *Let $d \geq 1$. There are constants $c = c(d)$ such that if $\mu$ and $\nu$ are probability measures on $\mathbb{R}^d$ with Fourier transforms $\hat{\mu}(\boldsymbol{t}) = \int e^{\mathrm{i}\boldsymbol{t}\cdot\boldsymbol{x}}\mu(\mathrm{d}x)$ and $\hat{\nu}(\boldsymbol{t}) = \int e^{\mathrm{i}\boldsymbol{t}\cdot\boldsymbol{x}}\nu(\mathrm{d}x)$, then for any $R, T > 0$ and any function $f : \mathbb{R}^d \to \mathbb{R}$ with Lipschitz constant $C$*

$$|\mu(f)-\nu(f)| \leq c\frac{C}{T} + \|f\|_\infty\left\{c(RT)^d\,\|\mathbb{1}_{(-T,T)^d}(\hat{\mu}-\hat{\nu})\|_\infty + \mu\left(([-R,R]^d)^c\right) + \nu\left(([-R,R]^d)^c\right)\right\}. \tag{2.18}$$

*Proof.* This follows from the quantitative Fourier inversion estimate (11.26) of Corollary 11.5 [8]. One uses a smoothing kernel $K_\varepsilon$ whose Fourier transform is supported on $[-c\varepsilon^{-1}, c\varepsilon^{-1}]^d$ (its existence is guaranteed by Theorem 10.1 [8]), with $\varepsilon = T^{-1}$. The quantity in the curly braces in (2.18) is the crude upper bound for the integral $\int_{\mathbb{R}^d}|(\mu-\nu)*K_\varepsilon|(\mathrm{d}x)$ one obtains from using the point-wise bound $|g(x)| \leq cT^d\|\mathbb{1}_{(-T,T)^d}(\hat{\mu}-\hat{\nu})\|_\infty$ on the density $g$ of $(\mu-\nu)*K_\varepsilon$, when $x \in [-R,R]^d$, and a trivial bound for the integral over the complement $([-R,R]^d)^c$. □

Pairs of points $h_1$ and $h_2$ that are macroscopically (or almost macroscopically) separated are the main contribution to $\mathbb{E}(Z^2)$. For such $h_1$ and $h_2$ we expect the events $J_x(h_1)$ and $J_x(h_2)$ to be essentially independent, and the bounds (2.19)-(2.20) that now follow make this quantitative.

**Proposition 2.7** (Two-point bound; Decoupling)**.** *Let $h_1, h_2 \in \mathbb{R}$ be such that $h_1 \wedge h_2 \le \frac{\log N}{K}$. Then for $0 < x \le \log N$, we have that*

$$\mathbb{P}\left(J_x(h_1) \cap J_x(h_2)\right) \le C \exp\left(-\sum_{m=2}^{K-1} \frac{x^2}{\sigma_m^2}\right) \ . \tag{2.19}$$

*Furthermore, if $h_1 \wedge h_2 \le \frac{\log N}{2K}$, then we have the more precise bound*

$$\mathbb{P}\left(J_x(h_1) \cap J_x(h_2)\right) \le (1 + \mathrm{O}_K\left(N^{-c}\right)) e^{-\sum_{m=2}^{K-1} \frac{x^2}{\sigma_m^2}} \left(\prod_{m=2}^{K-1} \eta_{0,\sigma_m^2}\left(e^{-\frac{xy}{\sigma_m^2}} \mathbb{1}_{[0,\infty)}(y)\right)\right)^2 \ , \tag{2.20}$$

*where $\eta_{0,\sigma^2}$ denotes the centered Gaussian measure on $\mathbb{R}$ with variance $\sigma^2$.*

Before starting the proof, we note that the Gaussian expectation in parentheses satisfies

$$C(\log N)^{-1/2} \le \eta_{0,\sigma_m^2}\left(e^{-\frac{xy}{\sigma_m^2}} \mathbb{1}_{[0,\infty)}(y)\right) \le 1 \ , \tag{2.21}$$

because of the bound

$$\eta_{0,\sigma_m^2}\left(e^{-\frac{xy}{\sigma_m^2}} \mathbb{1}_{[0,\infty)}(y)\right) = e^{\frac{x^2}{2\sigma_m^2}} \int_{x/\sigma_m}^{\infty} \frac{e^{-z^2/2}}{\sqrt{2\pi}} \mathrm{d}z \ge c\frac{\sigma_m}{x} \ ,$$

the estimate (2.8) on $\sigma_m$ and the assumption $0 \le x \le \log N$.

*Proof.* Consider the probability measure $\mathbb{Q}$ constructed from $\mathbb{P}$ through the density

$$\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} = \frac{e^{\sum_{m=2}^{K-1} \boldsymbol{\xi}_m \cdot \boldsymbol{Y}_m}}{\mathbb{E}\left(e^{\sum_{m=2}^{K-1} \boldsymbol{\xi}_m \cdot \boldsymbol{Y}_m}\right)}. \tag{2.22}$$

for $\boldsymbol{\xi}_m = \xi_m(1,1)$ to be picked later. We write $\mathbb{E}_{\mathbb{Q}}$ for the expectation under $\mathbb{Q}$ and $\mathbb{E}$ for the expectation under $\mathbb{P}$. For $\boldsymbol{x} = (x, x)$, the probability can be written as

$$\mathbb{P}\left(J_x(h_1) \cap J_x(h_2)\right) = \mathbb{E}\left(e^{\sum_{m=2}^{K-1} \boldsymbol{\xi}_m \cdot \boldsymbol{Y}_m}\right) e^{-\sum_{m=2}^{K-1} \boldsymbol{\xi}_m \cdot \boldsymbol{x}} \mathbb{E}_{\mathbb{Q}}\left(e^{-\sum_{m=2}^{K-1} \boldsymbol{\xi}_m \cdot (\boldsymbol{Y}_m - \boldsymbol{x})}; J_x(h_1) \cap J_x(h_2)\right) \ . \tag{2.23}$$

The first factor is evaluated using Lemma 2.5 on exponential moments. For the choice,

$$\xi_m = \frac{x}{\sigma_m^2}, m = 2, \ldots, K-1, \tag{2.24}$$

we have $0 < \xi_m < 2$ by the assumption on $x$, so the Lemma can be applied. We get

$$\mathbb{E}\left(e^{\sum_{m=2}^{K-1} \boldsymbol{\xi}_m \cdot \boldsymbol{Y}_m}\right) = \left(1 + \mathrm{O}\left(e^{-N^{1/(10K)}}\right)\right) \exp\left(\frac{1}{2} \sum_{m=2}^{K-1} \boldsymbol{\xi}_m \cdot \boldsymbol{\Sigma}_m \boldsymbol{\xi}_m\right) \ . \tag{2.25}$$

The estimate (2.10) on the covariance gives that $\sum_{m \ge 2} \rho_m(h_1, h_2) = O(1)$ for $h_1 \wedge h_2 \le \log N/K$. Therefore, the quadratic form reduces to

$$\frac{1}{2} \sum_{m=2}^{K-1} \boldsymbol{\xi}_m \cdot \boldsymbol{\Sigma}_m \boldsymbol{\xi}_m = \sum_{m=2}^{K-1} \xi_m^2 (\sigma_m^2 + \rho_m) = \sum_{m=2}^{K-1} \xi_m^2 \sigma_m^2 + \mathrm{O}(1) \ . \tag{2.26}$$

Putting this in (2.25), we get

$$\mathbb{E}\left(e^{\sum_{m=2}^{K-1} \boldsymbol{\xi}_m \cdot \boldsymbol{Y}_m}\right) e^{-\sum_{m=2}^{K-1} \boldsymbol{\xi}_m \cdot \boldsymbol{x}} = e^{-\sum_{m=2}^{K-1} \xi_m x + \mathrm{O}(1)} \ .$$

Equation (2.19) follows from this since the third factor of (2.23) is smaller than 1 by the definition of the event.

A more careful analysis of (2.23) is needed to prove (2.20). First, note that if $h_1 \wedge h_2 < \log N/(2K)$, then $\sum_{m\geq 2}\rho_m = O\left(N^{-1/(2K)}\right)$ by (2.10). Therefore for the same choice of $\xi_m$, we have

$$\mathbb{E}\left(e^{\sum_{m=2}^{K-1}\boldsymbol{\xi}_m\cdot\boldsymbol{Y}_m}\right)e^{-\sum_{m=2}^{K-1}\boldsymbol{\xi}_m\cdot\boldsymbol{x}} = (1+\mathrm{O}(N^{-1/(2K)}))e^{-\sum_{m=2}^{K-1}\xi_m x} \ .$$

We will thus be done once we show

$$\mathbb{E}_{\mathbb{Q}}\left(e^{-\sum_{m=2}^{K-1}\boldsymbol{\xi}_m\cdot(\boldsymbol{Y}_m-\boldsymbol{x})}; J_x(h_1)\cap J_x(h_2)\right) = \left(\prod_{m=2}^{K-1}\eta_{0,\sigma_m^2}\left(e^{-\frac{xy}{\sigma_m^2}}\mathbb{1}_{[0,\infty)}(y)\right)+\mathrm{O}_K\left(N^{-c}\right)\right)^2 . \tag{2.27}$$

Note that the product in (2.27) is the dominant term since it is at least $c\log N^{-(K-1)}$ by (2.21). We prove (2.27) using Fourier inversion.

Let $\boldsymbol{t}_m = (t_{1,m}, t_{2,m})$ and $t_{j,m}\in\mathbb{R}$ for $m = 2,\ldots,K-1$ and consider $\boldsymbol{\xi}_m+\mathrm{i}\boldsymbol{t}_m$. Suppose $|t_{j,m}| < N^{1/(32K)}$ so that $|\xi_m+\mathrm{i}t_{j,m}| < N^{1/(16K)}$. Let $\mu$ be the law of $(\boldsymbol{Y}_m-\boldsymbol{x}; m = 2,\ldots,K-1)$ under $\mathbb{Q}$. Its Fourier transform $\hat{\mu}$ becomes:

$$\mathbb{E}_{\mathbb{Q}}\left(e^{\mathrm{i}\sum_{m=2}^{K-1}\boldsymbol{t}_m\cdot(\boldsymbol{Y}_m-\boldsymbol{x})}\right) = \frac{\mathbb{E}\left(e^{\sum_{m=2}^{K-1}(\boldsymbol{\xi}_m+i\boldsymbol{t}_m)\cdot\boldsymbol{Y}_m}\right)s}{\mathbb{E}\left(e^{\sum_{m=2}^{K-1}\boldsymbol{\xi}_m\cdot\boldsymbol{Y}_m}\right)}e^{-\mathrm{i}\sum_{m=2}^{K-1}\boldsymbol{t}_m\cdot\boldsymbol{x}} \ . \tag{2.28}$$

We apply Lemma 2.5 with $\boldsymbol{\xi}_m+\mathrm{i}\boldsymbol{t}_m$ in place of $\boldsymbol{\xi}_m$ to the numerator, and use (2.25) to bound the denominator. After cancellation, we obtain that (2.28) equals

$$\left(1+\mathrm{O}\left(e^{-N^{1/(10K)}}\right)\right)\exp\left(-\sum_{m=2}^{K-1}\frac{1}{2}\boldsymbol{t}_m\cdot\boldsymbol{\Sigma}_m\boldsymbol{t}_m + i\sum_{m=2}^{K-1}\boldsymbol{t}_m\cdot\boldsymbol{\Sigma}_m\boldsymbol{\xi}_m - \mathrm{i}\sum_{m=2}^{K-1}\boldsymbol{t}_m\cdot\boldsymbol{x}\right) . \tag{2.29}$$

As in (2.26), but using here $\sum_{m\geq 2}\rho_m = \mathrm{O}(N^{-1/(2K)})$, we have that

$$\boldsymbol{t}_m\cdot\boldsymbol{\Sigma}_m\boldsymbol{\xi}_m = \boldsymbol{t}_m\cdot\boldsymbol{x} + \mathrm{O}\left(\frac{\rho_m x}{\sigma_m^2}\|\boldsymbol{t}_m\|\right) = \boldsymbol{t}_m\cdot\boldsymbol{x}+\mathrm{O}(N^{-1/(4K)}) \ .$$

Thus (2.29) in fact equals

$$\left(1+\mathrm{O}\left(N^{-1/(4K)}\right)\right)\exp\left(-\frac{1}{2}\sum_{m=2}^{K-1}(t_{1,m}^2+t_{2,m}^2)\sigma_m^2\right). \tag{2.30}$$

The exponential above is precisely the Fourier transform $\hat{\nu}$ of $\nu = \otimes_{m=2}^{k-1}\eta_{0,\sigma_m^2}$. Thus we have shown that

$$\hat{\mu}(\boldsymbol{t}_2,\ldots,\boldsymbol{t}_{K-1}) = \left(1+\mathrm{O}(N^{-1/(4K)})\right)\hat{\nu}(\boldsymbol{t}_2,\ldots,\boldsymbol{t}_{K-1}), \quad \text{when } |t_{i,m}|\leq N^{1/(32K)}. \tag{2.31}$$

This suggests the decoupling in (2.27). To complete the argument, consider the function $g_\xi:\mathbb{R}\to\mathbb{R}$ where

$$g_\xi(y) = \begin{cases} 0 & \text{if } y\leq -N^{-1/(64K^2)}, \\ e^{-\xi y} & \text{if } y>0, \end{cases}$$

and $g_\xi$ is linearly interpolated on $[-N^{-1/(64K^2)}, 0]$. Note that $g_\xi$ is bounded by 1 and has Lipschitz constant $N^{1/(64K^2)}$. By definition,

$$\mathbb{E}_{\mathbb{Q}}\left(e^{-\sum_{m=2}^{K-1}\boldsymbol{\xi}_m\cdot(\boldsymbol{Y}_m-\boldsymbol{x})}; J_x(h_1)\cap J_x(h_2)\right)\leq \mathbb{E}_{\mathbb{Q}}\left(\prod_{m=2}^{K-1}\prod_{i=1,2}g_{\xi_m}\left(Y_m(h_i)-x\right)\right) . \tag{2.32}$$

Lemma 2.6 can be applied with $d = 2(K-2)$, $T = N^{1/(32K^2)}$, $R = N^{1/(32K^2)}$. The right-hand side of (2.32) becomes

$$\begin{aligned}&\left(\prod_{m=2}^{K-1} \eta_{0,\sigma_m^2}\left(g_{\xi_m}(y)\right)\right)^2 + \mathrm{O}\left(N^{1/\left(64K^2\right)-1/\left(32K^2\right)}\right)\\&+\mathrm{O}\left(N^{2(K-2)/\left(32K^2\right)}N^{2(K-2)/(32K^2)}N^{-1/(4K)}\right)\\&+\left(\otimes_{m=2}^{K-1}\eta_{0,\sigma_m^2}\left(\left(-N^{1/\left(32K^2\right)},N^{1/\left(32K^2\right)}\right)^c\right)\right)^2\\&+\mathbb{Q}\left(\exists m:|Y_m(h_1)-x|>N^{1/\left(32K^2\right)} \text{ or } |Y_m(h_2)-x|>N^{1/\left(32K^2\right)}\right)\ .\end{aligned} \tag{2.33}$$

A standard Gaussian estimate and (2.8) show that

$$\otimes_{m=2}^{K-1}\eta_{0,\sigma_m^2}\left(\left(-N^{1/(32K^2)},N^{1/(32K^2)}\right)^c\right) = \mathrm{O}\left(Ke^{-cN^{1/(16K^2)}/\log N}\right). \tag{2.34}$$

Lemma 2.5 and the definition of $\mathbb{Q}$ imply the exponential moment $\mathbb{Q}(\exp(\lambda(Y_m(h)-x)))s \le c\exp(\lambda^2\sigma_m^2)$, valid for all $m$, $h$ and $1 \le |\lambda| \le N^{1/(10K)}$, where we have used that $\rho_m \le \sigma_m^2$. The choice $\lambda = N^{1/(32K^2)}/\sigma_m^2$ and the exponential Markov's inequality shows that for all $m$ and $h$,

$$\mathbb{Q}\left(|Y_m(h)-x| > N^{1/(32K^2)}\right) \le c\exp(-cN^{1/(16K^2)}/\log N)\ . \tag{2.35}$$

This means that last term of (2.33) is also bounded by the right-hand side of (2.34). We conclude that

$$\mathbb{E}_{\mathbb{Q}}\left(e^{-\sum_{m=2}^{K-1}\boldsymbol{\xi}_m\cdot(\boldsymbol{Y}_m-\boldsymbol{x})}; J_x(h_1)\cap J_x(h_2)\right) \le \left(\prod_{m=2}^{K-1}\eta_{0,\sigma_m^2}\left(g_{\xi_m}(y)\right)\right)^2 + \mathrm{O}_K(N^{-c})\ . \tag{2.36}$$

Note that $\left|\eta_{0,\sigma_m^2}(g_{\xi_m}) - \eta_{0,\sigma_m^2}\left(e^{-\xi y}\mathbb{1}_{[0,\infty)}(y)\right)\right| \le N^{-1/\left(64K^2\right)}$ and recall (2.21). This together with (2.36) shows (2.27), and completes the proof of (2.20). □

We now turn to bounding $P\left(J_x(h_1)\cap J_x(h_2)\right)$ when $h_1$ and $h_2$ are "close". In this regime we do not need such a precise bound, so Fourier inversion is not needed. The bound (2.37) reflects that if $h_1 \wedge h_2 \in [(j-1)\log N/K, j\log N/K]$ the increments $Y_m(h_1), Y_m(h_2), m = 2,\dots,j-1$, are essentially perfectly correlated, while the increments $Y_m(h_1), Y_m(h_2), m = j+1,\dots,K-1$ are essentially independent. The increments $Y_j(h_1), Y_j(h_2)$ are partially correlated, but we ignore this and dominate by the scenario where the correlation is perfect. This leads to a loss in the bound, which turns out to be irrelevant in the second moment computation.

**Proposition 2.8** (Two-point bound; Coupling)**.** *Let $h_1, h_2 \in \mathbb{R}$ such that $\frac{j-1}{K}\log N < h_1 \wedge h_2 \le \frac{j}{K}\log N$ for some $j = 2,\dots,K-1$. Then for $0 < x \le \log N$, we have that*

$$\mathbb{P}\left(J_x(h_1)\cap J_x(h_2)\right) \le C\exp\left(-\sum_{m=2}^{j}\frac{x^2}{2\sigma_m^2} - \sum_{m=j+1}^{K-1}\frac{x^2}{\sigma_m^2}\right)\ . \tag{2.37}$$

*Proof.* The proof is very similar to the proof of (2.19). As in (2.23), we use a change of measure $\mathbb{Q}$

$$\mathbb{P}\left(J_x(h_1)\cap J_x(h_2)\right) = e^{-\sum_{m=2}^{K-1}\boldsymbol{\xi}_m\cdot\boldsymbol{x}}\mathbb{E}\left(e^{\sum_{m=2}^{K-1}\boldsymbol{\xi}_m\cdot\boldsymbol{Y}_m}\right)\mathbb{E}_{\mathbb{Q}}\left(e^{-\sum_{m=2}^{K-1}\boldsymbol{\xi}_m\cdot(\boldsymbol{Y}_m-\boldsymbol{x})}; J_x(h_1)\cap J_x(h_2)\right) \tag{2.38}$$

where $\boldsymbol{\xi}_m = (\xi_m,\xi_m)$, $\xi_m \ge 0$ and $\boldsymbol{x} = (x,x)$, Note that the last factor is again smaller than 1 by the definition of $J_x(h)$. As in (2.25), we have using Lemma 2.5 that

$$\mathbb{E}\left(e^{\sum_{m=2}^{K-1}\boldsymbol{\xi}_m\cdot\boldsymbol{Y}_m}\right) \le C\exp\left(\frac{1}{2}\sum_{m=2}^{K-1}\boldsymbol{\xi}_m\cdot\boldsymbol{\Sigma}_m\boldsymbol{\xi}_m\right)\ .$$

By (2.10) and the assumption $(j-1)\log N/K < h_1 \wedge h_2 \le j \log N/K$, we have for $m \neq j$

$$\frac{1}{2}\boldsymbol{\xi}_m \cdot \boldsymbol{\Sigma}_m \boldsymbol{\xi}_m = \xi_m^2\big(\sigma_m^2 + \rho_m(h_1,h_2)\big) = \begin{cases} \xi_m^2\big(2\sigma_m^2 + \mathrm{O}(1)\big) & \text{if } m \le j-1 \\ \xi_m^2\big(\sigma_m^2 + \mathrm{O}(1)\big) & \text{if } m \ge j+1. \end{cases} \tag{2.39}$$

For $m = j$, since $\rho_m \le \sigma_m^2$, we have that

$$\frac{1}{2}\boldsymbol{\xi}_j \cdot \boldsymbol{\Sigma}_j \boldsymbol{\xi}_j \le 2\xi_j^2 \sigma_j \ . \tag{2.40}$$

To optimize the bound we pick

$$\xi_m = \begin{cases} \frac{x}{2\sigma_m^2} & \text{if } m \le j, \\ \frac{x}{\sigma_m^2} & \text{if } m \ge j+1. \end{cases} \tag{2.41}$$

Using (2.39)-(2.41) in (2.38) we obtain (2.37). □

Finally, we bound the one point probability $\mathbb{P}(J_x(h))$ from below. Here we again need a precise bound which uses Fourier inversion.

**Proposition 2.9** (One-point bound)**.** *For every $h \in \mathbb{R}$ and $0 < x \le \log N$, we have that*

$$\mathbb{P}(J_x(h)) \ge \big(1 + \mathrm{O}_K\left(N^{-c}\right)\big)\, e^{-\sum_{m=2}^{K-1} \frac{x^2}{2\sigma_m^2}} \prod_{m=2}^{K-1} \eta_{0,\sigma_m^2}\left(e^{-\frac{xy}{\sigma_m^2}} \mathbb{1}_{[0,\infty)}(y)\right) \ , \tag{2.42}$$

*where $\sigma_m^2$ is defined in* (2.8)*.*

*Proof.* By rotational invariance, it suffices to consider the case $h = 0$. We write $Y_m = Y_m(0)$, $m = 2, \dots, K-1$ for simplicity. The proof relies on a change of measure followed by a Fourier inversion as in the proof of Proposition 2.7. Let $\xi_m = x/\sigma_m^2$ as in (2.24), so that $0 < \xi_m < 1$ by the assumption on $x$. Consider the probability measure $\mathbb{Q}$ constructed from $\mathbb{P}$ via the density

$$\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} = \frac{e^{\sum_{m=2}^{K-1}\xi_m Y_m}}{\mathbb{E}\left(e^{\sum_{m=2}^{K-1}\xi_m Y_m}\right)}.$$

Again, we write $\mathbb{E}_{\mathbb{Q}}$ for the expectation under $\mathbb{Q}$ and $\mathbb{E}$ for the expectation under $\mathbb{P}$. We have

$$\mathbb{P}(J_x(0)) = \mathbb{E}\left(e^{\sum_{m=2}^{K-1}\xi_m Y_m}\right) e^{-\sum_{m=2}^{K-1}\xi_m x}\, \mathbb{E}_{\mathbb{Q}}\left(e^{-\sum_{m=2}^{K-1}(\xi_m Y_m - x)}; J_x(0)\right) \ . \tag{2.43}$$

Lemma 2.5 is applied to evaluate the exponential moment (with $\boldsymbol{\xi}_m = (\xi_m, 0)$). It yields

$$\mathbb{E}\left(e^{\sum_{m=2}^{K-1}\lambda_m Y_m}\right) = \left(1 + \mathrm{O}\left(e^{-N^{1/(10K)}}\right)\right) e^{\sum_{m=2}^{K-1}\lambda_m^2\sigma_m^2} \ .$$

In view of (2.43) and the above, it remains to show that

$$\mathbb{E}_{\mathbb{Q}}\left(e^{-\sum_{m=2}^{K-1}\xi_m(Y_m - x)}; J_x(0)\right) \ge \prod_{m=2}^{K-1} \eta_{0,\sigma_m^2}\left(e^{-\frac{xy}{\sigma_m^2}} \mathbb{1}_{[0,\infty)}(y)\right) + \mathrm{O}_K\left(N^{-c}\right). \tag{2.44}$$

This is done by Fourier inversion.

Let $t_m \in \mathbb{R}$ for $m = 2, \dots, K-1$ with $|t_m| < N^{1/(32K)}$. Then $|\xi_m + \mathrm{i}t_m| < N^{1/(16K)}$ so that Lemma 2.5 can be applied with $\xi_m + \mathrm{i}t_m$ in place of $\xi_m$. The Fourier transform of $(Y_m - x; m = 2, \dots, K-1)$ under $\mathbb{Q}$ becomes:

$$\begin{aligned} \mathbb{E}_{\mathbb{Q}}\left(e^{\mathrm{i}\sum_{m=2}^{K-1}t_m(Y_m - x)}\right) &= \frac{\mathbb{E}\left(e^{\sum_{m=2}^{K-1}(\xi_m + \mathrm{i}t_m)Y_m}\right)}{\mathbb{E}\left(e^{\sum_{m=2}^{K-1}\xi_m Y_m}\right)} e^{-i\sum_{m=2}^{K-1}t_m \cdot x} \\ &= \left(1 + \mathrm{O}\left(e^{-N^{1/(10K)}}\right)\right) \exp\left(-\frac{1}{2}\sum_{m=2}^{K-1} t_m^2 \sigma_m^2\right) \ . \end{aligned} \tag{2.45}$$

To complete the argument, consider the function $g_\lambda : \mathbb{R} \to \mathbb{R}$ where

$$g_\xi(y) = \begin{cases} 0 & \text{if } y \leq 0, \\ e^{-\xi y} & \text{if } y > N^{1/(32K)}. \end{cases}$$

and $g_\xi$ is linearly interpolated on $[0, N^{-1/(32K)}]$. Note that $g_\xi$ is bounded by 1 and has Lipschitz constant at most $N^{1/(32K)}$. By definition,

$$\mathbb{E}_\mathbb{Q}\left(e^{-\sum_{m=2}^{K-1}\xi_m(Y_m - x)}; J_x(0)\right) \geq \mathbb{E}_\mathbb{Q}\left(\prod_{m=2}^{K-1} g_{\xi_m}(Y_m - x)\right). \tag{2.46}$$

Lemma 2.6 can be applied with $T = N^{1/(16K)}$, $R = N^{(1/16K)}$ and (2.45). The right-hand side of (2.46) becomes

$$\begin{aligned} &\prod_{m=2}^{K-1} \eta_{0,\sigma_m^2}(g_{\xi_m}(y)) + \mathrm{O}\left(N^{1/(32K)-1/(16K)}\right) + \mathrm{O}\left(N^{2/(16K)}N^{2/(16K)}e^{-N^{1/(10K)}}\right) \\ &+ \otimes_{m=2}^{K-1}\eta_{0,\sigma_m^2}\left(\left\{\left(-N^{1/(16K)}, N^{1/(16K)}\right)\right\}^c\right) + \mathbb{Q}\left(\exists m : |Y_m - x| > N^{1/(16K)}\right) . \end{aligned} \tag{2.47}$$

We can proceed as in (2.34) and in (2.35) to conclude that

$$\mathbb{E}_\mathbb{Q}\left(e^{-\sum_{m=2}^{K-1}\xi_m(Y_m - x)}; J_x(0)\right) \geq \prod_{m=2}^{K-1} \eta_{0,\sigma_m^2}(g_{\xi_m}(y)) + \mathrm{O}(N^{-1/(32K)}) . \tag{2.48}$$

Note that $\left|\eta_{0,\sigma_m^2}(g_{\xi_m}) - \eta_{0,\sigma_m^2}\left(e^{-\xi y}\mathbb{1}_{[0,\infty)}(y)\right)\right| \leq N^{-1/(32K)}$. This with (2.48) shows (2.44). □

We now have all estimates needed to prove the second moment estimate in Proposition 2.4.

*Proof of Proposition 2.4.* Linearity of expectations and Proposition 2.9 directly imply that

$$\mathbb{E}(Z)^2 \geq N^2(1 + \mathrm{o}_K(1))\, e^{-\sum_{m=2}^{K-1}\frac{x^2}{\sigma_m^2}}\left(\prod_{m=2}^{K-1}\eta_{0,\sigma_m^2}\left(e^{-\frac{xy}{\sigma_m^2}}\mathbb{1}_{[0,\infty)}(y)\right)\right)^2 . \tag{2.49}$$

The choice (2.12) of $x$ and the bounds (2.21) on the Gaussian probability implies

$$\mathbb{E}(Z)^2 \geq c(\log N)^{-(K-1)/2}\; N^2 e^{-\sum_{m=2}^{K-1}\frac{x^2}{\sigma_m^2}} \tag{2.50}$$

The second moment can be split in terms of $h_1 \wedge h_2$:

$$\mathbb{E}(Z^2) = \Bigg(\sum_{\substack{h_1,h_2\in\mathcal{H}_N \\ h_1\wedge h_2 \leq \frac{1}{2K}\log N}} + \sum_{\substack{h_1,h_2\in\mathcal{H}_N \\ \frac{1}{2K}\log N \leq h_1\wedge h_2 \leq \frac{1}{K}\log N}} + \sum_{j=2}^{K}\sum_{\substack{h_1,h_2\in\mathcal{H}_N \\ \frac{j-1}{K}\log N < h_1\wedge h_2 \leq \frac{j}{K}\log N}}\Bigg)\,\mathbb{P}\left(J_x(h_1)\cap J_x(h_2)\right) . \tag{2.51}$$

The first sum is smaller than $(1 + \mathrm{o}_K(1))\mathbb{E}\left(Z^2\right)$ by (2.20) and (2.49), and the fact that there are less than $N^2$ terms in the sum. It remains to show that the second and third terms of (2.51) are $\mathrm{o}(\mathbb{E}\,(Z)^2)$. There are $\mathrm{O}(N^{2-1/(2K)})$ terms in the second sum. We can thus apply the rough estimate (2.19) to get that this sum is smaller than

$$\mathrm{O}\left(N^{2-1/(2K)}\exp\left(-\sum_{m=2}^{K-1}\frac{x^2}{\sigma_m^2}\right)\right) .$$

The estimate (2.50) then directy implies that the second sum is $\mathrm{o}(\mathbb{E}\,(Z)^2)$. As for the third term of (2.51), for a fixed $j = 2, \ldots, K-1$, there are at most $\mathrm{O}(N^{2-(j-1)/K})$ pairs such that $\frac{j-1}{K}\log N < h_1\wedge h_2 \leq \frac{j}{K}\log N$, Applying the two-point probability estimate (2.37), one gets that the contribution of such pairs for a fixed $j$ is

$$\mathrm{O}\left(\left(N^2 e^{-\sum_{m=2}^{K-1}\frac{x^2}{\sigma_m^2}}\right)\ N^{-(j-1)/K}\exp\left(\sum_{m=2}^{j}\frac{x^2}{2\sigma_2^2}\right)\right)\ . \tag{2.52}$$

The term in parenthesis is smaller than $C(\log N)^{(K-1)/2}\mathbb{E}(Z)^2$ by (2.50), whereas the other terms are, by the choice of $x$ and (2.8), of the order of

$$N^{-\frac{j-1}{K}}N^{(1-\varepsilon/2)^2\frac{j-1}{K}} = \mathrm{O}_K(N^{-c(j-1)})\ ,$$

for $c = c(\varepsilon)$. The sum of (2.52) from $j = 2$ to $K-1$ is therefore of order $\mathrm{O}_K(N^{-c})$. Altogether this implies that the third term of (2.51) is $\mathrm{o}_K(\mathbb{E}\,(Z)^2)$, and concludes the proof of the proposition. □

## 3 Extension to the full sum and to the continuous interval

*3.1 Lower bound.* In this section, we strengthen the lower bound (2.7) of the previous section to get:

**Proposition 3.1.** *For any $\varepsilon > 0$,*

$$\lim_{N\to\infty}\mathbb{P}\left(\max_{h\in[0,2\pi]}\sum_{j=1}^{\infty}-\frac{\mathrm{Re}(e^{-\mathrm{i}jh}\mathrm{Tr}\mathrm{U}_N^j)}{j}\geq(1-\varepsilon)\log N\right)=1\ . \tag{3.1}$$

To prove this, we need the following exponential moment bound for the tail of the sum.

**Lemma 3.2.** *For any fixed $\delta\in(0,1)$ and any $C>0$ we have for $N$ large enough that for all $h\in\mathbb{R}$ and $|\alpha|\leq C$,*

$$\mathbb{E}\left(\exp\left(\alpha\sum_{j\geq N^{1-\delta}}-\frac{\mathrm{Re}(e^{-\mathrm{i}jh}\mathrm{Tr}\,\mathrm{U}_N^j)}{j}\right)\right)=\exp\left(\left(\frac{1}{4}+\mathrm{o}(1)\right)\alpha^2\delta\log N\right), \tag{3.2}$$

The bound is proved in Section 5 in the form of Proposition 5.1.

*Proof of Proposition 3.1.* We show that

$$\lim_{N\to\infty}\mathbb{P}\left(\max_{h\in\mathcal{H}_N}\sum_{j=1}^{\infty}-\frac{Re(e^{-\mathrm{i}jh}\mathrm{Tr}\mathrm{U}_N^j)}{j}\geq(1-\varepsilon)\log N\right)=1, \tag{3.3}$$

from which (3.1) trivially follows. Using (3.2) with $\alpha=-2x/(\delta\log N)$ we get that

$$\mathbb{P}\left(\sum_{j\geq N^{1-\delta}}-\frac{\mathrm{Re}(e^{-\mathrm{i}jh}\mathrm{Tr}\mathrm{U}_N^j)}{j}\leq -x\right)\leq\exp(-cx^2/(\delta\log N)),\quad\text{for all } 0\leq x\leq\log N. \tag{3.4}$$

A simple union bound over the $N$ points of $\mathcal{H}_N$, like in (2.6), now shows that

$$\mathbb{P}\left(\min_{h\in\mathcal{H}_N}\sum_{j\geq N^{1-\delta}}-\frac{\mathrm{Re}(e^{-\mathrm{i}jh}\mathrm{Tr}\mathrm{U}_N^j)}{j}\leq-\frac{1}{2}\varepsilon\log N\right)\leq N^{1-c\varepsilon^2/\delta}. \tag{3.5}$$

Thus given $\varepsilon>0$, $\delta$ can be set small enough such that (3.5) is o(1), and such that (2.7), with $\frac{1}{2}\varepsilon$ in place of $\varepsilon$, is satisfied. Combining these implies (3.3), and thus also (3.1). □

*3.2 Upper bound* In this section, we strengthen the upper bound (2.5) by removing the discretization and truncation, to arrive at:

**Proposition 3.3.** *For any $\varepsilon > 0$,*

$$\lim_{N\to\infty} \mathbb{P}\left( \max_{h\in[0,2\pi]} \sum_{j=1}^{\infty} -\frac{\mathrm{Re}(e^{-\mathrm{i}jh}\mathrm{Tr}\mathrm{U}_N^j)}{j} \geq (1+\varepsilon)\log N \right) = 0 \ . \tag{3.6}$$

The proof is split into three steps. The first step is to extend (2.5) to a bound for the truncated sum over all of $[0, 2\pi]$ using a chaining argument. In the second step we restrict once again to a discrete set, but one containing $N^C$ equidistant points in $[0, 2\pi]$ for a large $C$, and show that the largest error made in the truncation over this denser discrete set is negligible compared to the leading order $\log N$. Thus we obtain a bound for the full sum over the denser discrete set. Finally we use a rough control of the derivative of the characteristic polynomial to show that the maximum over the denser set is close to the maximum over $[0, 2\pi]$.

To carry out the first step, we need a tail estimate for the difference between the truncated sum at two different but close points (at distance at most $N^{-(1-\delta)}$), that is for

$$\sum_{j=1}^{N^{1-\delta}} -\frac{\mathrm{Re}(e^{-\mathrm{i}jh}\mathrm{Tr}\mathrm{U}_N^j)}{j} - \sum_{j=1}^{N^{1-\delta}} -\frac{\mathrm{Re}(\mathrm{Tr}\mathrm{U}_N^j)}{j}, \ |h| \leq N^{-1-\delta}. \tag{3.7}$$

Using (1.8) one can compute the covariance matrix $\mathbf{\Sigma}$ of the two sums in (3.7) exactly; it turns out to be:

$$\mathbf{\Sigma} = \begin{pmatrix} \sigma^2 & \rho \\ \rho & \sigma^2 \end{pmatrix} \ , \text{ for } \sigma^2 = \frac{1}{2}\sum_{j=1}^{N^{1-\delta}} \frac{1}{j} \text{ and } \rho = \frac{1}{2}\sum_{j=1}^{N^{1-\delta}} \frac{\cos(jh)}{j}. \tag{3.8}$$

If $|h| \leq N^{-(1-\delta)}$ then $|jh| \leq 1$ for $j \leq N^{1-\delta}$, which implies that $\cos(jh) = 1 + \mathrm{O}(j^2h^2)$ and consequently $\rho \geq \sigma^2 - cN^{2(1-\delta)}h^2$. This reflects the fact that as $h$ decreases below scale $N^{-(1-\delta)}$ the correlation no longer behaves as the log of the inverse of $h$, but rather approaches 1 as a quadratic, so that the difference (3.7) has variance $cN^{2(1-\delta)}h^2$, decreasing quadratically in $h$. To obtain a corresponding tail bound, we need the following exponential moment bound (as the similar Lemma 2.5, it follows from the Riemann-Hilbert techniques of Section 5.

**Lemma 3.4.** *Let $\delta > 0$, $\varepsilon \in (0, \delta)$ be fixed. There exists $C > 0$ such that for all $h \in \mathbb{R}$ and real $\xi$ with $|\xi h N^{1-\delta}| \leq N^{\delta-\varepsilon}$, we have*

$$\mathbb{E}\left( \exp\left( \xi \left( \sum_{j=1}^{N^{1-\delta}} -\frac{\mathrm{Re}(e^{-\mathrm{i}jh}\mathrm{Tr}\mathrm{U}_N^j)}{j} - \sum_{j=1}^{N^{1-\delta}} -\frac{\mathrm{Re}(\mathrm{Tr}\mathrm{U}_N^j)}{j} \right)\right)\right) \leq C\exp\left(C\xi^2|\rho-\sigma^2|\right), \tag{3.9}$$

*where we used the definitions* (3.8).

We have $|\rho - \sigma^2| \leq C\sum_{j=1}^{N^{1-\delta}} \frac{(jh)^2}{j} \leq CN^{2-2\delta}h^2$. Hence, using (3.9) with $\xi = (hN^{1-\delta})^{-1}$ and the exponential Chebyshev inequality we get that

$$\mathbb{P}\left( \sum_{j=1}^{N^{1-\delta}} -\frac{\mathrm{Re}(e^{-\mathrm{i}jh}\mathrm{Tr}\mathrm{U}_N^j)}{j} - \sum_{j=1}^{N^{1-\delta}} -\frac{\mathrm{Re}(\mathrm{Tr}\mathrm{U}_N^j)}{j} \geq xhN^{1-\delta} \right) \leq C\exp\left(-cx\right), \tag{3.10}$$

for all $x \geq 1$. The bound (2.5) can now be extended to the set $[0, 2\pi]$.

**Lemma 3.5.** *There is a constant $c$ such that for $0 < \delta < 1$*

$$\lim_{N\to\infty} \mathbb{P}\left( \max_{h\in[0,2\pi]} \sum_{\ell=1}^{(1-\delta)\log N} W_\ell(h) \geq \log N + c \right) = 0 \ . \tag{3.11}$$

*Proof.* The proof uses a chaining argument on dyadic intervals. We actually show (3.11) with the maximum over $[0, 2\pi]$ replaced by a maximum over the set $\cup_{n\geq 0}\mathcal{H}_{N2^n}$. Since this set is dense in $[0, 2\pi]$ and $h \to \sum_{\ell=1}^{(1-\delta)\log N} W_\ell(h)$ is continuous this implies (3.11).

For simplicity, define the random variable

$$X(h) = \sum_{\ell=1}^{(1-\delta)\log N} W_\ell(h) \ .$$

Consider $h \in \cup_{n\geq 0} H_{N2^n}$. For $k \geq 0$ consider the sets $\mathcal{H}_{N2^k}$ with $N2^k$ equidistant points on $[0, 2\pi]$. Define the sequence $(h_k, k \geq 0)$ as follows: if $h \in [\frac{2\pi j}{N2^k}, \frac{2\pi(j+1)}{N2^k})$ for some $j = 0, \dots, N2^k - 1$, then $h_k = \frac{2\pi j}{N2^k}$. Note that $h_{k+1} - h_k$ is 0 or $\frac{2\pi}{N2^{k+1}}$. It holds trivially that

$$X(h) - X(h_0) = \sum_{k=0}^{\infty} \left(X(h_{k+1}) - X(h_k)\right) \ . \tag{3.12}$$

Consider the event

$$A = \left\{ \left| X(h') - X\left(h' + \frac{2\pi}{N2^{k+1}}\right)\right| \leq \frac{k+1}{2^k} \quad \forall h' \in \mathcal{H}_{N2^k},\ \forall k \geq 0 \right\} \ .$$

Since the sequence $(k+1)/2^k$ is summable, it is clear from (3.12) that for all $h \in \cup_{n\geq 0}\mathcal{H}_{N2^n}$,

$$X(h) = X(h_0) + \mathrm{O}(1) \ , \text{ on the event } A.$$

Therefore, the maximum over $\cup_{n\geq 0}\mathcal{H}_{N2^n}$ can only differ by a constant from the one on $\mathcal{H}_N$. The conclusion thus follows from Proposition 2.2 after it is shown that $\mathbb{P}(A^c)$ tends to 0. A straightforward union bound yields

$$\mathbb{P}(A^c) = \sum_{k=0}^{\infty} \sum_{h' \in \mathcal{H}_{N2^k}} \mathbb{P}\left( \left| X(h') - X\left(h' + \frac{2\pi}{N2^{k+1}}\right)\right| > \frac{k+1}{2^k} \right) \ . \tag{3.13}$$

The bound (3.10) is used with $h = \frac{2\pi}{N2^{k+1}}$ and $x = (\frac{k+1}{2^k})/(hN^{1-\delta}) \geq c(k+1)N^\delta$ to obtain that

$$\mathbb{P}\left( \left| X(h') - X\left(h' + \frac{2\pi}{N2^{k+1}}\right)\right| > \frac{k+1}{2^k} \right) \leq 2\exp\left(-c(k+1)N^\delta\right) \ .$$

Therefore, we get from (3.13) the estimate

$$\mathbb{P}(A^c) \leq \sum_{k=0}^{\infty} N2^{k+1} \exp\left(-c(k+1)N^\delta\right) \leq ce^{-cN^\delta},$$

which goes to 0 as $N \to \infty$. □

Next we bound the "tail" on a the set $\mathcal{H}_{N^{100}}$ (the choice $N^{100}$ is somewhat arbitrary, but since our proof is based on a simple union bound we can not obtain the result over all of $[0, 2\pi]$).

**Lemma 3.6.** *For every $0 < \varepsilon < 1$ there exists $\delta = \delta(\varepsilon) < 1$ such that,*

$$\lim_{N\to\infty} \mathbb{P}\left( \max_{h \in \mathcal{H}_{N^{100}}} \sum_{j \geq N^{1-\delta}} -\frac{\mathrm{Re}(e^{-\mathrm{i}jh}\mathrm{Tr}\mathrm{U}_N^j)}{j} \geq \varepsilon \log N \right) = 0 \ . \tag{3.14}$$

*Proof.* By a union bound and rotational invariance, this probability is smaller than

$$N^{100}\mathbb{P}\left(\sum_{j\geq N^{1-\delta}} -\frac{\mathrm{Re}(\mathrm{Tr}\mathrm{U}_N^j)}{j} \geq \varepsilon \log N\right) .$$

Using the exponential Chebyshev inequality and Lemma 3.2 with $\alpha = \varepsilon/\delta$ this is at most

$$N^{100} N^{-c\varepsilon^2/\delta}.$$

This tends to zero if $\delta$ is chosen small enough, depending on $\varepsilon$. □

Lemmas 3.5 and 3.6 combine to give us

**Lemma 3.7.** *For every* $0 < \varepsilon < 1$,

$$\lim_{N\to\infty} \mathbb{P}\left(\max_{h\in\mathcal{H}_{N^{100}}} \sum_{j=1}^{\infty} -\frac{\mathrm{Re}(e^{-\mathrm{i}jh}\mathrm{Tr}\mathrm{U}_N^j)}{j} \geq (1+\varepsilon)\log N\right) = 0 . \tag{3.15}$$

It is now possible to prove the upper bound for the maximum from a crude control on the derivative.

*Proof of Proposition 3.3.* To estimate the derivative in a neighborhood of a maximizer, we need to estimate how close the maximizer can be to an eigenvalue. Let $e^{\mathrm{i}\theta_1}, e^{\mathrm{i}\theta_2}, \ldots, e^{\mathrm{i}\theta_N}$ denote the eigenvalues of the random matrix $\mathrm{U}_N$. It is helpful to consider the event where the eigenvalues are not too close to each other:

$$B = \left\{\inf_{j\neq k} \|\theta_j - \theta_k\| > N^{-90}\right\} .$$

It was shown in [10] (see Theorem 1.1) that $\lim_{N\to\infty}\mathbb{P}(B^c) = 0$. (In fact, the authors show that the smallest gap is of order $N^{-4/3}$).

It remains to estimate the probability restricted on the event $B$. We will show that on this event

$$\frac{\mathrm{d}}{\mathrm{d}h}\log|\mathrm{P}_N(e^{\mathrm{i}h})| \leq cN^{92}, \qquad \forall h \text{ such that } |h-h^\star| \leq N^{-100} , \tag{3.16}$$

where $h^\star$ is a maximizer of $\log|P_N(e^{\mathrm{i}h})|$. In particular,

$$\log|\mathrm{P}_N(e^{\mathrm{i}h^\star})| \leq \log|\mathrm{P}_N(e^{\mathrm{i}h})| + cN^{-98} \qquad \forall h \text{ such that } |h-h^\star| \leq N^{-100} .$$

Since there must be an $h\in\mathcal{H}_{N^{100}}$ with $|h-h^\star|\leq N^{-100}$, this implies that

$$\mathbb{P}\left(\{\log|\mathrm{P}_N(e^{\mathrm{i}h^\star})| \geq (1+\varepsilon)\log N\}\cap B\right) \leq \mathbb{P}\left(\max_{h\in\mathcal{H}_{N^{100}}} \log|\mathrm{P}_N(e^{\mathrm{i}h^\star})| \geq (1+\varepsilon)\log N + cN^{-98}\right) \to 0$$

by Lemma 3.7.

To prove (3.16), notice that the derivative is

$$\frac{\mathrm{d}}{\mathrm{d}h}\log|\mathrm{P}_N(e^{\mathrm{i}h})| = \mathrm{Re}\sum_{j=1}^{N} \frac{\mathrm{i}e^{\mathrm{i}h}}{e^{\mathrm{i}h}-e^{\mathrm{i}\theta_j}} \tag{3.17}$$

Suppose without loss of generality that $\theta_1$ is the closest eigenangle to $h^\star$. Then we must have for $j\neq 2$ that $|\theta_j - h^\star| > \frac{1}{2}\inf_{j\neq k}|\theta_j-\theta_k| > \frac{1}{2}N^{-90}$ on the event $B$. Since the derivative is 0 at $h^\star$, this can be used to bound $|\theta_1 - h^\star|$, namely

$$\left|\mathrm{Re}\frac{\mathrm{i}e^{\mathrm{i}h^\star}}{e^{\mathrm{i}h^\star}-e^{\mathrm{i}\theta_1}}\right| = \left|\mathrm{Re}\sum_{j=2}^{N}\frac{\mathrm{i}e^{\mathrm{i}h^\star}}{e^{\mathrm{i}h^\star}-e^{\mathrm{i}\theta_j}}\right| \leq \sum_{j=2}^{N}\frac{1}{|e^{\mathrm{i}h^\star}-e^{\mathrm{i}\theta_j}|} \leq cN^{91}.$$

This shows by a Taylor expansion that $|\theta_1 - h^\star| > N^{-91}$. This also means that for every $h$ such that $|h - h^\star| \le N^{-100}$, we must have $|\theta_j - h| > cN^{-91}$. Putting this back in (3.17) shows that

$$\frac{\mathrm{d}}{\mathrm{d}h} \log |\mathrm{P}_N(e^{\mathrm{i}h})| \le cN^{92} ,$$

as claimed. □

To complete the proof of the main result Theorem 1.2 it now only remains to show the exponential moment/characteristic function bounds in Lemmas 2.1, 2.5, 3.2 and 3.4. These will be proved in Section 5.

## 4 High points and free energy

In this section we prove Theorem 1.3 about the Lebesgue measure of high points, and derive from it Corollary 1.4 about the free energy.

Recall the definition (1.3) of the set $\mathcal{L}_N(\gamma)$ of $\gamma-$high points. The first goal is to prove Theorem 1.3, i.e. to show that $\mathrm{Leb}(\mathcal{L}_N(\gamma)) = N^{-\gamma^2+\mathrm{o}(1)}$ with high probability. For the upper bound we will be able to work with the full sum (i.e. the logarithm of the characteristic polynomial without truncation). The following exponential moment bound for the full sum, which is obtained from the Selberg integral (see Lemma 5.10), will be used.

**Lemma 4.1.** *For any fixed $C > 0$ we have uniformly for $|\alpha| \le C$ that*

$$\mathbb{E}\left(\exp\left(\alpha \sum_{i=1}^{\infty} -\frac{\mathrm{Re}\left(\mathrm{Tr}\mathrm{U}_N^j\right)}{j}\right)\right) = e^{\alpha^2\left(\frac{1}{4}+\mathrm{o}(1)\right)\log N}. \tag{4.1}$$

*Proof of Theorem 1.3.* The upper bound is direct: Fubini's theorem and rotational invariance show that

$$\mathbb{E}\Big(\mathrm{Leb}\big(\mathcal{L}_N(\gamma)\big)\Big) = 2\pi\mathbb{P}\left(\sum_{j=1}^{\infty} -\frac{\mathrm{Re}\left(\mathrm{Tr}\mathrm{U}_N^j\right)}{j} \ge \gamma \log N\right).$$

The exponential Chebyshev inequality and (4.1) with $\alpha = 2\gamma$ show that the latter probability is at most $cN^{-\gamma^2+\frac{\varepsilon}{2}}$ for any $\varepsilon > 0$ and large enough $N$, so that

$$\text{for all } \varepsilon > 0,\ \mathbb{P}\Big(\mathrm{Leb}\big(\mathcal{L}_N(\gamma)\big) \ge N^{-\gamma^2+\varepsilon}\Big) \le cN^{-\frac{\varepsilon}{2}} \to 0, \text{ as } N \to \infty .$$

This gives the upper bound of (1.4).

The proof of the lower bound is very similar to the proof of Proposition 2.3, which gave a lower bound for the maximum of the truncated sum. Fix a $\gamma \in (0,1)$ and an $\varepsilon > 0$ and recall the coarse increments $Y_m(h), m = 1, \ldots, K$, from (1.18), as well as the event $J_x(h)$ from (2.11). Here we will use

$$x = \frac{\gamma}{K}\left(1 + \frac{\varepsilon}{3}\right)\log N,$$

and apply the second moment method to the measure of the set

$$\mathcal{L}_N^K = \{h \in [0, 2\pi] : J_x(h) \text{ occurs}\}.$$

Note that if $h \in \mathcal{L}_N^K$ and $K$ is large enough depending on $\varepsilon$, then

$$\sum_{N^{1/K} \le j < N^{1-1/K}} -\frac{\mathrm{Re}\left(\mathrm{Tr}\mathrm{U}_N^j\right)}{j} \ge \frac{K-2}{K}\gamma\left(1 + \frac{\varepsilon}{3}\right)\log N \ge \gamma\left(1 + \frac{\varepsilon}{4}\right)\log N. \tag{4.2}$$

Consider the subsets

$$\mathcal{A}=\left\{h\in[0,2\pi]:\sum_{j\geq N^{1-1/K}}-\frac{\mathrm{Re}\left(\mathrm{Tr}\mathrm{U}_N^j\right)}{j}\leq-\frac{\varepsilon}{8}\log N\right\},$$

and

$$\mathcal{B}=\left\{h\in[0,2\pi]:\sum_{j<N^{1/K}}-\frac{\mathrm{Re}\left(\mathrm{Tr}\mathrm{U}_N^j\right)}{j}\leq-\frac{\varepsilon}{8}\log N\right\}.$$

Because of (4.2) we have (for $K$ large enough)

$$\mathcal{L}_N^K\subset\mathcal{L}_N(\gamma)\cup\mathcal{A}\cup\mathcal{B},$$

so that

$$\mathrm{Leb}(\mathcal{L}_N(\gamma))\geq\mathrm{Leb}\left(\mathcal{L}_N^K\right)-\mathrm{Leb}(\mathcal{A})-\mathrm{Leb}(\mathcal{B}).$$

The bounds (2.4) (with $\delta=1-\log N/K$) and (3.4) together with Fubini's theorem show that $\mathbb{E}(\mathrm{Leb}(\mathcal{A}))\leq cN^{-cK\varepsilon^2}$ and $\mathbb{E}(\mathrm{Leb}(\mathcal{B}))\leq cN^{-cK\varepsilon^2}$, so that if $K$ is chosen large enough depending on $\varepsilon$,

$$\lim_{N\to\infty}\mathbb{P}\left(\mathrm{Leb}(\mathcal{A})+\mathrm{Leb}(\mathcal{B})\leq N^{-\gamma^2-\varepsilon}\right)=1.$$

It thus suffices to show that for large enough $K$,

$$\lim_{N\to\infty}\mathbb{P}\left(\mathrm{Leb}\left(\mathcal{L}_N^K\right)\geq 2N^{-\gamma^2-\varepsilon}\right)=1. \tag{4.3}$$

To apply the second moment method to $\mathrm{Leb}\left(\mathcal{L}_N^K\right)$, we first note that by (2.42) the first moment satisfies

$$\mathbb{E}\left(\mathrm{Leb}\left(\mathcal{L}_N^K\right)\right)\geq 2\pi\left(1+\mathrm{o}_K(1)\right)e^{-\sum_{m=2}^{K-1}\frac{x^2}{2\sigma_m^2}}\prod_{m=2}^{K-1}\eta_{0,\sigma_m^2}\left(e^{-\frac{xy}{\sigma_m^2}}\mathbb{1}_{[0,\infty)}(y)\right). \tag{4.4}$$

As in the last inequality of (4.2), the right-hand side is at least $N^{-\frac{K-2}{K}\gamma\left(1+\frac{\varepsilon}{3}\right)^2}c\left(\log N\right)^{-(K-2)}\geq N^{-\gamma^2-\frac{\varepsilon}{2}}$, where the last inequality follows for $K$ large enough depending on $\varepsilon$. Therefore,

$$\mathbb{E}\left(\mathrm{Leb}\left(\mathcal{L}_N^K\right)\right)\geq N^{-\gamma^2-\frac{\varepsilon}{2}}.$$

Using the Paley-Zygmund inequality as for the maximum, this proves that

$$\mathbb{P}\left(\mathrm{Leb}\left(\mathcal{L}_N^K\right)\geq N^{-\gamma^2-\varepsilon}\right)\geq\mathbb{P}\left(\mathrm{Leb}\left(\mathcal{L}_N^K\right)\geq N^{-\frac{\varepsilon}{2}}\,\mathbb{E}\left(\mathrm{Leb}\left(\mathcal{L}_N^K\right)\right)\right)\geq(1-N^{-\frac{\varepsilon}{2}})^2\frac{\mathbb{E}\left(\mathrm{Leb}\left(\mathcal{L}_N^K\right)^2\right)}{\mathbb{E}\left(\mathrm{Leb}\left(\mathcal{L}_N^K\right)\right)^2}. \tag{4.5}$$

To obtain the lower bound of (1.4) it only remains to show that for all $K\geq 1$,

$$\mathbb{E}\left(\mathrm{Leb}\left(\mathcal{L}_N^K\right)^2\right)=\left(1+\mathrm{o}_K(1)\right)\mathbb{E}\left(\mathrm{Leb}\left(\mathcal{L}_N^K\right)\right)^2. \tag{4.6}$$

To prove (4.6), we write the second moment as

$$\mathbb{E}\left(\mathrm{Leb}\left(\mathcal{L}_N^K\right)^2\right)=\int_{[0,2\pi]^2}\mathbb{P}\left(J_x(h_1)\cap J_x(h_2)\right)\mathrm{d}h_1\mathrm{d}h_2,$$

and split the integral in analogy with (2.51):

$$\underbrace{\int_{h_1\wedge h_2\le\frac{1}{2K}\log N}(\cdot)\,\mathrm{d}h_1\mathrm{d}h_2}_{\text{(I)}}+\underbrace{\int_{\frac{1}{2K}\log N<h_1\wedge h_2\le\frac{1}{2K}\log N}(\cdot)\,\mathrm{d}h_1\mathrm{d}h_2}_{\text{(II)}}$$
$$+\underbrace{\sum_{j=2}^{K-1}\int_{\frac{j-1}{K}\log N<h_1\wedge h_2\le\frac{j}{K}\log N}(\cdot)\,\mathrm{d}h_1\mathrm{d}h_2}_{\text{(III)}}+\underbrace{\int_{h_1\wedge h_2\ge\frac{K-1}{K}\log N}(\cdot)\,\mathrm{d}h_1\mathrm{d}h_2}_{\text{(IV)}}\,.$$

The bound (2.20) provides a uniform bound on the integrand in (I), showing that

$$\text{(I)}\le(2\pi)^2\left(1+\mathrm{o}_K(1)\right)e^{-\sum_{m=2}^{K-1}\frac{x^2}{\sigma_m^2}}\left(\prod_{m=2}^{K-1}\eta_{0,\sigma_m^2}\left(e^{-\frac{xy}{\sigma_m^2}}\mathbb{1}_{[0,\infty)}(y)\right)\right)^2=\left(1+\mathrm{o}_K(1)\right)\mathbb{E}\left(\mathrm{Leb}\left(\mathcal{L}_N^K\right)\right)^2,$$

where the equality follows by (4.4). By (2.19) the integrand in (II) is uniformly bounded by $e^{-\sum_{m=2}^{K-1}\frac{x^2}{\sigma_m^2}}\le c\mathbb{E}\left(\mathrm{Leb}\left(\mathcal{L}_N^K\right)\right)^2(\log N)^{\frac{K-2}{2}}$, and since the measure of the set integrated over is at most $(2\pi)^2N^{-\frac{1}{2K}}$ the integral (II) is $\mathrm{o}\left(\mathbb{E}\left(\mathrm{Leb}\left(\mathcal{L}_N^K\right)\right)^2\right)$. Similarily, the measure of the set integrated over in the $j-$th term of (III) is most $(2\pi)^2N^{-\frac{j-1}{K}}$. Therefore by bounding the integrand using (2.37), we see similarly to (2.52) that the integral (III) is $\mathrm{o}\left(\mathbb{E}\left(\mathrm{Leb}\left(\mathcal{L}_N^K\right)\right)^2\right)$. Lastly (IV) can be bounded by

$$\int_{h_1\wedge h_2\ge\frac{K-1}{K}\log N}\mathbb{P}\left(J_x(h_1)\right)\mathrm{d}h_1\mathrm{d}h_2=\mathbb{E}\left(\mathrm{Leb}\left(\mathcal{L}_N^K\right)\right)(2\pi)^2N^{-\frac{K-1}{K}},$$

which is also $\mathrm{o}\left(\mathbb{E}\left(\mathrm{Leb}\left(\mathcal{L}_N^K\right)\right)^2\right)$, see (4.4). This proves (4.6), and therefore also the lower bound of (1.4). □

We can now derive Corollary 1.4 about the free energy.

*Proof of Corollary 1.4.* The proof uses Laplace's method. Let $\varepsilon>0$ and fix $M\in\mathbb{N}$. We define levels

$$\gamma_j=(1+\varepsilon)\frac{j}{M}\qquad j=0,\dots,M\ ,$$

and the event

$$E=\bigcap_{j=1}^{M-1}\left\{N^{-\gamma_j^2-\varepsilon}\le\mathrm{Leb}\left(\mathcal{L}_N(\gamma_j)\right)\le N^{-\gamma_j^2+\varepsilon}\right\}\cap\left\{\mathrm{Leb}\left(\mathcal{L}_N(\gamma_M)\right)=0\right\}.$$

Theorems 1.2 and 1.3 imply that $\mathbb{P}(E^c)\to0$ as $N\to\infty$. Therefore it suffices to prove the result on the event $E$. The integrals $\frac{1}{2\pi}\int_0^{2\pi}|\mathrm{P}_N(h)|^\beta\mathrm{d}h$ can be split in terms of the subsets $\{h\in[0,2\pi]:|\mathrm{P}_N(h)|<1\}$ and $\{h\in[0,2\pi]:N^{\gamma_{j-1}}\le|\mathrm{P}_N(h)|<N^{\gamma_j}\}$. We get the upper bound

$$\begin{aligned}\frac{1}{2\pi}\int_0^{2\pi}|\mathrm{P}_N(h)|^\beta\mathrm{d}h&\le\frac{1}{2\pi}\int_0^{2\pi}|\mathrm{P}_N(h)|^\beta\mathbb{1}_{\{|\mathrm{P}_N(h)|<1\}}\mathrm{d}h+\sum_{j=1}^{M}\frac{1}{2\pi}\int_0^{2\pi}|\mathrm{P}_N(h)|^\beta\mathbb{1}_{\{N^{\gamma_{j-1}}\le|\mathrm{P}_N(h)|<N^{\gamma_j}\}}\mathrm{d}h\\&\le1+\sum_{j=1}^{M}N^{\beta\gamma_j}\mathrm{Leb}\left(\mathcal{L}_N(\gamma_{j-1})\right)\le1+\sum_{j=1}^{M}N^{\beta\gamma_j-\gamma_{j-1}^2+\varepsilon}\ ,\end{aligned}\tag{4.7}$$

where the last inequality holds on the event $E$. Similarly we have on this event that

$$\frac{1}{2\pi}\int_0^{2\pi}|\mathrm{P}_N(h)|^\beta\mathrm{d}\ge\sum_{j=1}^{M}N^{\beta\gamma_{j-1}}\mathrm{Leb}\left(\mathcal{L}_N(\gamma_j)\right)\ge\sum_{j=1}^{M}N^{\beta\gamma_{j-1}-\gamma_j^2-\varepsilon}\ .\tag{4.8}$$

Equations (4.7) and (4.8) imply for $M$ fixed

$$\max_{j=1,\dots,M}\left\{1+\beta\gamma_{j-1}-\gamma_j^2-\varepsilon\right\}+\mathrm{o}(1)\leq \frac{1}{\log N}\log\left(\frac{N}{2\pi}\int_0^{2\pi}|\mathrm{P}_N(h)|^\beta \mathrm{d}h\right)\leq \max_{j=1,\dots,M}\left\{1+\beta\gamma_j-\gamma_{j-1}^2+\varepsilon\right\}+\mathrm{o}(1)\,.$$

After taking the limit $N\to\infty$, $M\to\infty$, and finally $\varepsilon\to 0$, we get

$$\lim_{M\to\infty}\frac{1}{\log N}\log\left(\frac{N}{2\pi}\int_0^{2\pi}|\mathrm{P}_N(h)|^\beta \mathrm{d}h\right)=\max_{\gamma\in[0,1]}\{1+\beta\gamma-\gamma^2\}=\begin{cases}1+\frac{\beta^2}{4} & \text{if } \beta<2\\ \beta & \text{if } \beta\geq 2\end{cases},$$

which proves the corollary. □

Up to Lemma 4.1 (and the other Gaussian estimates from the previous sections), we have now completed the proof of Theorem 1.3 and Corollary 1.4.

## 5 Estimates on increments and tails

This section proves two important estimates.

(i) Proposition 5.1 (which is a reformulation of Lemma 3.2) justifies the approximation of the characteristic polynomial by partial sums, by bounding the contribution of high powers in the Fourier expansion. We follow the proof of asymptotics of Toeplitz determinants when the symbols have a Fisher-Hartwig singularity as given in [24, 25, 47]. In particular we rely on the Riemann-Hilbert problem developped in these works. We will also make use of a key differential identity from [25].

(ii) Proposition 5.11 gives appropriate asymptotics for the joint Laplace and Fourier transforms of sums involving only traces of small powers, at possibly different evaluation points $h$. It will be used to prove the previously stated bounds Lemmas 2.1, 2.5 and 3.4 involving the truncated sum of traces and the increments $Y_m(h)$. The proof of (ii) is easier than that of (i): here the symbols have no singularity

There are several technical differences with [24,25,47]. First, we only need to consider one Fisher-Hartwig singularity (at $z=1$), which simplifies the analysis. Second, our external potential $V$ depends on $N$ and it is close to singular: it corresponds to a singularity smoothed on the mesoscopic scale $N^{-1+\delta}$, $\delta>0$ arbitrarily small. Consequently, the contour of the Riemann-Hilbert problem is $N$-dependent. Third, a small modification of the Riemann-Hilbert problem considered in [24, 25, 47] will be necessary to obtain better error terms (see (5.27)).

*5.1 Setting.* The main task of this section is the following exponential moment estimate.

**Proposition 5.1.** *For any fixed (small) $\delta\in(0,1)$ and (large) $\alpha$, for large enough $N$ we have*

$$\mathbb{E}\left(\exp\left(2\alpha\sum_{j\geq N^{1-\delta}}-\frac{\mathrm{Re}(\mathrm{Tr}\mathrm{U}_N^j)}{j}\right)\right)=\exp\left((1+\mathrm{o}(1))\alpha^2\delta\log N\right), \tag{5.1}$$

Note that this implies Lemma 3.2 by rotational invariance. For the proof, we first introduce the following notations. Let $V$ be analytic in a neighborhood of the unit circle, $\mathscr{C}$. We will actually consider

$$V(z)=\lambda\sum_{j=1}^{N^{1-\delta}}\frac{z^j+z^{-j}}{j},\ \text{i.e. } V_j=\frac{\lambda}{|j|}\mathbb{1}_{|j|\leq N^{1-\delta}}, \tag{5.2}$$

for $0\leq\lambda\leq\alpha$, but most of subsections 5.1 and 5.2 is independent of this specific form of $V$. We also define

$$e^{V^{(t)}(z)}=1-t+te^{V(z)},\ 0\leq t\leq 1 \tag{5.3}$$

$$f^{(t)}(z)=e^{V^{(t)}(z)}|z-1|^{2\alpha}\ . \tag{5.4}$$

Note that $V^{(1)}(z) = V(z)$ and $V^{(t)}(z)$ is clearly well defined for $|z| = 1$, because $1 - t + te^{V(z)} > 0$. With these definitions in mind, instead of (5.55) we will prove the equivalent form (for $\lambda = \alpha$)

$$\mathbb{E}\left(\prod_{k=1}^{N} f^{(1)}(e^{\mathrm{i}\theta_k})\right) = \exp\left((1 + \mathrm{o}(1))\alpha^2\delta\log N\right) \ . \tag{5.5}$$

For the proof, we will use a Riemann-Hilbert approach and therefore need the following lemma about an analytic extension of $V^{(t)}$.

**Lemma 5.2.** *There exists $\varepsilon > 0$ small enough (depending on the fixed constants $\delta$ and $\alpha$) such that for any $0 \le t \le 1$, $V^{(t)}$ admits an analytic extension to $||z| - 1| < \varepsilon N^{-1+\delta}$.*

*Proof.* We only need to prove

$$1 - t + te^{V(z)} \neq 0 \tag{5.6}$$

for any $||z|-1| < \varepsilon N^{-1+\delta}$, $0 \le t \le 1$. Note that $\sup_{||z|-1|<N^{-1+\delta}} |V'(z)| = \mathrm{O}(N^{1-\delta})$ and $V(z) \in \mathbb{R}$ if $|z| = 1$, so that for small enough $\varepsilon$ we have

$$\sup_{||z|-1|<\varepsilon N^{-1+\delta}} |\mathrm{Im}V(z)| < 1/10. \tag{5.7}$$

The above condition clearly implies (5.6). ☐

In the following, we will abbreviate $\kappa = \varepsilon N^{-1+\delta}$, with $\varepsilon$ so that the above lemma holds. Define the Wiener-Hopf factorization of $e^{V^{(t)}}$:

$$b_+^{(t)}(z) = \exp\left(-V_0 + \int_{\mathscr{C}} \frac{1}{2\pi\mathrm{i}} \frac{V^{(t)}(s)}{s - z}\mathrm{d}s\right) \text{ for } |z| < 1, \tag{5.8}$$

$$b_-^{(t)}(z) = \exp\left(\frac{1}{2\pi\mathrm{i}} \int_{\mathscr{C}} \frac{V^{(t)}(s)}{s - z}\mathrm{d}s\right) \text{ for } |z| > 1, \tag{5.9}$$

where $\mathscr{C}$ is the unit circle. Then

$$b_+^{(t)}(z) = e^{\sum_{k=1}^{\infty} V_k^{(t)} z^k}, b_-^{(t)}(z) = e^{\sum_{k=-\infty}^{-1} V_k^{(t)} z^k} \tag{5.10}$$

where $V_k^{(t)} = \frac{1}{2\pi}\int_0^{2\pi} V^{(t)}(e^{\mathrm{i}\theta})e^{-\mathrm{i}k\theta}\mathrm{d}\theta$. By Lemma 5.2, both functions can be slightly extended analytically: $b_+^{(t)}$ to $|z| < 1 + \kappa$ and $b_-^{(t)}$ to $|z| > 1 - \kappa$. In the domain $||z| - 1| < \kappa$, they satisfy the following properties:

$$e^{V^{(t)}(z)} = b_+^{(t)}(z)e^{V_0^{(t)}}b_-^{(t)}(z), \tag{5.11}$$

$$b_+^{(t)}(z) = \overline{b_-^{(t)}(\bar{z}^{-1})}. \tag{5.12}$$

The proof of (5.5) proceeds by interpolation through the parameter $t$. Indeed, writing

$$\mathrm{D}_N(f^{(t)}) = \mathbb{E}\left(\prod_{k=1}^{N} f^{(t)}(e^{\mathrm{i}\theta_k})\right), \tag{5.13}$$

a formula from [25] gives

$$\log \mathrm{D}_N(f^{(1)}) - \log \mathrm{D}_N(f^{(0)}) = NV_0 + \sum_{k=1}^{\infty} kV_kV_{-k} - \alpha\log(b_+^{(1)}(1)b_-^{(1)}(1)) + \int_0^1 E(t)\mathrm{d}t \ ,$$

where the error term $E(t)$ can be expressed from the solution of a Riemann-Hilbert problem, see Corollary 5.4. Hence the main part of the analysis consists in bounding this solution, which is performed in the next two subsections. The proof is then concluded, up to the initial value $\log \mathrm{D}_N(f^{(0)})$ in the interpolation. This term is a Selberg integral, and its asymptotics are given in Lemma 5.10.

*5.2 The Riemann-Hilbert problem.* Consider the Szegő function

$$\exp(g(z)) = \exp\left(\frac{1}{2\pi \mathrm{i}}\int_{\mathscr{C}} \frac{\ln f^{(t)}(s)}{s-z}\mathrm{d}s\right) = \begin{cases} e^{V_0^{(t)}} b_+^{(t)}(z)(z-1)^{\alpha} e^{-\mathrm{i}\pi\alpha} & \text{if } |z|<1, \\ b_-^{(t)}(z)^{-1}(z-1)^{-\alpha} z^{\alpha} & \text{if } |z|>1. \end{cases} \tag{5.14}$$

In the above definition and for further use, the cut of $(z-1)^\alpha$ goes from 1 to $\infty$ along $\theta = 0$. We fix the branches by $0 < \arg(z-1) < 2\pi$, and for $z^{\alpha/2}$, $0 < \arg(z-1) < 2\pi$ as well. Let

$$\zeta = N \log z,$$

where $\log x > 0$ for $x > 1$, and has a cut on the negative half of the real axis. We define the analytic continuation of the function

$$h_\alpha(z) = |z-1|^\alpha$$

through $|z-1|^\alpha = (z-1)^{\alpha/2}(z^{-1}-1)^{\alpha/2} = \frac{(z-1)^\alpha}{z^{\alpha/2}e^{\mathrm{i}\pi\alpha/2}}$, in a neighborhood of the open arc $\{|z|=1, |z|\neq 1\}$. The factor $e^{\mathrm{i}\pi\alpha/2}$ is chosen so that $h_\alpha(z)$ has null argument on the unit circle. Moreover, let

$$F(z) = \begin{cases} e^{\frac{V^{(t)}(z)}{2}} h_\alpha(z) e^{-\mathrm{i}\pi\alpha} & \text{if } \xi \in \mathrm{I}, \mathrm{II}, \mathrm{V}, \mathrm{VI}, \\ e^{\frac{V^{(t)}(z)}{2}} h_\alpha(z) e^{\mathrm{i}\pi\alpha} & \text{if } \xi \in \mathrm{III}, \mathrm{IV}, \mathrm{VII}, \mathrm{VIII}, \end{cases} \tag{5.15}$$

where we used the definitions of Figure 4, and the conventions for cuts and branches were explained previously.

Let $\psi(a,b,x)$ be the confluent hypergeometric function of the second kind. Define $\Psi$ to be the analytic function such that for any $\zeta \in \mathrm{I}$ (see Figure 4) we have

$$\Psi(\zeta) = \begin{pmatrix} \zeta^\alpha \psi(\alpha, 2\alpha+1, \zeta) e^{\mathrm{i}\pi\alpha} e^{-\zeta/2} & -\zeta^\alpha \psi(\alpha+1, 2\alpha+1, e^{-\mathrm{i}\pi}\zeta) e^{\mathrm{i}\pi\alpha} e^{\zeta/2}\alpha \\ -\zeta^{-\alpha}\psi(-\alpha+1, -2\alpha+1, \zeta) e^{-3\mathrm{i}\pi\alpha} e^{-\zeta/2}\alpha & \zeta^{-\alpha}\psi(-\alpha, -2\alpha+1, e^{-\mathrm{i}\pi}\zeta) e^{-\mathrm{i}\pi\alpha} e^{\zeta/2} \end{pmatrix} \tag{5.16}$$

and

$$\Psi_+(\zeta) = \Psi_-(\zeta) K(\zeta)$$

where

$$K(\zeta) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad \text{if } \zeta \in \Gamma_1 \cup \Gamma_5$$
$$K(\zeta) = \begin{pmatrix} e^{\mathrm{i}\pi\alpha} & 0 \\ 0 & e^{-\mathrm{i}\pi\alpha} \end{pmatrix} \quad \text{if } \zeta \in \Gamma_3 \cup \Gamma_7$$
$$K(\zeta) = \begin{pmatrix} 1 & 0 \\ e^{\mathrm{i}\pi 2\alpha} & 1 \end{pmatrix} \quad \text{if } \zeta \in \Gamma_4 \cup \Gamma_8$$
$$K(\zeta) = \begin{pmatrix} 1 & 0 \\ e^{-\mathrm{i}\pi 2\alpha} & 1 \end{pmatrix} \quad \text{if } \zeta \in \Gamma_2 \cup \Gamma_6 \ .$$

Let $\sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ and denote $z^{\sigma_3} = \begin{pmatrix} z & 0 \\ 0 & z^{-1} \end{pmatrix}$.

Define

$$N(z) = e^{g(z)\sigma_3} \quad \text{if } |z|>1$$
$$N(z) = e^{g(z)\sigma_3}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad \text{if } |z|<1 \ .$$

We will also need the notation

$$E(z) = N(z)F(z)^{\sigma_3}\begin{pmatrix} e^{-\mathrm{i}\pi\alpha} & 0 \\ 0 & e^{2\mathrm{i}\pi\alpha} \end{pmatrix} \quad \text{if } z \in \mathrm{I},\ \mathrm{II}$$
$$E(z) = N(z)F(z)^{\sigma_3}\begin{pmatrix} e^{-\mathrm{i}\pi 2\alpha} & 0 \\ 0 & e^{\mathrm{i}\pi 3\alpha} \end{pmatrix} \quad \text{if } z \in \mathrm{III},\ \mathrm{IV}$$
$$E(z) = N(z)F(z)^{\sigma_3}\begin{pmatrix} 0 & e^{\mathrm{i}\pi 3\alpha} \\ -e^{-\mathrm{i}\pi 2\alpha} & 0 \end{pmatrix} \quad \text{if } z \in \mathrm{V},\ \mathrm{VI}$$
$$E(z) = N(z)F(z)^{\sigma_3}\begin{pmatrix} 0 & e^{\mathrm{i}\pi 2\alpha} \\ -e^{-\mathrm{i}\pi\alpha} & 0 \end{pmatrix} \quad \text{if } z \in \mathrm{VII},\ \mathrm{VIII}\ . \tag{5.17}$$

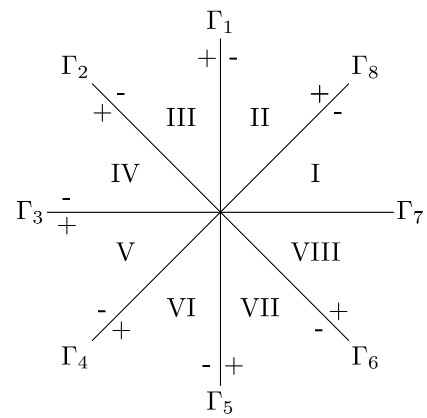


Figure 4: Auxiliary contours in the variable $\zeta$ around 0 (i.e. a neighborhood of the singularity $z=1$)

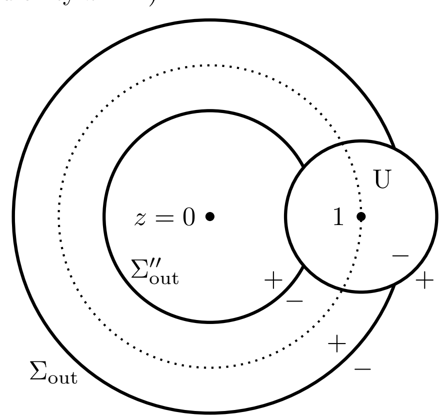


Figure 5: Contour $\Gamma$ for the $R$-Riemann-Hilbert problem

Finally, consider the jump matrix

$$M(z) = E(z)\Psi(z)F(z)^{-\sigma_3}z^{\pm N\sigma_3}N(z)^{-1} \tag{5.18}$$

where the plus sign is taken for $|z| < 1$, and minus for $|z| > 1$.

We define the contour $\Gamma$ in $\mathbb{C}$ as follows (see Figure 5): it consists in the boundary $\partial\mathrm{U}$ of a disk U centered at 1 with radius $\kappa$, the arc of circle $\Sigma_{\mathrm{out}}$ (resp. $\Sigma''_{\mathrm{out}}$) centered at 0 with radius $1+(2/3)\kappa$ (resp. $1-(2/3)\kappa$), outside U with extremities on U.

Consider the following Riemann-Hilbert problem for the $2\times 2$ matrix valued function $R$.

1. $R$ is analytic for $z \in \mathbb{C}\backslash\Gamma$.
2. The boundary values of $R$ are related by the jump condition

$$\begin{aligned}
R_+(z) &= R_-(z)N(z)\begin{pmatrix} 1 & 0 \\ f^{(t)}(z)^{-1}z^{-N} & 1\end{pmatrix}N(z)^{-1} && \text{if } z \in \Sigma_{\mathrm{out}}\\
R_+(z) &= R_-(z)N(z)\begin{pmatrix} 1 & 0 \\ f^{(t)}(z)^{-1}z^{N} & 1\end{pmatrix}N(z)^{-1} && \text{if } z \in \Sigma''_{\mathrm{out}}\\
R_+(z) &= R_-(z)M(z), && \text{if } z \in \partial\mathrm{U}\ .
\end{aligned}$$

3. $R(z) = \mathrm{Id} + \mathrm{O}(1/z)$ as $z \to \infty$.

It was proved in [24] that there exists a unique solution to this Riemann-Hilbert problem and the solution satisfies

$$\det R(z) = 1. \tag{5.19}$$

The following decomposes the ratio $\mathrm{D}_N(f^{(1)})/\mathrm{D}_N(f^{(0)})$ into the main contribution and some error term. It is just a restatement of a differential identity from [25], keeping all error terms explicit. We denote $\dot f = \partial f/\partial t$ and $f' = \partial f/\partial z$.

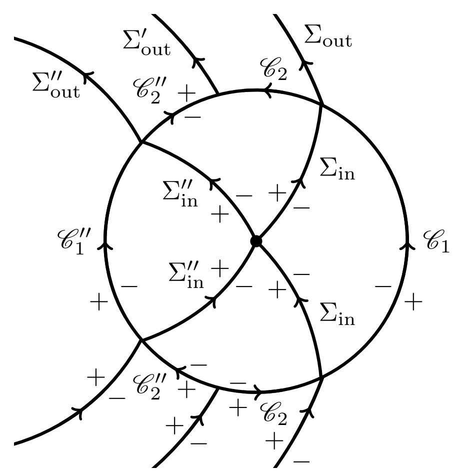

Figure 6: Contours definition and orientation in Proposition 5.3

**Proposition 5.3.** *We have*

$$\log \mathrm{D}_N(f^{(1)}) - \log \mathrm{D}_N(f^{(0)}) = NV_0 + \sum_{k=1}^{\infty} kV_kV_{-k} - \alpha\log(b^1_+(1)b^1_-(1)) + \int_0^1 E(t)\mathrm{d}t. \tag{5.20}$$

*The error term $E(t)$ is defined as (see Figure 6 for the definition and orientation of the contours)*

$$\begin{aligned}
E(t) = &-\int_{\mathscr{C}_1}(R_{11}R'_{22} - R'_{12}R_{21})_-\frac{\dot f^{(t)}}{f^{(t)}}\frac{\mathrm{d}z}{\mathrm{i}2\pi} + \int_{\mathscr{C}''_1}(R'_{11}R_{22} - R_{12}R'_{21})_-\frac{\dot f^{(t)}}{f^{(t)}}\frac{\mathrm{d}z}{\mathrm{i}2\pi}\\
&-\int_{\mathscr{C}_2}(R_{11}R'_{22} - R'_{12}R_{21})_+\frac{\dot f^{(t)}}{f^{(t)}}\frac{\mathrm{d}z}{\mathrm{i}2\pi} + \int_{\mathscr{C}''_2}(R'_{11}R_{22} - R_{12}R'_{21})_+\frac{\dot f^{(t)}}{f^{(t)}}\frac{\mathrm{d}z}{\mathrm{i}2\pi}\\
&+\int_{\Sigma_{\mathrm{out}}} z^{-N}\frac{e^{-2g}}{f^{(t)}}(R'_{22}R_{12} - R'_{12}R_{22})_+\frac{\dot f^{(t)}}{f^{(t)}}\frac{\mathrm{d}z}{\mathrm{i}2\pi} + \int_{\Sigma''_{\mathrm{out}}} z^{N}\frac{e^{2g}}{f^{(t)}}(R'_{21}R_{11} - R'_{11}R_{21})_-\frac{\dot f^{(t)}}{f^{(t)}}\frac{\mathrm{d}z}{\mathrm{i}2\pi}\\
&+\int_{\Sigma'_{\mathrm{out}}}((R'_{11}R_{22} - R_{12}R'_{21}) - (R_{11}R'_{22} - R'_{12}R_{21}))\frac{\dot f^{(t)}}{f^{(t)}}\frac{\mathrm{d}z}{\mathrm{i}2\pi}\ .
\end{aligned}$$

*Proof.* Define

$$S(z)=R(z)N(z),\ I=(S'_{22}S_{12}-S'_{12}S_{22})/f^{(t)},\ J=S'_{22}S_{11}-S'_{12}S_{21}.$$

Simple calculations give

$$\begin{array}{ll} I=\frac{e^{-2g}}{f^{(t)}}(R'_{22}R_{12}-R_{22}R'_{12}) & \text{for } |z|>1\\ I=\frac{e^{2g}}{f^{(t)}}(R_{11}R'_{21}-R'_{11}R_{21}) & \text{for } |z|<1\\ J=-g'+(R_{11}R'_{22}-R'_{12}R_{21}) & \text{for } |z|>1\\ J=g'+(R'_{11}R_{22}-R_{12}R'_{21}) & \text{for } |z|<1 \end{array}$$

From [25, equations (5.69), (5.70), (5.71)] we have

$$\frac{\partial}{\partial t}\log D_N(f^{(t)})=N\int_{\mathscr{C}}\frac{\dot f^{(t)}}{f^{(t)}}\frac{\mathrm{d}z}{\mathrm{i}2\pi z}+\int_{\Sigma}(-J_++z^{-N}I_+)\frac{\dot f^{(t)}}{f^{(t)}}\frac{\mathrm{d}z}{\mathrm{i}2\pi}+\int_{\Sigma''}(J_-+z^NI_-)\frac{\dot f^{(t)}}{f^{(t)}}\frac{\mathrm{d}z}{\mathrm{i}2\pi},\tag{5.21}$$

where $\Sigma=\Sigma_{\rm out}\cup\Sigma_{\rm in}$, $\Sigma''=\Sigma''_{\rm out}\cup\Sigma''_{\rm in}$ We first consider the contribution from $\Sigma_{\rm in}$ and $\Sigma''_{\rm in}$. From [25, equations (5.73), (5.77)],

$$\int_{\Sigma''_{\rm in}}(J_-+z^NI_-)\frac{\dot f^{(t)}}{f^{(t)}}\frac{\mathrm{d}z}{\mathrm{i}2\pi}=\int_{\mathscr{C}''_1}J\frac{\dot f^{(t)}}{f^{(t)}}\frac{\mathrm{d}z}{\mathrm{i}2\pi}=\int_{\mathscr{C}''_1}g'\frac{\dot f^{(t)}}{f^{(t)}}\frac{\mathrm{d}z}{\mathrm{i}2\pi}+\int_{\mathscr{C}''_1}(R'_{11}R_{22}-R_{12}R'_{21})_-\frac{\dot f^{(t)}}{f^{(t)}}\frac{\mathrm{d}z}{\mathrm{i}2\pi}.\tag{5.22}$$

In the same way we have

$$\int_{\Sigma_{\rm in}}(-J_++z^{-N}I_+)\frac{\dot f^{(t)}}{f^{(t)}}\frac{\mathrm{d}z}{\mathrm{i}2\pi}=\int_{\mathscr{C}_1}g'\frac{\dot f^{(t)}}{f^{(t)}}\frac{\mathrm{d}z}{\mathrm{i}2\pi}-\int_{\mathscr{C}_1}(R'_{11}R'_{22}-R'_{12}R'_{21})_-\frac{\dot f^{(t)}}{f^{(t)}}\frac{\mathrm{d}z}{\mathrm{i}2\pi}.\tag{5.23}$$

Concerning the contribution from $\Sigma_{\rm out}$ and $\Sigma''_{\rm out}$ in (5.21), we first note that

$$\int_{\Sigma''_{\rm out}}z^{-N}I_+\frac{\dot f^{(t)}}{f^{(t)}}\frac{\mathrm{d}z}{\mathrm{i}2\pi}=\int_{\Sigma''_{\rm out}}z^{-N}\frac{e^{-2g}}{f^{(t)}}(R'_{22}R_{12}-R_{22}R'_{12})\frac{\dot f^{(t)}}{f^{(t)}}\frac{\mathrm{d}z}{\mathrm{i}2\pi},\tag{5.24}$$

$$\int_{\Sigma_{\rm out}}z^{N}I_-\frac{\dot f^{(t)}}{f^{(t)}}\frac{\mathrm{d}z}{\mathrm{i}2\pi}=\int_{\Sigma''_{\rm out}}z^{N}\frac{e^{2g}}{f^{(t)}}(R_{11}R'_{21}-R'_{11}R_{11})\frac{\dot f^{(t)}}{f^{(t)}}\frac{\mathrm{d}z}{\mathrm{i}2\pi}.\tag{5.25}$$

Finally, by a contour deformation,

$$\begin{aligned}&\int_{\Sigma_{\rm out}}(-J_+)\frac{\dot f^{(t)}}{f^{(t)}}\frac{\mathrm{d}z}{\mathrm{i}2\pi}+\int_{\Sigma''_{\rm out}}(J_-)\frac{\dot f^{(t)}}{f^{(t)}}\frac{\mathrm{d}z}{\mathrm{i}2\pi}\\ =&\int_{\Sigma'_{\rm out}}(J_+-J_-)\frac{\dot f^{(t)}}{f^{(t)}}\frac{\mathrm{d}z}{\mathrm{i}2\pi}-\int_{\mathscr{C}_2}J_+\frac{\dot f^{(t)}}{f^{(t)}}\frac{\mathrm{d}z}{\mathrm{i}2\pi}+\int_{\mathscr{C}''_2}J_+\frac{\dot f^{(t)}}{f^{(t)}}\frac{\mathrm{d}z}{\mathrm{i}2\pi}\\ =&\int_{\Sigma'_{\rm out}}(g'_++g'_-)\frac{\dot f^{(t)}}{f^{(t)}}\frac{\mathrm{d}z}{\mathrm{i}2\pi}-\int_{\mathscr{C}_2\cup\mathscr{C}''_2}g'\frac{\dot f^{(t)}}{f^{(t)}}\frac{\mathrm{d}z}{\mathrm{i}2\pi}+\int_{\Sigma'_{\rm out}}((R'_{11}R_{22}-R_{12}R'_{21})-(R_{11}R'_{22}-R'_{12}R_{21}))\frac{\dot f^{(t)}}{f^{(t)}}\frac{\mathrm{d}z}{\mathrm{i}2\pi}\\ &-\int_{\mathscr{C}_2}(R_{11}R'_{22}-R'_{12}R_{21})\frac{\dot f^{(t)}}{f^{(t)}}\frac{\mathrm{d}z}{\mathrm{i}2\pi}+\int_{\mathscr{C}'_2}(R'_{11}R_{22}-R_{12}R'_{21})\frac{\dot f^{(t)}}{f^{(t)}}\frac{\mathrm{d}z}{\mathrm{i}2\pi}\ .\end{aligned}\tag{5.26}$$

Injecting the estimates (5.22), (5.23), (5.24), (5.25), (5.26) into the integrated form of (5.21), we obtain

$$\begin{aligned}&\log \mathrm{D}_N(f^{(1)})-\log \mathrm{D}_N(f^{(0)})\\ =&N\int_0^1\int_{\mathscr{C}}\frac{\dot f^{(t)}}{f^{(t)}}\frac{\mathrm{d}z}{\mathrm{i}2\pi}\mathrm{d}t+\int_0^1\left(\int_{\mathscr{C}_1\cup\mathscr{C}''_1\cup\mathscr{C}_2\cup\mathscr{C}''_2}g'\frac{\dot f^{(t)}}{f^{(t)}}\frac{\mathrm{d}z}{\mathrm{i}2\pi}\right)\mathrm{d}t+\int_0^1\left(\int_{\Sigma'_{\rm out}}(g'_++g'_-)\frac{\dot f^{(t)}}{f^{(t)}}\frac{\mathrm{d}z}{\mathrm{i}2\pi}\right)\mathrm{d}t+\int_0^1E(t)\mathrm{d}t\\ =&NV_0+\int_0^1\left(\int_{\mathscr{C}_{1-\kappa}\cup\mathscr{C}_{1+\kappa}}g'\frac{\dot f^{(t)}}{f^{(t)}}\frac{\mathrm{d}z}{\mathrm{i}2\pi}\right)\mathrm{d}t+\int_0^1E(t)\mathrm{d}t,\end{aligned}$$

where the last equation holds by contour deformation, and we denoted $\mathscr{C}_r$ the circle entered at 0 with radius $r$. Finally, it was proved in [25, equations (5.81) to (5.93)] that the above double integral coincides with $\sum_{k=1}^{\infty}kV_kV_{-k}-\alpha\log(b_+^0(1)b_-^0(1))$. This concludes the proof. □

Let

$$X(z) = \begin{pmatrix} 1 & 0 \\ 0 & e^{V_0^{(t)}} \end{pmatrix} R(z) \begin{pmatrix} 1 & 0 \\ 0 & e^{-V_0^{(t)}} \end{pmatrix} \tag{5.27}$$

As we will see in the following corollary, this conjugacy establishes symmetry between $|z| < 1$ and $|z| > 1$. This symmetry was initially broken in (5.14). This small adjustment will be important to us to optimize error terms, as mentionned after Corollary 5.4.

The matrix $X$ satisfies the following Riemann-Hilbert problem:

1. $X$ is analytic for $z \in \mathbb{C}\backslash\Gamma$.
2. The boundary values of $R$ are related by the jump condition

$$X_+(z) = X_-(z)Q(z)$$

where the jump matrix $Q$ is given by

$$\begin{aligned} Q(z) &= \begin{pmatrix} 1 & 0 \\ 0 & e^{V_0^{(t)}} \end{pmatrix} N(z) \begin{pmatrix} 1 & 0 \\ f^{(t)}(z)^{-1}z^{-N} & 1 \end{pmatrix} N(z)^{-1} \begin{pmatrix} 1 & 0 \\ 0 & e^{-V_0^{(t)}} \end{pmatrix} && \text{if } z \in \Sigma_{\text{out}} \\ Q(z) &= \begin{pmatrix} 1 & 0 \\ 0 & e^{V_0^{(t)}} \end{pmatrix} N(z) \begin{pmatrix} 1 & 0 \\ f^{(t)}(z)^{-1}z^{N} & 1 \end{pmatrix} N(z)^{-1} \begin{pmatrix} 1 & 0 \\ 0 & e^{-V_0^{(t)}} \end{pmatrix} && \text{if } z \in \Sigma''_{\text{out}} \\ Q(z) &= \begin{pmatrix} 1 & 0 \\ 0 & e^{V_0^{(t)}} \end{pmatrix} M(z) \begin{pmatrix} 1 & 0 \\ 0 & e^{-V_0^{(t)}} \end{pmatrix}, && \text{if } z \in \partial \mathrm{U} \end{aligned}$$

3. $X(z) = \mathrm{Id} + \mathrm{O}(1/z)$ as $z \to \infty$.

Proposition 5.3 can be written in terms of $X$ as follows.

**Corollary 5.4.** *The identity (5.20) holds with the error term expressed as*

$$\begin{aligned} E(t) = &- \int_{\mathscr{C}_1} (X_{11}X'_{22} - X'_{12}X_{21})_- \frac{\dot{f}^{(t)}}{f^{(t)}} \frac{\mathrm{d}z}{\mathrm{i}2\pi} + \int_{\mathscr{C}''_1} (X'_{11}X_{22} - X_{12}X'_{21})_- \frac{\dot{f}^{(t)}}{f^{(t)}} \frac{\mathrm{d}z}{\mathrm{i}2\pi} \\ &- \int_{\mathscr{C}_2} (X_{11}X'_{22} - X'_{12}X_{21})_+ \frac{\dot{f}^{(t)}}{f^{(t)}} \frac{\mathrm{d}z}{\mathrm{i}2\pi} + \int_{\mathscr{C}''_2} (X'_{11}X_{22} - X_{12}X'_{21})_+ \frac{\dot{f}^{(t)}}{f^{(t)}} \frac{\mathrm{d}z}{\mathrm{i}2\pi} \\ &+ \int_{\Sigma_{\text{out}}} z^{-N} \frac{e^{-2g+V_0^{(t)}}}{f^{(t)}} (X'_{22}X_{12} - X'_{12}X_{22})_+ \frac{\dot{f}^{(t)}}{f^{(t)}} \frac{\mathrm{d}z}{\mathrm{i}2\pi} + \int_{\Sigma''_{\text{out}}} z^{N} \frac{e^{2g-V_0^{(t)}}}{f^{(t)}} (X'_{21}X_{11} - X'_{11}X_{21})_- \frac{\dot{f}^{(t)}}{f^{(t)}} \frac{\mathrm{d}z}{\mathrm{i}2\pi} \\ &+ \int_{\Sigma'_{\text{out}}} ((X'_{11}X_{22} - X_{12}X'_{21}) - (X_{11}X'_{22} - X'_{12}X_{21})) \frac{\dot{f}^{(t)}}{f^{(t)}} \frac{\mathrm{d}z}{\mathrm{i}2\pi} . \end{aligned}$$

One advantage of this writing of $E$ is on $\Sigma_{\text{out}}$ and $\Sigma''_{\text{out}}$: the terms $e^{-2g+V_0^{(t)}}/f^{(t)}$ and $e^{2g-V_0^{(t)}}/f^{(t)}$ have smaller order than the corresponding terms in Proposition 5.3. As we will see in the next subsection, the conjugacy (5.27) also gives better bounds on the jump matrix $Q - \mathrm{Id}$ than on $M - \mathrm{Id}$.

Before performing these estimates, we will use the following rewriting of $|e^{-2g+V_0^{(t)}}/f^{(t)}|$ (when $|z| > 1$) and $|e^{2g-V_0^{(t)}}/f^{(t)}|$ (when $|z| < 1$).

**Lemma 5.5.** *We have*

$$\left| \frac{e^{-2g(z)+V_0^{(t)}}}{f^{(t)}(z)} \right| = \left| \frac{b_-^{(t)}(z)}{b_-^{(t)}(\bar{z}^{-1})} z^{-2\alpha} \right|, \text{ for } 1 < |z| < 1 + \kappa,$$

$$\left| \frac{e^{2g(z)-V_0^{(t)}}}{f^{(t)}(z)} \right| = \left| \frac{b_-^{(t)}(\bar{z}^{-1})}{b_-^{(t)}(z)} \right| \text{ for } 1 - \kappa < |z| < 1.$$

*Proof.* This is an elementary combination of equations (5.4), (5.8), (5.9), (5.11), (5.12), (5.14). □

*5.3 Asymptotic analysis of the Riemann-Hilbert problem.* The proof of Proposition 5.1 relies on bounding the error estimate in Corollary 5.4. For this, we will show that the matrix $X(z)$ is close to the constant Id, by first proving that the jump matrix $Q(z)$ is approximately Id. Before that, we need to prove that terms appearing in the previous Lemma 5.5 are close to 1.

**Lemma 5.6.** *There is a $C > 0$ such that for any $||z| - 1| < \kappa$ we have*

$$C^{-1} < \left| \frac{e^{-2g(z)+V_0^{(t)}}}{f^{(t)}(z)} \right| < C \text{ if } |z| > 1,$$

$$C^{-1} < \left| \frac{e^{2g(z)-V_0^{(t)}}}{f^{(t)}(z)} \right| < C \text{ if } |z| < 1.$$

*Proof.* Assume that $|z| > 1$. From Lemma 5.5, we need to estimate $\left| \frac{b_-^{(t)}(z)}{b_-^{(t)}(\bar{z}^{-1})} z^{-2\alpha} \right|$. Clearly, $\log |z^{-2\alpha}| < \alpha N^{-1+\delta} = \mathrm{o}(1)$. Moreover, let $\mathscr{C}'$ be the circle centered at 0 with radius $r = 1 + \kappa$. Then by analyticity we can write

$$\log b_-^{(t)}(z) - \log b_-^{(t)}(\bar{z}^{-1}) = \frac{1}{2\pi \mathrm{i}} \int_{\mathscr{C}'} \left( \frac{V^{(t)}(s)}{s - z} - \frac{V^{(t)}(s)}{s - \bar{z}^{-1}} \right) \mathrm{d}s. \tag{5.28}$$

We denote $s = re^{\mathrm{i}\theta}$. Note that

$$\int_{\mathscr{C}'} |V^{(t)}(re^{\mathrm{i}\theta}) - V^{(t)}(e^{\mathrm{i}\theta})| \frac{|z - \bar{z}^{-1}|}{|z - e^{\mathrm{i}\theta}||\bar{z}^{-1} - e^{\mathrm{i}\theta}|} \mathrm{d}\theta \le C\kappa \sup_{||s|-1|<\kappa} |{V^{(t)}}'(s)|. \tag{5.29}$$

Moreover, using (5.3) and (5.7), we have

$$\sup_{||s|-1|<\kappa} |{V^{(t)}}'(s)| \le CN^{1-\delta}. \tag{5.30}$$

Using (5.28), (5.29), and (5.30), we proved

$$\log b_-^{(t)}(z) - \log b_-^{(t)}(\bar{z}^{-1}) = \frac{1}{2\pi \mathrm{i}} \int_{\mathscr{C}'} \left( \frac{V^{(t)}(s/r)}{s - z} - \frac{V^{(t)}(s/r)}{s - \bar{z}^{-1}} \right) \mathrm{d}s + \mathrm{O}(1).$$

Define $z_0 = z/|z|$. Then from the above equation we also have

$$\log b_-^{(t)}(z) - \log b_-^{(t)}(\bar{z}^{-1}) = \frac{1}{2\pi \mathrm{i}} \int_{\mathscr{C}'} \left( \frac{V^{(t)}(s/r) - V^{(t)}(z_0)}{s - z} - \frac{V^{(t)}(s/r) - V^{(t)}(z_0)}{s - \bar{z}^{-1}} \right) \mathrm{d}s + \mathrm{O}(1),$$

so that the result will be proved if

$$\int_{\mathscr{C}'} |V^{(t)}(s/r) - V^{(t)}(z_0)| \frac{|z - \bar{z}^{-1}|}{|s - z|^2} \mathrm{d}s = \mathrm{O}(1). \tag{5.31}$$

Note that from (5.3) we have

$$e^{V^{(t)}(s/r) - V^{(t)}(z_0)} = \frac{1 - t + te^{V(s/r)}}{1 - t + te^{V(z_0)}},$$

so that

$$e^{V^{(t)}(s/r) - V^{(t)}(z_0)} \le C + Ce^{V(s/r) - V(z_0)}.$$

Here we used that $(a + b)/(c + d) < \max(a/c, b/d)$ for positive numbers, and that on the unit circle $V$ has real values. This implies that

$$|V^{(t)}(s/r) - V^{(t)}(z_0)| < 1 + |V(s/r) - V(z_0)|.$$

Moreover, Abel summation gives

$$V_0(z) = -\lambda \log \max(|1-z|, \kappa) + \mathrm{O}(1) \quad \text{uniformly in } ||z|-1| < \kappa. \tag{5.32}$$

Both previous equations give (let $z_0 = e^{\mathrm{i}\theta_0}$)

$$\int_{\mathscr{C}'} |V^{(t)}(s/r) - V^{(t)}(z_0)| \frac{|z-\bar{z}^{-1}|}{|s-z|^2} \mathrm{d}s \leq C \int_{\mathscr{C}'} (1 + |\log\max(\theta,\kappa) - \log\max(\theta_0,\kappa)|) \frac{\kappa}{|e^{\mathrm{i}\theta}-z|^2} \mathrm{d}\theta + \mathrm{O}(1) = \mathrm{O}(1).$$

This proves (5.31) and therefore the lemma for $|z| > 1$. The proof for $|z| < 1$ is similar. □

**Lemma 5.7.** *Entries of the jump matrix satisfy the bounds*

$$Q(z) - \mathrm{Id} = \mathrm{O}\left(N^{-\delta}\right) \quad \text{uniformly in } z \in \partial \mathrm{U},$$
$$Q(z) - \mathrm{Id} = \mathrm{O}\left(e^{-cN^{\delta}}\right) \quad \text{uniformly in } z \in \Sigma_{\text{out}} \cup \Sigma''_{\text{out}}.$$

*Proof.* We first consider the jump matrix on $\partial U$. It was proved in [47, Proposition 3] that the jump matrix of the $R$-Riemann-Hilbert problem satisfies the asymptotic expansion (notice that we changed the notations from [47] to our setting), in the case $\delta = 1$, i.e. when $V$ and the contour do not depend on $N$: if $|z| > 1$,

$$M(z) = \mathrm{Id} + \sum_{k=1}^{\infty} \frac{\mathrm{i}^k}{2^{k+1}\zeta^k} \begin{pmatrix} s_{\alpha,k} & (-1)^k \left(\frac{e^{-2g(z)}}{f^{(t)}(z)}\right)^{-1} t_{\alpha,k} \\ \frac{e^{-2g(z)}}{f^{(t)}(z)} t_{\alpha,k} & (-1)^k s_{\alpha,k} \end{pmatrix}$$

meaning that the remainder associated with partial sums does not exceed the first neglected term in absolute value. Here, the coefficients are $s_{\beta,k} = (\alpha + \frac{1}{2}, k) + (\alpha - \frac{1}{2}, k)$, $t_{\beta,k} = (\alpha + \frac{1}{2}, k) - (\alpha - \frac{1}{2}, k)$ where $(\nu, k) = \frac{(4\nu^2-1)(4\nu^2-9)\dots(4\nu^2-(2k-1)^2)}{2^{2k}k!}$. The above expansion is equivalent (by simple conjugacy) to

$$Q(z) = \mathrm{Id} + \sum_{k=1}^{\infty} \frac{\mathrm{i}^k}{2^{k+1}\zeta^k} \begin{pmatrix} s_{\alpha,k} & (-1)^k \left(\frac{e^{-2g+V_0^{(t)}}}{f^{(t)}}\right)^{-1} t_{\alpha,k} \\ \frac{e^{-2g+V_0^{(t)}}}{f^{(t)}} t_{\alpha,k} & (-1)^k s_{\alpha,k} \end{pmatrix}. \tag{5.33}$$

If we assume $|z-1| \asymp \kappa$ (hence $|\zeta| \asymp N^{\delta}$), and that $V$ depends on $N$ (through (5.2) and (5.3)), this expansion still holds: the proof in [47] only requires (1) known asymptotics of confluent hyperbolic fuctions (these still hold in our context as $\alpha$ is a fixed parameter independent of $N$, like in [47]) and (2) that the above coefficient $\frac{e^{-2g+V_0^{(t)}}}{f^{(t)}}$ is of order one (this property holds as proved in Lemma 5.6).

Thus (5.33) holds in our regime of interest and at first order, this approximation is $Q(z) = \mathrm{Id} + \mathrm{O}\left(N^{-\delta}\right)$, as expected. This approximation still holds when $z \in \partial\mathrm{U}$ but $|z| < 1$, by just changing the coefficient $e^{-2g+V_0^{(t)}}/f^{(t)}$ by $e^{2g-V_0^{(t)}}/f^{(t)}$ in the reasoning.

On $\Sigma_{\text{out}}$, the result follows from

$$Q(z) - \mathrm{Id} = \begin{pmatrix} 0 & 0 \\ \frac{e^{-2g+V_0^{(t)}}}{f^{(t)}} z^{-N} & 0 \end{pmatrix} = \mathrm{O}(z^{-N}) = \mathrm{O}(e^{-cN^{\delta}}). \tag{5.34}$$

A similar calculation gives the estimate on $\Sigma''_{\text{out}}$. □

**Proposition 5.8.** *The matrix $X$ satisfies the bounds*

$$X(z)_{\pm} - \mathrm{Id} = \mathrm{O}\left(\frac{1}{N|z-1|}\right) \quad \text{uniformly in } z \in \Gamma. \tag{5.35}$$

*The matrix $X' = \frac{\partial}{\partial z} X$ satisfies the bounds*

$$X'(z)_{\pm} = \mathrm{O}\left(\frac{N^{-\delta}}{|z-1|}\right) \quad \text{uniformly in } z \in \Gamma. \tag{5.36}$$

*The same bounds hold for $X(z)$ and $X'(z)$ away from $\Gamma$ uniformly in $||z|-1| < \kappa$.*

*Proof.* Consider the Cauchy operator

$$\mathrm{C}f(z)=\frac{1}{2\pi\mathrm{i}}\int_{\Gamma}\frac{f(\xi)}{\xi-z}\mathrm{d}\xi,\ \ z\in\mathbb{C}\backslash\Gamma,$$

where the orientation of inegration is clockwise on $\Sigma_{\rm out}$ and $\Sigma''_{\rm out}$ and clockwise on U. For $\xi\in\Gamma$, let $\mathrm{C}_f(\xi)=\lim_{z\to\xi^-}\mathrm{C}f(z)$. It is well known (see e.g. [49]) that this limit exists in $\mathrm{L}^2(\Gamma)$ and

$$\|C_-f\|_{\mathrm{L}^2(\Gamma)}\leq c_\Gamma\|f\|_{\mathrm{L}^2(\Gamma)}. \tag{5.37}$$

For our contour, the constant $c_\Gamma$ can actually be chosen uniform in $N$ as shown by a simple scaling argument (the radii of our circles do not matter).

Let $\Delta_X(z)=Q(z)-\mathrm{Id}$. From Lemma 5.7, $\|\Delta_X\|_{L^\infty(\Gamma)}=\mathrm{O}(N^{-\delta}+e^{-cN^\delta})$ converges to 0. Together with (5.37) it implies that

$$\mathrm{C}_{\Delta_X}:f\mapsto C_-(f\Delta_X)$$

is a bounded operator from $\mathrm{L}^2(\Gamma)$ to itself with norm $\|\mathrm{C}_{\Delta_X}\|\to0$. Hence $1-\mathrm{C}_{\Delta_X}$ has an inverse for large enough $N$, and we have

$$\|(1-\mathrm{C}_{\Delta_X})^{-1}f\|_{\mathrm{L}^2(\Gamma)}\leq C\|f\|_{\mathrm{L}^2(\Gamma)}, \tag{5.38}$$

still for large enough $N$. Let

$$\mu_X=(1-\mathrm{C}_{\Delta_X})^{-1}(\mathrm{C}_-\Delta_X)\in\mathrm{L}^2(\Gamma).$$

It is well known (see e.g. [26, Theorem 7.8]) that for any $z\notin\Gamma$

$$X(z)=\mathrm{Id}+\mathrm{C}(\Delta_X+\mu_X\Delta_X)(z). \tag{5.39}$$

We first bounds similar to (5.35) and (5.36) under a stronger assumption on $z$, namely $\mathrm{dist}(z,\Gamma)=N^{-1+\delta}/100$. To bound the first term in (5.39), note that

$$|\mathrm{C}\Delta_X(z)|\lesssim\int_{\partial\mathrm{U}}\frac{|\Delta_X(\xi)|}{|z-\xi|}|\mathrm{d}\xi|+\int_{\Sigma_{\rm out}\cup\Sigma''_{\rm out}}\frac{|\Delta_X(\xi)|}{|z-\xi|}|\mathrm{d}\xi|\lesssim\frac{1}{N|z-1|}+e^{-cN^\delta}. \tag{5.40}$$

where we used Lemma 5.7 to bound the jump matrix $\Delta_X$. Concerning the other term from (5.39), the Schwarz inequality yields

$$|\mathrm{C}(\mu_X\Delta_X)(z)|\lesssim\|\mu_X\|_{\mathrm{L}^2(\partial\mathrm{U})}\left(\int_{\partial\mathrm{U}}\frac{|\Delta_X(\xi)|^2}{|z-\xi|^2}|\mathrm{d}\xi|\right)^{1/2}+\|\mu_X\|_{\mathrm{L}^2(\Sigma_{\rm out}\cup\Sigma''_{\rm out})}\left(\int_{\Sigma_{\rm out}\cup\Sigma''_{\rm out}}\frac{|\Delta_X(\xi)|^2}{|z-\xi|^2}|\mathrm{d}\xi|\right)^{1/2}. \tag{5.41}$$

From (5.38), we have

$$\begin{aligned}\|\mu_X\|_{\mathrm{L}^2(\partial\mathrm{U})}&\lesssim\|C_-\Delta_X\|_{\mathrm{L}^2(\partial\mathrm{U})}\lesssim\|C_-(\Delta_X\mathbb{1}_{\partial\mathrm{U}})\|_{\mathrm{L}^2(\partial\mathrm{U})}+\|C_-(\Delta_X\mathbb{1}_{\Sigma_{\rm out}\cup\Sigma''_{\rm out}})\|_{\mathrm{L}^2(\partial\mathrm{U})}\\&\lesssim\|\Delta_X\|_{\mathrm{L}^2(\partial\mathrm{U})}+\|\Delta_X\|_{\mathrm{L}^\infty(\Sigma_{\rm out}\cup\Sigma''_{\rm out})}\left(\int_{\partial\mathrm{U}}|\log\mathrm{dist}(\xi,\Sigma_{\rm out}\cup\Sigma''_{\rm out})|^2|\mathrm{d}\xi|\right)^{1/2}\\&\lesssim N^{-\delta}N^{-(1-\delta)/2}+e^{-cN^\delta}\end{aligned} \tag{5.42}$$

where we used (5.37) for the third inequality. In the same manner we have

$$\|\mu_X\|_{\mathrm{L}^2(\Sigma_{\rm out}\cup\Sigma''_{\rm out})}\lesssim e^{-N^\delta}+N^{-\delta}N^{-(1-\delta)/2}\log N. \tag{5.43}$$

Equations (5.41), (5.42) and (5.43) imply that

$$|\mathrm{C}(\mu_X\Delta_X)(z)|\lesssim\frac{N^{-\delta}}{N|z-1|}+e^{-cN^\delta}. \tag{5.44}$$

Equations (5.40) and (5.44) conclude the proof of (5.35) when $z$ is far enough from $\Gamma$. As $\mathrm{dist}(z,\Gamma)=N^{-1+\delta}/100$, the estimate (5.36) follows from (5.35) by Cauchy's integral formula.

Remarkably, these estimates also hold up to $z\in\Gamma_\pm$ thanks to the classical contour deformation argument as explained in the proof of [26, Corollary 7.77]. This concludes the proof. □

**Corollary 5.9.** *We have*

$$\int_0^1 E(t)\mathrm{d}t = \mathrm{O}\left(N^{-\delta}(\log N)^3\right).$$

*Proof.* By Lemma 5.6, $|e^{-2g+V_0^{(t)}}/f^{(t)}| = |e^{2g-V_0^{(t)}}/f^{(t)}| = \mathrm{O}(1)$ in Corollary 5.4. Moreover, $\dot f^{(t)}/f^{(t)}$ has a constant sign for fixed $z$ and $\int_0^1 \dot f^{(t)}/f^{(t)}\mathrm{d}t = V(z)$. These observations injected in Corollary 5.4 give

$$\left|\int_0^1 E(t)\mathrm{d}t\right| \lesssim \sup_{0\le t\le 1}\left(\|X\|_{\mathrm{L}^\infty(\partial\mathrm{U})}\|X'\|_{\mathrm{L}^\infty(\partial\mathrm{U})}N^{-1+\delta} + \int_{\Sigma_{\mathrm{out}}\cup\Sigma'_{\mathrm{out}}\cup\Sigma''_{\mathrm{out}}} |X(z)||X'(z)||\mathrm{d}z|\right)\sup_{z\in\Gamma}\int_0^1\left|\frac{\dot f^{(t)}}{f^{(t)}}\right|$$
$$\lesssim N^{-\delta}(\log N)\sup_{z\in\Gamma}\int_0^1\left|\frac{\dot f^{(t)}}{f^{(t)}}\right|,$$

where we used Proposition 5.8. Moreover, from (5.30) and (5.32)

$$V(z) = -\lambda\log|1-z| + \mathrm{O}(1) \tag{5.45}$$

uniformly in $\Gamma$, so that if $t\le e^{-(\log N)^2}$ we have $|\dot f^{(t)}/f^{(t)}|\le CN^\lambda$. If $t > e^{-(\log N)^2}$ we use (5.7) to conclude that

$$\left|\frac{\dot f^{(t)}}{f^{(t)}}\right| \le C\left|\frac{1+e^V}{te^V}\right| \le \frac{C}{t}.$$

All together, we obtain that for any $||z|-1| < \varepsilon N^{-1+\delta}$ we have

$$\int_0^1\left|\frac{\dot f^{(t)}}{f^{(t)}}(z)\right|\mathrm{d}t \le \int_0^{e^{-(\log N)^2}} N^\delta + C\int_{e^{-(\log N)^2}}^1 \frac{\mathrm{d}t}{t} \le C(\log N)^2,$$

which concludes the proof. ☐

*5.4 Proof of Proposition 5.1.* We choose $\lambda = \alpha$. Note that

$$V_0 = 0,$$
$$\sum_{k\ge 1} kV_kV_{-k} = \alpha^2(1-\delta)\log N + \mathrm{O}(1),$$
$$\log(b_+^1(1)b_-^1(1)) = 2\alpha(1-\delta)\log N + \mathrm{O}(1).$$

Therefore Corollary 5.4 and Corollary 5.9 yield

$$\log D_N(f^{(1)}) - \log D_N(f^{(0)}) = -\alpha^2(1-\delta)\log N + \mathrm{O}(1).$$

Together with the following Lemma 5.10, this gives

$$\log D_N(1) = \alpha^2\delta\log N + \mathrm{O}(1),$$

which concludes the proof.

**Lemma 5.10.** *The Toeplitz determinant assocated to the pure Fisher-Hartwig symbol satisfies the estimates*

$$\log D_N(f^{(0)}) = \alpha^2\log N + \mathrm{O}(1).$$

*Proof.* This is an elementary consequence of Selberg's integral. The following exact formula holds [44]:

$$\mathrm{D}_N(f^{(0)}) = \prod_{k=1}^N \frac{\Gamma(k)\Gamma(k+2\alpha)}{\Gamma(k+\alpha)^2}.$$

Let $G$ be the Barnes function, which satisfies in particular $G(z+1) = \Gamma(z)G(z)$, $G(1) = 1$. Then

$$\log D_N(f^{(0)}) = \log G(1+N+2\alpha)+\log G(1+N)-2\log G(1+N+\alpha)-\log G(1+2\alpha)+2\log G(1+\alpha). \quad (5.46)$$

The Barnes function satisfies the asymptotic expansion [51]

$$\log G(1+z) = \frac{z^2}{2}\log z - \frac{3}{4}z^2 + \frac{1}{2}\log(2\pi)z - \frac{1}{12}\log z + \zeta'(-1) + \mathrm{O}\left(\frac{1}{z^2}\right). \quad (5.47)$$

From (5.46) and (5.47) we get the result of the lemma. □

Note that Lemma 5.10 also proves Lemma 4.1.

*5.5 Gaussian approximation of the increments.* We now prove Lemmas 2.1, 2.5 and 3.4, which will all follow from the following Proposition 5.11.

Let $m \in \mathbb{N}$ be fixed, let $\boldsymbol{h} = (h_1, \dots, h_m)$ (for any $1 \leq k \leq m$, $h_1, \dots, h_m \in [0, 2\pi]$), and let $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_m)$, $\boldsymbol{\beta} = (\beta_1, \dots, \beta_m)$ (for any $1 \leq k \leq m$, $0 < \alpha_k < \beta_k < 1$) have real entries. Let $\boldsymbol{\xi} = (\xi_1, \dots, \xi_m)$ have possibly complex entries. In this subsection, we change the definition of $V$ and are interested in functions of type

$$V(z) = \frac{1}{2}\sum_{k=1}^{m}\xi_k \sum_{j=N^{\alpha_k}}^{N^{\beta_k}} \frac{z^j e^{-\mathrm{i}jh_k} + z^{-j}e^{\mathrm{i}jh_k}}{j}, \quad (5.48)$$

so that $\sum_{\ell=1}^N V(e^{\mathrm{i}\theta_\ell}) = \sum_{k=1}^m \xi_k \sum_{j=N^{\alpha_k}}^{N^{\beta_k}} \mathrm{Re}\left(\frac{e^{-\mathrm{i}jh_k}\mathrm{Tr}(U_N^j)}{j}\right)$. We will also consider in linear combinations of such functions.

For a general smooth function on the unit circle we define

$$\sigma^2(V) = \sum_{j=1}^{\infty} jV_jV_{-j}. \quad (5.49)$$

We have the following estimates, for general $V$, not necessarily of type (5.48).

**Proposition 5.11** (Fourier transform asymptotics)**.** *Let $0 < \delta < 1$ be fixed. Assume that $V$ is analytic in $||z|-1| < N^{-1+\delta}$ and $|V(z)| < N^{\delta-\varepsilon}$ in this domain, for some fixed $\varepsilon > 0$. Then*

$$\mathbb{E}\left(e^{\sum_{\ell=1}^N V(e^{\mathrm{i}\theta_\ell})}\right) = e^{\sigma^2(V)}\left(1+\mathrm{O}(e^{-cN^\delta})\right). \quad (5.50)$$

*Proof.* We follow the method of Subsections 5.1, 5.2, 5.3 and 5.4 with a notable difference: there is no singularity at 1, that is $\alpha = 0$. The contour $\Gamma$ of the $X$-Riemann-Hilbert problem is just $\Sigma_{\mathrm{out}} \cup \Sigma''_{\mathrm{out}}$ (their closed version, i.e. two full circles).

We cannot apply the previous method directly. Our new choice for $V$ has much greater amplitude than (5.2), so if we adopt the interpolation (5.3), there is no not guarantee that $1-t+te^{V(z)} \neq 0$ for all $0 \leq t \leq 1$ and $||z|-1| < N^{-1+\delta}$. Lemma 5.2 does not hold anymore and we cannot consider any Riemann-Hilbert problem.

To circumvent this problem, we adopt a different interpolation in many steps, following an idea from [25, Section 5.4]. Define $N$ functions $(V^{(k)})_{1\leq k\leq N}$ (any number on a polynomial scale greater than $N^\delta$ would actually work) simply defined as $V^{(k)}(z) = \frac{k}{N}V(z)$. For fixed $k$, we consider the interpolation

$$e^{V^{(k,t)}(z)} = (1-t)e^{V^{(k-1)}(z)} + te^{V^{(k)}(z)}.$$

Then the analogue of Lemma 5.2 holds: for any $k$ and $t$, $V^{(k,t)}$ admits an analytic continuation to $||z|-1| < \kappa$, because, on that domain, we have $|\mathrm{Im}V^{(k-1)}(z) - \mathrm{Im}V^{(k)}(z)| \leq \frac{1}{N}\sup_{||z|-1|<\kappa}|V(z)| \leq N^{-1+\delta} < 1/10$. We therefore can consider the following Riemann-Hilbert problem for the matrix $X$ (we do not mention the dependence of $X$ in $t$ and $k$):

1. $X$ is analytic for $z \in \mathbb{C}\backslash\Gamma$.
2. The boundary values of $R$ are related by the jump condition

$$X_+(z) = X_-(z)Q(z)$$

where the jump matrix $Q$ is given by

$$Q(z) = \begin{pmatrix} 1 & 0 \\ 0 & e^{V_0^{(k,t)}} \end{pmatrix} N(z) \begin{pmatrix} 1 & 0 \\ f^{(k,t)}(z)^{-1}z^{-N} & 1 \end{pmatrix} N(z)^{-1} \begin{pmatrix} 1 & 0 \\ 0 & e^{-V_0^{(k,t)}} \end{pmatrix} \quad \text{if } z \in \Sigma_{\text{out}},$$
$$Q(z) = \begin{pmatrix} 1 & 0 \\ 0 & e^{V_0^{(k,t)}} \end{pmatrix} N(z) \begin{pmatrix} 1 & 0 \\ f^{(k,t)}(z)^{-1}z^{N} & 1 \end{pmatrix} N(z)^{-1} \begin{pmatrix} 1 & 0 \\ 0 & e^{-V_0^{(k,t)}} \end{pmatrix} \quad \text{if } z \in \Sigma''_{\text{out}},$$

where we used the notations

$$f^{(k,t)}(z) = e^{V^{(k,t)}(z)}$$
$$N(z) = \begin{cases} e^{g(z)\sigma_3} & \text{if } |z| > 1 \\ e^{g(z)\sigma_3}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} & \text{if } |z| < 1 \end{cases}$$
$$\exp(g(z)) = \exp \frac{1}{2\pi \mathrm{i}} \int_{\mathscr{C}} \frac{\ln V^{(k,t)}(s)}{s-z} \mathrm{d}s = \begin{cases} e^{V_0^{(k,t)}} b_+^{(k,t)}(z) & \text{if } |z| < 1, \\ b_-^{(k,t)}(z)^{-1} & \text{if } |z| > 1. \end{cases}$$

3. $X(z) = \mathrm{Id} + \mathrm{O}(1/z)$ as $z \to \infty$.

Define

$$\mathrm{D}_N(V^{(k)}) = \mathbb{E}\left(\prod_{j=1}^{N} e^{V^{(k)}(e^{\mathrm{i}\theta_j})}\right) . \tag{5.51}$$

From [25, Equation (5.106)], for any $1 \leq k \leq N$ we have

$$\log \mathrm{D}_N(f^{(k)}) - \log \mathrm{D}_N(f^{(k-1)}) = V_0 + \frac{2k-1}{N^2}\sum_{k=1}^{\infty} kV_kV_{-k} + \int_0^1 E(t)\mathrm{d}t. \tag{5.52}$$

where

$$E(t) = \int_{\Sigma_{\text{out}}} z^{-N}\frac{e^{-2g+V_0^{(k,t)}}}{f^{(k,t)}}(X'_{22}X_{12} - X'_{12}X_{22})_+ \frac{\dot f^{(k,t)}}{f^{(k,t)}}\frac{\mathrm{d}z}{\mathrm{i}2\pi} + \int_{\Sigma''_{\text{out}}} z^{N}\frac{e^{2g-V_0^{(k,t)}}}{f^{(k,t)}}(X'_{21}X_{11} - X'_{11}X_{21})_- \frac{\dot f^{(k,t)}}{f^{(k,t)}}\frac{\mathrm{d}z}{\mathrm{i}2\pi}$$
$$+ \int_{\Sigma'_{\text{out}}} ((X'_{11}X_{22} - X_{12}X'_{21}) - (X_{11}X'_{22} - X'_{12}X_{21}))\frac{\dot f^{(k,t)}}{f^{(k,t)}}\frac{\mathrm{d}z}{\mathrm{i}2\pi}. \tag{5.53}$$

The formula (5.53) comes from [25, Section 5.4], in the same way as we followed [25, Section 5.3] for the proof of Proposition 5.3 and Corollary 5.4.

Note that $\sum_{k=1}^{N}(2k-1)/N^2 = 1$, so that Proposition 5.11 will be proved by summation of (5.52) if, uniformly in $k$ and $t$, we have

$$|E(t)| \leq e^{-cN^{\delta}}. \tag{5.54}$$

We now bound all terms in (5.48), uniformly in $k$ and $t$.

(a) Reproducing the reasoning in Lemma 5.6, we have, on $\Sigma_{\text{out}}$,

$$\log\left|\frac{e^{-2g+V_0^{(k,t)}}}{f^{(t)}}(z)\right| \leq C\int_{\mathscr{C}_{1+2\kappa}}\left(\frac{V^{(k,t)}(s)}{s-z} - \frac{V^{(k,t)}(s)}{s-\bar z^{-1}}\right)\mathrm{d}s \leq C\|V^{(k,t)}\|_{\mathrm{L}^\infty(\Sigma_{\text{out}})} \leq CN^{\delta-\varepsilon}.$$

(b) Clearly $|\dot f^{(k,t)}/f^{(k,t)}| < \frac{1}{N}\sup_{||z|-1|<\kappa}|V(z)| \leq 1$.

(c) The analogue of Lemma 5.7 now just involves $\Sigma_{\rm out}$ and $\Sigma''_{\rm out}$. By equation (5.34), using (a) we now have $Q(z) - \mathrm{Id} = \mathrm{O}(e^{CN^{\delta-\varepsilon}}e^{-cN^{\delta}}) \leq e^{-cN^{\delta}}$. Following the reasoning of Proposition 5.8, this implies that

$$X(z)_\pm - \mathrm{Id} = \mathrm{O}\left(e^{-cN^{\delta}}\right), X'(z)_\pm = \mathrm{O}\left(e^{-cN^{\delta}}\right)$$

uniformly in $z \in \Sigma_{\rm out} \cup \Sigma''_{\rm out}$, and the same bounds hold for $X(z)$ and $X'(z)$ uniformly in $||z|-1| < \kappa$.

The estimates (a), (b) and (c) in (5.52) imply the result (5.54) and therefore conclude the proof of the proposition. □

The Fourier decomposition of $V$ is

$$V(e^{\mathrm{i}\theta}) = \sum_{j\in\mathbb{Z}} V_j e^{\mathrm{i}j\theta},\ V_j = \frac{1}{2}\sum_{k=1}^{m}\xi_k\frac{e^{-\mathrm{i}jh_k}}{|j|}\mathbb{1}_{N^{\alpha_k}\leq|j|\leq N^{\beta_k}}.$$

*Proof of Lemma 2.1.* We apply Proposition 5.11 to

$$V(z) = \frac{1}{2}\xi \sum_{1\leq j<N^{1-\delta}} -\frac{e^{-\mathrm{i}jh}z^j + e^{\mathrm{i}jh}z^{-j}}{j}$$

We have $|V(z)| < C|\xi|\log N < N^{\delta/2}$ on $||z|-1| < N^{-1+\delta}$. Therefore (5.50) holds and the proof is concluded by noting that $2\sigma^2(V) = \xi^2\sigma^2$, where we used both notations (5.49) and (2.3). □

*Proof of Lemma 2.5.* The proof is identical to the above proof of Lemma 2.1. □

*Proof of Lemma 3.4.* We apply Proposition 5.11 to

$$V(z) = \frac{1}{2}\xi\left(\sum_{j=1}^{N^{1-\delta}} -\frac{(e^{-\mathrm{i}jh}z^j + e^{\mathrm{i}jh}z^{-j})}{j} - \sum_{j=1}^{N^{1-\delta}} -\frac{(z^j+z^{-j})}{j}\right)$$

As $|\xi h N^{1-\delta}| \leq N^{\delta-\varepsilon}$, we have $|V(z)| < N^{\delta-\frac{\varepsilon}{2}}$ for $||z|-1| < N^{-1+\delta}$. Therefore (5.50) holds and the proof is concluded by noting that $\sigma^2(V)/2 = 2\xi^2(\sigma^2-\rho)$, where we used both notations (5.49) and (3.8). □

*5.6 Tail estimate for the imaginary part.* The following is the strict analogue of Proposition 5.1, for the imaginary part of the series.

**Proposition 5.12.** *For any fixed (small) $\delta \in (0,1)$ and (large) $\alpha$, for large enough $N$ we have*

$$\mathbb{E}\left(\exp\left(2\alpha\sum_{j\geq N^{1-\delta}}\frac{\mathrm{Im}(\mathrm{Tr}\mathrm{U}_N^j)}{j}\right)\right) = \exp\left((1+\mathrm{o}(1))\alpha^2\delta\log N\right). \tag{5.55}$$

*Proof.* We follow the ideas from subsections 5.1 to 5.4. The previous definition of $V$ in (5.2) is now replaced with

$$V(z) = \lambda\sum_{j=1}^{N^{1-\delta}}\frac{z^j - z^{-j}}{\mathrm{i}j},\ \text{i.e. } V_j = \frac{\lambda}{\mathrm{i}j}\mathbb{1}_{|j|\leq N^{1-\delta}}, \tag{5.56}$$

where we will choose $\lambda = -\alpha$. Denoting $z = e^{\mathrm{i}\theta}$ with $0 \leq \theta < 2\pi$, we now have

$$f^{(t)}(z) = e^{V^{(t)}(z)}e^{-\alpha(\theta-\pi)}.$$

With the notations in [24, equation (1.2)], this coincides with the choice $\beta_0 = \mathrm{i}\alpha$, $\beta_j = 0$ for $j \geq 1$ and $\alpha_j \equiv 0$. Proposition 5.3 is now changed into

$$\log \mathrm{D}_N(f^{(1)}) - \log \mathrm{D}_N(f^{(0)}) = NV_0 + \sum_{k=1}^{\infty} kV_kV_{-k} + \mathrm{i}\alpha \log(b_+^1(1)/b_-^1(1)) + \int_0^1 \widetilde{E}(t)\mathrm{d}t, \quad (5.57)$$

as shown in [25, equation (5.93)]. The new error term $\widetilde{E}(t)$ will be defined and shown to be negligible at the end of this proof. We first estimate all other terms in (5.57). Clearly, $V_0 = 0$, $\sum_{k\geq 1} kV_kV_{-k} = \alpha^2(1-\delta)\log N + \mathrm{O}(1)$, and from (5.10) we have $\log(b_+^1(1)/b_-^1(1)) = \mathrm{i}2\alpha(1-\delta)\log N + \mathrm{O}(1)$. Moreover, from Selberg's integrals it is well known that (see [44])

$$\mathrm{D}_N(f^{(0)}) = \prod_{k=1}^{N} \frac{\Gamma(k)^2}{\Gamma(k+\mathrm{i}\alpha)\Gamma(k-\mathrm{i}\alpha)}.$$

With the asymptotic expansion (5.47), this yields $\log \mathrm{D}_N(f^{(0)}) = \alpha^2 \log N + \mathrm{O}(1)$. To summarize, we just proved that

$$\log \mathrm{D}_N(f^{(1)}) = \alpha^2(1-\delta)\log N + \mathrm{O}(1) + \int_0^1 \widetilde{E}(t)\mathrm{d}t,$$

so that the proof of Proposition 5.12 will be complete after bounding $\widetilde{E}(t)$.

This error term $\widetilde{E}(t)$ is defined exactly as in Corollary 5.4 after replacing $X$ with $\widetilde{X}$, the latter being the solution of the same Riemann-Hilbert problem as $X$ (see before Corollary 5.4) except that in the definition of the jump matrix $M$ (see (5.18)) one needs to be replace $\exp(g)$ from (5.14), $F$ from (5.15), $E$ from (5.17) and $\Psi$ from (5.16) as follows: we now choose $m = 0$, $z_0 = 1$, $\alpha_0 = 0$ and $\beta_0 = \mathrm{i}\alpha$ in the formulas [24, equations (4.9) and (4.10)] for $\mathcal{D} = \exp(g)$, [24, equation (4.17)] for $F$, [24, equations (4.47) till (4.50)] for $E$, and [24, Proposition 4.1 and equations (4.25) till (4.29)] for $\Psi$.

For the new definition of $\exp(g)$, Lemmas 5.5 and 5.6 do not hold anymore, but the following can be shown easily: for some $C = C(\delta) > 0$ we have $N^{-C} < \left|\frac{e^{-2g(z)+V_0^{(t)}}}{f^{(t)}(z)}\right| < N^C$ for $|z| > 1$, and $N^{-C} < \left|\frac{e^{2g(z)-V_0^{(t)}}}{f^{(t)}(z)}\right| < N^C$ for $|z| < 1$. As a consequence, in the definition $\widetilde{E}(t)$, the integrals along $\Sigma_{\mathrm{out}}$ and $\Sigma''_{\mathrm{out}}$ are easily shown to be negligible.

To control the other terms, we need to bound the solution of the Riemann-Hilbert problem. Lemma 5.7 still holds for our new jump matrix: this is a simple calculation from [24, equation (4.55)]. Once this estimate is known, Proposition 5.8 and Corollary 5.9 still hold and their proof is unchanged. This concludes the proof of Proposition 5.12. □

## References


[1] L. Addario-Berry and B. A. Reed, *Ballot theorems, old and new*, Bolyai Soc. Math. Stud., vol. 17, Springer, Berlin, 2008.

[2] ———, *Minima in branching random walks*, Ann. Probab. **37** (2009), no. 3, p. 1044–1079.

[3] E. Aïdékon, *Convergence in law of the minimum of a branching random walk*, Ann. Probab. **41** (2013), no. 3A, 1362–1426.

[4] L.-P. Arguin, D. Belius, and A.J. Harper, *Maxima of a randomized Riemann zeta function, and branching random walks*, Preprint arXiv:1506.00629 (2015).

[5] Louis-Pierre Arguin and Olivier Zindy, *Poisson-Dirichlet statistics for the extremes of the two-dimensional discrete Gaussian free field*, Electron. J. Probab. **20** (2015), no. 59, 19.

[6] ———, *Poisson-Dirichlet statistics for the extremes of a log-correlated Gaussian field*, Ann. Appl. Probab. **24** (2014), no. 4, 1446–1481.

[7] M. Bachmann, *Limit theorems for the minimal position in a branching random walk with independent logconcave displacements*, Adv. in Appl. Probab. **32** (2000), no. 1, 159–176.

[8] R. N. Bhattacharya and R. Ranga Rao, *Normal approximation and asymptotic expansions*, John Wiley & Sons, New York-London-Sydney, 1976. Wiley Series in Probability and Mathematical Statistics.

[9] D. Belius and N. Kistler, *The subleading order of two dimensional cover times*, Preprint arXiv:1405.0888 (2014).

[10] G. Ben Arous and P. Bourgade, *Extreme gaps between eigenvalues of random matrices*, Ann. Probab. **41** (2013), 2648–2681.

[11] M. Biskup and O. Louidor, *Extreme local extrema of two-dimensional discrete Gaussian free field*, Preprint arXiv:1306.2602 (2013).

[12] E. Bolthausen, J.-D. Deuschel, and G. Giacomin, *Entropic repulsion and the maximum of the two-dimensional harmonic crystal*, Ann. Probab. **29** (2001), no. 4, 1670–1692.

[13] Anton Bovier and Irina Kurkova, *Derrida's generalized random energy models. II. Models with continuous hierarchies*, Ann. Inst. H. Poincaré Probab. Statist. **40** (2004), no. 4, 481–495 (English, with English and French summaries).

[14] P. Bourgade, *Mesoscopic fluctuations of the zeta zeros*, Probab. Theory Related Fields **148** (2010), no. 3-4, 479–500.

[15] M. D. Bramson, *Maximal displacement of branching Brownian motion*, Comm. Pure Appl. Math. **31** (1978), no. 5, 531–581.

[16] M. Bramson, J. Ding, and O. Zeitouni, *Convergence in law of the maximum of the two-dimensional discrete Gaussian free field*, Preprint arXiv:1301.6669 (2013).

[17] ______, *Convergence in law of the maximum of nonlattice branching random walk*, Preprint arxiv: 1404.3423 (2014).

[18] M. Bramson and O. Zeitouni, *Tightness for a family of recursion equations*, Ann. Probab. **37** (2009), no. 2, 615–653.

[19] ______, *Tightness of the recentered maximum of the two-dimensional discrete Gaussian free field*, Comm. Pure Appl. Math. **65** (2012), no. 1, 1–20.

[20] C. Cacciapuoti, A. Maltsev, and B. Schlein, *Bounds for the Stieltjes transform and the density of states of Wigner matrices*, Probability Theory and Related Fields **163** (2015), no. 1, 1–59.

[21] D. Carpentier and P. Le Doussal, *Glass transition of a particle in a random potential, front selection in nonlinear renormalization group, and entropic phenomena in Liouville and sinh-Gordon models*, Phys. Rev. E **63** (2001), 026110, 33pp.

[22] T. Claeys and I. Krasovsky, *Toeplitz determinants with merging singularities*, to appear in Duke Math. Journal, Preprint arxiv:1403.3639 (2014).

[23] P. Deift, *Integrable operators*, Amer. Math. Soc. Transl. **189** (1999), no. 2, 69–84.

[24] P. Deift, A. Its, and I. Krasovsky, *Asymptotics of Toeplitz, Hankel, and Toeplitz+Hankel determinants with Fisher-Hartwig singularities*, Annals of Mathematics **174** (2011), 1243–1299.

[25] ______, *On the asymptotics of a Toeplitz determinant with singularities*, Mathematical Sciences Research Institute Publications **65** (2014), 93–146.

[26] P. Deift, T. Kriecherbauer, K. T-R McLaughlin, S. Venakides, and X. Zhou, *Strong asymptotics of orthogonal polynomials with respect to exponential weights*, Comm. Pure Appl. Math. **52** (1999), no. 12, 1491–1552.

[27] A. Dembo, Y. Peres, and O. Zeitouni, *Cover times for Brownian motion and random walks in two dimensions*, Ann. of Math. (2) **160** (2004), no. 2, 433–464.

[28] B. Derrida and H. Spohn, *Polymers on disordered trees, spin glasses, and traveling waves*, J. Statist. Phys. **51** (1988), no. 5-6, 817–840. New directions in statistical mechanics (Santa Barbara, CA, 1987).

[29] P. Diaconis and S. N. Evans, *Linear functionals of eigenvalues of random matrices*, Trans. Amer. Math. Soc. **353** (2001), no. 7, 2615–2633.

[30] P. Diaconis and M. Shahshahani, *On the eigenvalues of random matrices*, J. Appl. Probab. **31A** (1994), 49–62. Studies in applied probability.

[31] J. Ding, R. Roy, and O. Zeitouni, *Convergence of the centered maximum of log-correlated Gaussian fields*, Preprint arXiv:1503.04588 (2015).

[32] J. Ding and O. Zeitouni, *Extreme values for two-dimensional discrete Gaussian free field*, Ann. Probab. **42** (2014), no. 4, 1480–1515.

[33] L. Erdős, H.-T. Yau, and J. Yin, *Rigidity of eigenvalues of generalized Wigner matrices*, Adv. Math. **229** (2012), no. 3, 1435–1515.

[34] Y. V. Fyodorov and J.-P. Bouchaud, *Freezing and extreme-value statistics in a random energy model with logarithmically correlated potential*, J. Phys. A **41** (2008), no. 37, 372001, 12 pp.

[35] Y. V. Fyodorov, G. A. Hiary, and J. P. Keating, *Freezing Transition, Characteristic Polynomials of Random Matrices, and the Riemann Zeta Function*, Phys. Rev. Lett. **108** (2012), 170601, 5pp.

[36] Y. V. Fyodorov and J. P. Keating, *Freezing transitions and extreme values: random matrix theory, and disordered landscapes*, Phil. Trans. R. Soc. A **372** (2014), 20120503, 32 pp.

[37] Y. V. Fyodorov, B. Khoruzhenko, and N. Simm, *Fractional Brownian motion with Hurst index $H = 0$ and the Gaussian Unitary Ensemble*, to appear in Annals of Probability (2016).

[38] Y. V. Fyodorov, P. Le Doussal, and A. Rosso, *Statistical mechanics of logarithmic REM: duality, freezing and extreme value statistics of $1/f$ noises generated by Gaussian free fields*, J. Stat. Mech. Theory Exp. **10** (2009), P10005, 32 pp.

[39] ———, *Counting function fluctuations and extreme value threshold in multifractal patterns: the case study of an ideal* $1/f$ *noise*, J. Stat. Phys. **149** (2012), no. 5, 898–920.

[40] Y. V. Fyodorov and N. J. Simm, *On the distribution of maximum value of the characteristic polynomial of GUE random matrices*, preprint, arXiv:1503.07110 (2015).

[41] A. J. Harper, *A note on the maximum of the Riemann zeta function, and log-correlated random variables*, Preprint arXiv:1304.0677 (2013).

[42] H. Heine, *Kugelfunktionen*, Berlin, 1878 and 1881; reprinted, Physica Verlag, Würzburg, 1961.

[43] C. P. Hughes, J. Keating, and N. O'Connell, *On the Characteristic Polynomial of a Random Unitary Matrix*, Comm. Math. Phys. **2** (2001), 429–451.

[44] J. Keating and N. Snaith, *Random Matrix Theory and* $\zeta(1/2 + \mathrm{i}t)$, Comm. Math. Phys. **1** (2000), 57–89.

[45] N. Kistler, *Derrida's Random Energy Models*, Lecture Notes in Math., vol. 2143, Springer, Berlin, 2015.

[46] T. Madaule, *Maximum of a log-correlated Gaussian field*, Preprint arXiv:1307.1365 (2014).

[47] A. Martínez-Finkelshtein, K. T.-R. McLaughlin, and E. B. Saff, *Asymptotics of orthogonal polynomials with respect to an analytic weight with algebraic singularities on the circle*, Int. Math. Res. Not. (2006), Art. ID 91426, 43 pp.

[48] R. Rhodes and V. Vargas, *Gaussian multiplicative chaos and applications: A review*, Probab. Surveys **11** (2014), 315–392.

[49] E. M. Stein, *Singular integrals and differentiability properties of functions*, Princeton Mathematical Series, No. 30, Princeton University Press, Princeton, N.J., 1970.

[50] E. M. Stein and R. Shakarchi, *Fourier analysis*, Princeton Lectures in Analysis, vol. 1, Princeton University Press, Princeton, NJ, 2003. An introduction.

[51] A. Voros, *Spectral Functions, Special Functions and the Selberg Zeta Function*, Commun. Math. Phys. **110** (1987), 439–465.

[52] C. Webb, *The characteristic polynomial of a random unitary matrix and Gaussian multiplicative chaos - The* $\mathrm{L}^2$*-phase*, axXiv:1410.0939 (2014).