Fractional Brownian motion with Hurst index $H = 0$ and the Gaussian Unitary Ensemble

Y. V. Fyodorov*, **B. A. Khoruzhenko and N. J. Simm***

School of Mathematical Sciences, Queen Mary University of London
Mile End Road, London E1 4NS, United Kingdom

Abstract: The goal of this paper is to establish a relation between characteristic polynomials of $N \times N$ GUE random matrices $\mathcal{H}$ as $N \to \infty$, and Gaussian processes with logarithmic correlations. We introduce a regularized version of fractional Brownian motion with zero Hurst index, which is a Gaussian process with stationary increments and logarithmic increment structure. Then we prove that this process appears as a limit of $D_N(z) = -\log|\det(\mathcal{H} - zI)|$ on *mesoscopic scales* as $N \to \infty$. By employing a Fourier integral representation, we use this to prove a continuous analogue of a result by Diaconis and Shahshahani [15]. On the *macroscopic scale*, $D_N(x)$ gives rise to yet another type of Gaussian process with logarithmic correlations. We give an explicit construction of the latter in terms of a Chebyshev-Fourier random series.

1. Introduction

Suppose that $\mathcal{H}$ is a random Hermitian matrix of size $N \times N$ taken from the Gaussian Unitary Ensemble (GUE), with ensemble distribution given by the measure

$$Const. \exp\left[-2N\mathrm{Tr}(\mathcal{H}^2)\right] \prod_{j=1}^{N} d\mathcal{H}_{jj} \prod_{1\le j<k\le N} d\,\mathrm{Re}\,\mathcal{H}_{jk} d\,\mathrm{Im}\,\mathcal{H}_{jk}. \tag{1.1}$$

It is well known that in the limit of infinite matrix dimensions $N \to \infty$, the distribution of the eigenvalues of $\mathcal{H}$ is supported on the interval $[-1, 1]$ and has density $\frac{2}{\pi}\sqrt{1-x^2}$ there. This is known as Wigner's semicircle law, see e.g. [42] and [1] for precise statements. In this paper we are concerned with the *random process* in x defined by the logarithm

$$D_N(x) = -\log|\det(\mathcal{H} - xI)| \tag{1.2}$$

of the characteristic polynomial of $\mathcal{H}$ in the limit $N \to \infty$, with x varying in $(-1, 1)$. The quantity $D_N(x)$ is a particular case of linear eigenvalue statistics $X_N(f) = \sum_{k=1}^{N} f(x_k)$, where $x_1, \ldots, x_N$ are the eigenvalues of $\mathcal{H}$. It is well known that for suitably regular test functions f, $X_N(f)$ is asymptotically normal as $N \to \infty$ with variance $\sigma^2(f) = \frac{1}{4}\sum_{k=1}^{\infty} kc_k(f)^2$, where $c_k(f)$ are the *Chebyshev-Fourier coefficients*:

$$c_k(f) = \frac{2}{\pi}\int_{-1}^{1} \frac{f(u)T_k(u)}{\sqrt{1-u^2}}\,du, \qquad T_k(u) = \cos(k\arccos(u)). \tag{1.3}$$

In fact, the asymptotic normality of $X_N(f)$ for regular f has been established for a variety of random matrix ensembles, see for example [31, 38, 42] and references therein.

Since x lies in the bulk of the eigenvalue distribution, our test function, $f(u) = \log|u - x|$ is unbounded. Its Chebyshev-Fourier coefficients are proportional to $1/k$, so that $\sigma^2(f) = \infty$ and it is then natural to consider normalizing $D_N(x)$ before taking the limit $N \to \infty$. Indeed, for any fixed $x \in (-1, 1)$ the variance of $D_N(x)$ grows with N like $\frac{1}{2}\log N$, and for any finite number of distinct points $x_1, \ldots, x_m$ in $(-1, 1)$ the random vector $(D_N(x_1), \ldots, D_N(x_m))/(\frac{1}{2}\log N)^{1/2}$ converges

*Supported by the EPSRC grant EP/J002763/1 "Insights into Disordered Landscapes via Random Matrix Theory and Statistical Mechanics".

in distribution, after centering, to a collection of m independent standard Gaussians as $N \to \infty$. This can be inferred from the asymptotic identity due to Krasovsky [35]:

$$\begin{aligned} \mathbb{E}\left\{e^{-\sum_{k=1}^{m}\alpha_k D_N(x_k)}\right\} = \prod_{k=1}^{m}\left[C\left(\frac{\alpha_k}{2}\right)(1-x_k^2)^{\alpha_k^2/8}N^{\alpha_k^2/4}e^{\alpha_k N(2x_k^2-1-2\log(2))/2}\right] \\ \times \prod_{1\le\nu<\mu\le m}(2|x_\nu - x_\mu|)^{-\alpha_\nu\alpha_\mu/2}\left(1+O\left(\frac{\log N}{N}\right)\right), \end{aligned} \tag{1.4}$$

where $C(\alpha) = 2^{2\alpha^2}G(\alpha+1)^2/G(2\alpha+1)$ and $G(z)$ is the Barnes G-function. The most salient feature of the asymptotics in (1.4) is the product of differences on the second line which, when rewritten in the form

$$\exp\Big[-\sum_{1\le\nu<\mu\le m}\frac{\alpha_\nu\alpha_\mu}{2}\,\log|2(x_\nu - x_\mu)|\Big], \tag{1.5}$$

is suggestive of the existence of a logarithmic covariance structure in the Gaussian process $D_N(x)$. However, this term is of sub-leading order to the variance term. Clearly then, the normalization of the process (1.2) comes at a price, because the non-trivial covariance structure implied by (1.5) is too small to survive the limit $N \to \infty$.

This motivates the following question. How can we 'regularize' the process (1.2) so that it has a well-defined limit that 'feels' the covariance structure implied by (1.5)? Hughes, Keating and O'Connell [30] answered this question in the context of the Circular Unitary Ensemble (Haar unitary matrices). Employing convergence in functional spaces instead of point-wise convergence, they proved that the logarithm $V_N(\theta) = -2\log|p_N(\theta)|$ of the characteristic polynomial $p_N(\theta) = \det\left(I - U\,e^{-i\theta}\right)$ of Haar unitary matrices U converges as $N \to \infty$ to the stochastic process represented by the Fourier series

$$V(\theta) = \sum_{n=1}^{\infty}\frac{1}{\sqrt{n}}\left(v_n e^{in\theta} + \overline{v_n}e^{-in\theta}\right) . \tag{1.6}$$

Here, the coefficients $v_n, \overline{v}_n$ are independent standard *complex* Gaussians, $\mathbb{E}\{v_n\overline{v}_n\} = 1$, and the convergence of the series is understood in the sense of distributions in a suitable Sobolev space. This process has a logarithmic singularity in the covariance structure: $\mathbb{E}\{V(\theta_1)V(\theta_2)\} = -2\log|e^{i\theta_1} - e^{i\theta_2}|$.

At this point it is appropriate to mention that random processes and fields with logarithmic covariance structure appear with astonishing regularity in physics and also engineering applications, see e.g. [12] and more recently [26]. Those objects are intimately related to multifractal cascades emerging in turbulence, and from that angle attracted considerable mathematical interest within the last decade, see, e.g., [3] and [4]. In fact, closely related mathematical objects appear in the so-called "multiplicative chaos" construction going back to Kahane's work [32], also see [44] and references therein for recent research in that direction which was motivated, in particular, by Quantum Gravity applications. In two spatial dimensions, the most famous example of the random field of that type is the two-dimensional Gaussian Free Field [48]. A regularized version of this field appeared in a non-trivial way in the work of Rider and Virág [45], who showed that it describes the limiting law of the log-modulus of characteristic polynomials in the Ginibre ensemble. The Gaussian Free Field also appeared more recently as the limiting distribution of the eigenvalue counting function in general β-Jacobi ensembles and their principal subminors [7]. As for the one-dimensional processes with logarithmic correlations, they are known in natural sciences under the general name of *1/f noises* (see Section 2 in [26] for some general references) since, in the spectral representation, the Fourier transform of the covariance or structure function, interpreted as a "power" of the signal, is inversely proportional to the Fourier variable (i.e. the "frequency" f). The random process $V(\theta)$ is, arguably, the simplest time-periodic stationary version of $1/f$ noise. It was found to play an important role in the construction of conformally invariant planar random curves [2] and statistical mechanics of disordered systems [24]. We note in passing that from a different angle, discrete sequences with $1/f$ properties were considered heuristically in the physics literature, see e.g. [20] and [39].

The motivation for the work in [30] came from number theory, as for large N, $p_N(\theta)$ provides a good model for describing statistics of the values of the Riemann-zeta function high up the critical line [33]. The established relation of $p_N(\theta)$ to $V(\theta)$ turned out to be fruitful. It allowed one to put forward nontrivial conjectures about statistics of extreme and high values of characteristic polynomials of Haar unitary matrices emerging as $N \to \infty$, and eventually for the Riemann-zeta function [23, 25].

The main goal of this paper is to investigate further the relation between $1/f$-noises and the characteristic polynomials of random matrices in the limit $N \to \infty$. Significantly extending the picture found in [30] we will show that the limiting process depends on the *spectral scale* at which one allows the argument x of the characteristic polynomial $\det(\mathcal{H} - xI)$ to vary. To this end, let us remind the reader that, as is well known in random matrix theory, see e.g., [42], there exist three natural scales in the spectra of large random matrices. One, known as the global, or *macroscopic* scale is set for the GUE by the width of the support of the semicircle law and, in the normalization chosen in the present paper, see (1.1), remains of the order of unity as $N \to \infty$. Second, known as the local, or *microscopic* scale is set by the typical separation between neighbouring eigenvalues and is, in the chosen normalization, of order $1/N$ for large N. Finally, the third scale which is called *mesoscopic* can be defined as intermediate between those two.

Deferring precise statements to the next section, now we will outline the two instances of $1/f$ noise that emerge in the limit $N \to \infty$ for the GUE matrices. On the macroscopic scale, by adapting the arguments of [30] to our setting, we prove that, as $N \to \infty$, the process $\{D_N(x) : x \in (-1, 1)\}$ converges, after centering, to the (aperiodic) $1/f$ noise given by the random Chebyshev-Fourier series

$$F(x) = \sum_{n=1}^{\infty} \frac{1}{\sqrt{n}} a_n T_n(x), \qquad x \in (-1, 1), \tag{1.7}$$

where a_n, $n = 1, 2 \ldots$ is a sequence of independent standard real Gaussians. As with the Fourier series in (1.6), the convergence in (1.7) has to be understood in the sense of distributions in a suitable Sobolev space. The covariance structure associated with the generalized process (1.7) is given by an integral operator with kernel $\mathbb{E}\{F(x)F(y)\} = -\frac{1}{2}\log(2|x - y|)$.

The problem of finding a suitable model to describe the statistical properties of the characteristic polynomials of random matrices on the *mesoscopic* rather than macroscopic scale turned out to be much more challenging and is the main focus of the present paper. Our main finding is the emergence of fractional Brownian motion with Hurst index $H = 0$ in this context. To describe the latter, we recall that the conventional *fractional Brownian motion* (fBm) is a zero-mean Gaussian process $B_H(t)$, $B_H(0) = 0$, with stationary increments and the covariance structure given by

$$\mathbb{E}\left\{[B_H(t_1) - B_H(t_2)]^2\right\} = \sigma^2 \, |t_1 - t_2|^{2H}, \tag{1.8}$$

where $H \in (0, 1)$ and $\sigma^2 > 0$ are two parameters. Although first introduced by Kolmogorov in 1940, fBm became very popular after the seminal work of Mandelbrot and van Ness [40] and proved to be a very rich mathematical object of high utility, see e.g. articles by M. Taqqu and by G. Molchan in the book [43] for an introduction and further references and applications. The utility of fBm is related to its properties of being *self-similar*, i.e. $\{B_H(at) : t \in \mathbb{R}\} \stackrel{d}{=} a^H \{B_H(t) : t \in \mathbb{R}\}$ for any $a > 0$, and having *stationary increments*. These two properties characterize the corresponding Gaussian process uniquely, see, e.g., [43]. In the context of self-similarity parameter H is also known as the Hurst index H or the scaling exponent.

For $H = 1/2$, the fBm $B_{1/2}(t)$ is proportional to the usual Brownian motion (Wiener process). We will denote the latter simply as $B(t)$, with $B(dt)$ being the corresponding white noise measure, $\mathbb{E}\{B(dt)\} = 0$ and $\mathbb{E}\{B(dt)B(dt')\} = \delta(t - t')dtdt'$, where we have chosen the normalization corresponding to the choice of $\sigma = 1$ in (1.8).

It is apparent from (1.8) that the naive limit $H = 0$ of $B_H(t)$ is not well-defined. To overcome this problem, the first author proposed some time ago to regularize the fBm in the limit $H \to 0$ as follows.

Consider the stochastic Fourier integral

$$B_H^{(\eta)}(t) = \frac{1}{2\sqrt{2}} \int_0^\infty \frac{e^{-\eta s}}{s^{1/2+H}} \left[\left(e^{-its} - 1\right) B_c(ds) + \left(e^{its} - 1\right) \overline{B_c(ds)} \right], \quad \eta \geq 0\,, \tag{1.9}$$

where $B_c(t) = B_R(t) + iB_I(t)$ and $B_R(t)$ and $B_I(t)$ are two independent copies of the Brownian motion. For $H \in (0,1)$ the integral in (1.9) is well defined for all $\eta \geq 0$ and represents a zero-mean Gaussian process with stationary increments and covariance $\mathbb{E}\left\{[B_H^{(\eta)}(t_1) - B_H^{(\eta)}(t_2)]^2\right\} = 2\phi_H^{(\eta)}(t_1 - t_2)$, where

$$\begin{aligned} \phi_H^{(\eta)}(t) &= \frac{1}{2} \int_0^\infty \frac{e^{-2\eta s}}{s^{1+2H}} \left(1 - \cos(ts)\right) ds \\ &= \frac{1}{4H} \Gamma(1-2H) \left[(4\eta^2 + t^2)^H \cos\left(2H \arctan \frac{t}{2\eta}\right) - (2\eta)^{2H} \right]. \end{aligned} \tag{1.10}$$

For fixed $H \in (0,1)$, $\lim_{\eta \to 0} \phi_H^{(\eta)}(t) = \frac{1}{4H}\Gamma(1-2H)\cos(\pi H) t^{2H}$, where $\Gamma(z)$ is the Euler gamma-function. Hence $B_H^{(0)}(t)$ is fBm. This also follows from the so-called *harmonizable representation* of the fBm, which is precisely the integral on r.h.s. in (1.9) when $\eta = 0$, see Proposition 9.2 in [43], or Eq. (7.16) in [46]. On the other hand, for any fixed $\eta > 0$, the limit of $H = 0$ in (1.9) is well defined, and

$$\lim_{H \downarrow 0} \phi_H^{(\eta)}(t) = \frac{1}{4} \log\left(\frac{t^2}{4\eta^2} + 1\right). \tag{1.11}$$

We consider the resulting limiting process,

$$B_0^{(\eta)}(\tau) = \frac{1}{2\sqrt{2}} \int_0^\infty \frac{e^{-\eta s}}{\sqrt{s}} \left\{ [e^{-i\tau s} - 1]\, B_c(ds) + [e^{i\tau s} - 1]\, \overline{B_c(ds)} \right\} \tag{1.12}$$

as the most natural extension of the standard fBm to the case of zero Hurst index $H = 0$. This process can also be defined axiomatically.

Definition. *The regularized fBm with Hurst index $H = 0$ is a real-valued stochastic process $\{B_0^{(\eta)}(\tau), \tau \in \mathbb{R}\}$ with the following properties*

(i) $B_0^{(\eta)}(t)$ *is a Gaussian process with mean 0 and* $B_0^{(\eta)}(0) = 0$,

(ii) $\mathrm{Var}\{B_0^{(\eta)}(t)\} = \frac{1}{2}\log\left(\frac{t^2}{4\eta^2} + 1\right)$ *for some* $\eta > 0$,

(iii) $B_0^{(\eta)}(t)$ *has stationary increments.*

The increment structure of $B_0^{(\eta)}(t)$ depends logarithmically on the time separation:

$$\mathbb{E}\{[B_0^{(\eta)}(t_1) - B_0^{(\eta)}(t_2)]^2\} = \frac{1}{2} \log\left[\frac{(t_1 - t_2)^2}{4\eta^2} + 1\right], \tag{1.13}$$

and, hence the regularized fBm with $H = 0$ defines a *bona fide* version of the $1/f$ noise with stationary increments[1]. Therefore, the stochastic process $B_0^{(\eta)}(\tau)$ is of interest in its own right and deserves further study. We do not pursue this direction in the present paper except for noting for future reference that the regularized fBm has continuous sample paths.

Note. After posting the initial version of this paper to the arXiv, we learnt of the work [52], where a regularization of fBm essentially equivalent to our $B_H^{(\eta)}(t)$ was introduced for $H > 0$. Note that neither the limit $H \to 0$ nor the connection with random matrices were identified or investigated there.

[1] Compare (1.12) with a stationary version of fBm with $H = 0$ proposed in Eq. (16) of [47]

2. Main results

2.1. Macroscopic regime

We start with the simpler case of the macroscopic scale where we extend the analogous construction of [30] from unitary to Hermitian matrices. The relation between characteristic polynomials of Haar unitary matrices and the random Fourier series in (1.6) can be understood by expanding $\log|p_N(\theta)|$ into the Fourier series

$$V_N(\theta) = -2\log|\det(I - Ue^{-i\theta})| = \sum_{n=1}^{\infty} \frac{1}{\sqrt{n}} \left(v_{n,N} e^{in\theta} + \overline{v_{n,N}} e^{-in\theta} \right), \tag{2.1}$$

where $v_{n,N} = \frac{1}{\sqrt{n}} \operatorname{Tr}(U^{-n})$. Now, the coefficients $v_{n,N}$ converge in distribution as $N \to \infty$ to independent standard complex Gaussians. This is a result due to Diaconis and Shahshahani [15] from which it can be inferred [30] that (1.6) represents the limit of $V_N(\theta)$ in a suitable functional space.

An analogue of the Diaconis-Shahshahani result for the $N \times N$ GUE matrices $\mathcal{H}$ was obtained by Johansson [31]. He proved that for any fixed m the vector $\left(\frac{2}{\sqrt{n}} \operatorname{Tr} T_n(\mathcal{H})\right)_{n=1}^{m}$, with $T_n(x) = \cos(n \arccos(x))$ being Chebyshev polynomials, converges, after centering, to a collection of independent standard Gaussians in the limit $N \to \infty$. In view of the handy identity

$$-\log(2|x-y|) = \sum_{n=1}^{\infty} \frac{2}{n}\, T_n(x) T_n(y), \quad x, y \in [-1, 1], \quad x \neq y\,, \tag{2.2}$$

the desired analogue of Fourier expansion is an expansion in terms of Chebyshev polynomials,

$$D_N(x) = -\log|\det(\mathcal{H} - xI)| = \sum_{n=1}^{\infty} \frac{a_{n,N}}{\sqrt{n}} T_n(x) + N\log 2 + R_N(x), \quad a_{n,N} = \frac{2}{\sqrt{n}} \operatorname{Tr} T_n(\mathcal{H}), \tag{2.3}$$

where the error term $R_N(x)$ is due to the eigenvalues of $\mathcal{H}$ outside the support $[-1, 1]$ of the semicircle law. Since the probability of finding such an eigenvalue vanishes fast as $N \to \infty$ it can be shown that the error term does not contribute in the limit (see the proof of Proposition 5.2 for a more precise statement). One then concludes that the natural limit of $D_N(x)$, after centering, is given by the random Chebyshev-Fourier series (1.7).

We will make this picture mathematically rigorous by working in a suitable functional space. First, let us assign a formal meaning to the series in (1.7) and the corresponding stochastic process. Consider the space $L^2 = L^2((-1, 1), \mu(dx))$ with $\mu(dx) = dx/\sqrt{1-x^2}$. The Chebyshev polynomials form an orthogonal basis in this space, with $c_n(f)$ (1.3) being the coefficients of the corresponding Chebyshev-Fourier series. For $a > 0$, consider the space $V^{(a)}$ of functions f in L^2 such that $\sum_{n=0}^{\infty} |c_n(f)|^2 (1 + n^2)^a < \infty$. This is a Hilbert space with the inner product

$$\langle f, g \rangle_a = \sum_{n=0}^{\infty} c_n(f) c_n(g) (1 + n^2)^a\,.$$

Its dual, $V^{(-a)}$, is the Hilbert space of generalised functions $F(x) = \sum_{n=0}^{\infty} c_n T_n(x)$ with $||F||_{-a}^2 = \sum_{n=0}^{\infty} |c_n|^2 (1 + n^2)^{-a} < \infty$. Setting here $c_0 = 0$ and $c_n = a_n/\sqrt{n}$ with a_n, $n \geq 1$, being independent standard Gaussians, one obtains $F(x)$ of (1.7). In such case $||F||_{-a}^2$ is finite with probability one. This defines $F(x)$ in (1.7) as a generalised random function (stochastic process) which acts on a test function $f \in V^{(a)}$ in the usual way,

$$F[f] = \sum_{n=1}^{\infty} \frac{a_n}{\sqrt{n}}\, c_n(f) = \langle f, F \rangle_0\,.$$

This process is Gaussian with zero mean. Its covariance, $\mathbb{E}\{F[f]F[g]\}$, is given by

$$\mathbb{E}\{F[f]F[g]\} = \sum_{n=1}^{\infty} \frac{1}{n} \int_{-1}^{1} \int_{-1}^{1} f(x)g(y)T_n(x)T_n(y)\ \mu(dx)\mu(dy)\,. \tag{2.4}$$

It can be shown, see e.g. Lemma 3.1 in [27], that the order of summation and integration in (2.4) can be interchanged, and, in view of (2.2), one obtains the covariance operator in closed form:

$$\mathbb{E}\{F[f]F[g]\} = -\int_{-1}^{1} \int_{-1}^{1} \frac{1}{2} \log(2|x-y|) f(x)g(y)\ \mu(dx)\mu(dy), \quad f,g \in V^{(a)}\,.$$

We are now in a position to formulate our result. Consider the centered process

$$\tilde{D}_N(x) = -\log|\det(\mathcal{H}-xI)| + \mathbb{E}\{\log|\det(\mathcal{H}-xI)|\}, \quad x \in (-1,1)\,. \tag{2.5}$$

Since $\log|x|$ is locally integrable, $\tilde{D}_N \in V^{(-a)}$ for every N.

Theorem 2.1. *For every $a > 1/2$, $\tilde{D}_N(x) \Rightarrow F(x)$ in $V^{(-a)}$ as $N \to \infty$, where $F(x)$ given by (1.7).*

Our proof of this theorem in Section 5 involves solving at least two technical problems that did not arise in [30]. First, when proving convergence of the finite-dimensional distributions of $\tilde{D}_N(x)$, we are faced with a test function possessing square-root singularities at the edges of the spectrum, arising from the Chebyshev-Fourier coefficients of the logarithm outside $[-1,1]$, see Lemma 5.1. Most bounds and concentration inequalities for linear statistics rely on the test function having at least $C^1(\mathbb{R})$ regularity, see *e.g.* [42, 38, 1], while ours is only $C^{1/2}(\mathbb{R})$ (even the recent extension [50] of such bounds to test functions from the $C^{1/2+\epsilon}(\mathbb{R})$ class does not suffice here). Making use of fine asymptotics of orthogonal polynomials and Airy functions, we prove that this linear statistic converges to zero, a problem that did not appear in [30].

Secondly, when proving tightness of $(\tilde{D}_N(x))_{N=1}^{\infty}$ we need additional control over the variance of $\mathrm{Tr}(T_n(\mathcal{H}))$ for both large N *and* large n. In [30], the analogous quantity, namely $\mathrm{Var}\{\mathrm{Tr}(U^{-n})\}$, was known explicitly due to exact results for the unitary group obtained by Diaconis and Shashahani [15]. In contrast, for the GUE case, $\mathrm{Var}\{\mathrm{Tr}(T_n(\mathcal{H}))\}$ and related quantities need to be estimated asymptotically as $N \to \infty$, *uniformly* in the degree n of the Chebyshev polynomial.

2.2. Mesoscopic regime

Now we proceed to our next task of extending the relation between characteristic polynomials of random matrices and $1/f$-noises to the mesoscopic scale. In this case, instead of working directly with a generalised stochastic process, we find it more convenient to work with their *regularized* versions.

To formulate our results more precisely, fix a parameter $\eta > 0$ and consider the following sequence of stochastic processes $\{W_N^{(\eta)}(\tau) : \tau \in \mathbb{R}\}$, $N = 1, 2, \ldots$:

$$W_N^{(\eta)}(\tau) = -\log\left|\det\left[\mathcal{H} - \left(x_0 - \frac{\tau}{d_N}\right)I - \frac{i\eta}{d_N}I\right]\right| + \log\left|\det\left[\mathcal{H} - x_0 I - \frac{i\eta}{d_N}I\right]\right|. \tag{2.6}$$

Note that $W_N^{(\eta)}(\tau)$ also depends implicitly on three additional parameters: $\eta > 0$, $x_0 \in (-1,1)$ and $d_N > 0$; their importance is explained below, though for ease of notation we will not emphasize the dependence on x_0 when referring to $W_N^{(\eta)}(\tau)$. We use the parameter $d_N > 0$ to zoom into the appropriate spectral scale of $\mathcal{H}$ centered around a point x_0 inside the bulk of the limiting spectrum of the GUE matrices $\mathcal{H}$. On the macroscopic scale $d_N = 1$, on the microscopic scale $d_N = N$ whilst on the mesoscopic scale d_N is in between these two extremes, $1 \ll d_N \ll N$. The parameter η is an arbitrary but fixed positive real number, introduced to regularize the logarithmic singularity at zero.

Our main result shows that in the *mesoscopic limiting regime* where

$$d_N \to \infty \text{ and } d_N = o(N/\log N) \text{ as } N \to \infty \tag{2.7}$$

the stochastic process $W_N^{(\eta)}(\tau)$ converges, after centering, to $B_0^{(\eta)}(\tau)$; the regularized fractional Brownian motion with Hurst index $H = 0$. For finite-dimensional distributions this is the content of the following Theorem. Let

$$\tilde{W}_N^{(\eta)}(\tau) = W_N^{(\eta)}(\tau) - \mathbb{E}\{W_N^{(\eta)}(\tau)\}\,.$$

Theorem 2.2. *Consider GUE random matrices $\mathcal{H}$ in (1.1). Assume that the reference point x_0 is in the bulk of the limiting spectrum of $\mathcal{H}$, $x_0 \in (-1, 1)$, and the scaling factor d_N satisfies (2.7). Then for any fixed $\eta > 0$ and any finite number of times $(\tau_1, \dots, \tau_m) \in \mathbb{R}^m$ we have the convergence in distribution*

$$(\tilde{W}_N^{(\eta)}(\tau_1), \dots, \tilde{W}_N^{(\eta)}(\tau_m)) \stackrel{d}{\Longrightarrow} (B_0^{(\eta)}(\tau_1), \dots, B_0^{(\eta)}(\tau_m)), \qquad \textit{as } N \to \infty. \tag{2.8}$$

We prove this theorem in Section 3 by adopting Krasovsky's derivation of identity (1.4) to the mesoscopic scale. The characteristic function of the random vector on the l.h.s. in (2.8) is given by a Hankel determinant whose symbol possesses Fisher-Hartwig singularities. The Riemann-Hilbert problem provides a powerful tool to obtain asymptotics of such Hankel determinants [14, 37, 36, 35]. On the mesoscopic scale the Fisher-Hartwig singularities (these are located at points $x_0+(\tau_k+i\eta)/d_N$) are all at distance of order $1/d_N$ from the point $x_0 \in (-1, 1)$. Because of this, the system of contours defining the Riemann-Hilbert problem (inside of which the symbol is analytic) close onto the real line as $N \to \infty$. In this regime, the estimates become more delicate. In contrast, in the macroscopic regime the Fisher-Hartwig singularities are real and spaced out and one does not need to consider the case of shrinking contours.

Here it is appropriate to mention that linear eigenvalue statistics on the mesoscopic scale are more challenging to study compared to the macroscopic scale. Known results are sparse and mostly limited to regular test functions with compact support, see [9, 10, 49] and also more recent works [18, 19, 16, 8, 11]. One reason is that the majority of concentration inequalities involving derivatives, such as *e.g.* Lipschitz norm [1] or the Poincaré inequality [1, 42], that proved to be so useful on the macroscopic scale, get a factor of d_N in the mesoscopic case and, hence, no longer apply without appropriate modification. In this context, the Riemann-Hilbert problem proves to be a powerful tool for estimating the error terms down to very small scales (2.7).

One can extend Theorem 2.2 to an infinite-dimensional setting with a little bit more work. Let $L^2[a, b]$ denote the Hilbert space of square integrable functions on $[a, b]$ with the inner product

$$\langle f, g\rangle_2 = \int_a^b f(\tau)\overline{g(\tau)}\, d\tau. \tag{2.9}$$

Since the sample paths of $\tilde{W}_N^{(\eta)}$ are continuous, $\|\tilde{W}_N^{(\eta)}\|_2 < \infty$. Therefore, both $W_N^{(\eta)}$ and its $N \to \infty$ limit $B_0^{(\eta)}$ can be viewed as random elements in the space $L^2[a, b]$. We have,

Theorem 2.3. *Let $-\infty < a < b < \infty$. Then on mesoscopic scales (2.7), the process $\tilde{W}_N^{(\eta)}$ converges weakly (in the sense of probability law) to $B_0^{(\eta)}$ in $L^2[a, b]$ as $N \to \infty$. Furthermore, for every $h \in L^2[a, b]$, we have the convergence in distribution*

$$\int_a^b h(\tau)\tilde{W}_N^{(\eta)}(\tau)\, d\tau \stackrel{d}{\Longrightarrow} \int_a^b h(\tau)B_0^{(\eta)}(\tau)\, d\tau, \qquad N \to \infty\,. \tag{2.10}$$

This result follows from Theorem 3 in [28], which allows one to deduce weak convergence for general processes in $L^2[a, b]$ under the hypothesis that

(i) The finite-dimensional distributions of $\tilde{W}_N^{(\eta)}$ converge to those of $B_0^{(\eta)}$ as $N \to \infty$.

(ii) For some $C > 0$, the bound $\mathbb{E}\{|\tilde{W}_N^{(\eta)}(\tau)|^2\} \leq C$ holds for all N and $\tau \in [a, b]$ and

$$\lim_{N\to\infty} \mathbb{E}\{|\tilde{W}_N^{(\eta)}(\tau)|^2\} = \mathbb{E}\{|B_0^{(\eta)}(\tau)|^2\}. \tag{2.11}$$

Note that item (i) is a restatement of Theorem 2.2, while item (ii) will be shown to follow from our proof of Theorem 2.2.

Having established the relation between characteristic polynomials of GUE matrices and $1/f$ noise on the mesoscopic scale, let us revisit the series expansions of the macroscopic scale discussed at length in Sec. 2.1. Instead of expanding the process $W_N^{(\eta)}(\tau)$ in a Chebyshev-Fourier series and applying the Diaconis-Shahshahani result, in the mesoscopic regime it comes in handy to expand $W_N^{(\eta)}(\tau)$ as a Fourier *integral*.

To this end, we now provide a suitable Fourier-integral representation for $W_N^{(\eta)}(\tau)$. Such a representation can be derived by making use of the identity (see, e.g., Eq. 7.89 in [14])

$$\frac{1}{2}\log\left(\frac{t^2}{\epsilon^2}+1\right) = \int_0^\infty \frac{e^{-\epsilon s}}{s}[1-\cos(ts)]\,ds, \quad \epsilon > 0\,. \tag{2.12}$$

It follows from (2.12) that

$$W_N^{(\eta)}(\tau) = \frac{1}{2}\int_0^\infty \frac{e^{-\eta s}}{\sqrt{s}}\left\{[e^{-i\tau s}-1]\,b_N(s) + [e^{i\tau s}-1]\,\overline{b_N(s)}\right\}\,ds \tag{2.13}$$

where

$$b_N(s) = \frac{1}{\sqrt{s}}\,\mathrm{Tr}\,e^{-isd_N(\mathcal{H}-x_0 I)}\,. \tag{2.14}$$

The identity (2.13) can be thought of as the Fourier integral version of the Fourier series (2.1). Furthermore, comparison of the harmonizable representation (1.12) for $B_0^{(\eta)}(t)$ (which can be thought as a natural integral analogue of the series expansions in (1.6)) and (2.13)), suggests that the Fourier coefficients $b_N(s)$ converge in the mesoscopic regime to Gaussian white noise. Such a statement may be interpreted as a continuous analogue of the Diaconis-Shahshahani result [15] and is the content of our next theorem.

Let $C_0^\infty(\mathbb{R}_+)$ be the space of infinitely many times differentiable functions with compact support on $\mathbb{R}_+ = \{x \in \mathbb{R} : x > 0\}$. Denote

$$c_N(\xi) = \int_0^\infty \xi(s)\,b_N(s)\,ds. \tag{2.15}$$

Theorem 2.4. *Consider the mesoscopic regime where $d_N = N^\alpha$ with any $\alpha \in (0,1)$. Then for every $\xi \in C_0^\infty(\mathbb{R}_+)$*

$$\lim_{N\to\infty} \mathbb{E}\{e^{-i\,\mathrm{Re}\,c_N(\xi)}\} = \exp\left(-\frac{1}{4}\int_0^\infty |\xi(s)|^2\,ds\right). \tag{2.16}$$

Furthermore, for any finite number of $\xi_j \in C_0^\infty(\mathbb{R}_+)$, the vector $(c_N(\xi_1), \ldots, c_N(\xi_m))$ converges in distribution, as $N \to \infty$, to the centered complex Gaussian vector $Z \in \mathbb{R}^m$ having relation matrix $\mathbb{E}(ZZ^{\mathrm{T}}) = 0$ and covariance matrix $\Gamma = \mathbb{E}(ZZ^\dagger)$ given by

$$\Gamma_{j,k} = \int_0^\infty \xi_j(s)\overline{\xi_k(s)}\,ds, \qquad j, k = 1, \ldots, m. \tag{2.17}$$

Proof. See Section 4. □

Remark 2.5. As is often the case in random matrix theory, linear eigenvalue statistics such as (2.15) have variance of the order of unity due to strong correlations between the eigenvalues and converge to a Gaussian random variable after centering. One would typically expect that $\mathbb{E}\{c_N(\xi)\} = O(N/d_N)$ as $N \to \infty$. Instead, we find, see Section 4, that the smoothness of ξ and the rapid oscillations in (2.14) imply $\mathbb{E}\{c_N(\xi)\} = O(d_N^{-1})$ as $N \to \infty$ and, thus, centering is not really needed.

The rest of the paper is organized as follows. Section 3 is devoted to the proof of Theorem 2.2. To do this, we begin by adapting the differential identity used in [35] and then outline the relevant asymptotic analysis of the Riemann-Hilbert problem, leaving estimation of all error terms to Appendix A. Section 4 is devoted to proving the convergence of the Fourier coefficients $b_N(s)$ to the white noise. In the final section we focus on the macroscopic scale and prove Theorem 2.1.

3. Mesoscopic regime

In this section we prove Theorem 2.2. Let us fix $m-1$ distinct times $\tau_1, \ldots, \tau_{m-1}$, $m \geq 2$, and consider the characteristic function

$$\varphi_N(\alpha_1, \ldots, \alpha_{m-1}) = \mathbb{E}\left\{\exp\left(\sum_{k=1}^{m-1} \alpha_k W_N^{(\eta)}(\tau_k)\right)\right\}$$

of the random vector $(W_N^{(\eta)}(\tau_1), \ldots, W_N^{(\eta)}(\tau_{m-1}))$. Our strategy will be to prove that φ_N converges to the characteristic function of the multivariate Gaussian distribution in the limit $N \to \infty$. Theorem 2.2 will then follow by inspection of the quadratic form in the exponential.

To begin with, we will write the characteristic function φ_N as the partition function of a matrix model with Gaussian weight, modified by the singularities

$$\mu_k = \sqrt{2N}\left(x_0 + \frac{\tau_k + i\eta}{d_N}\right), \qquad \eta > 0, \tag{3.1}$$

where $k = 1, \ldots, m$ and $\tau_m \equiv 0$. A standard calculation (changing variables of integration from $\mathcal{H}$ to the eigenvalues and eigenvectors of $\mathcal{H}$ and integrating out the eigenvectors, see e.g. [42]) yields

$$\varphi_N(\alpha_1, \ldots, \alpha_{m-1}) = \frac{1}{C}\int_{\mathbb{R}^N} \prod_{j=1}^{N} w(x_j) \prod_{1\leq i<j\leq N} (x_i - x_j)^2 \, dx_1 \ldots dx_N \tag{3.2}$$

where the weight function is given by

$$w(x) = e^{-x^2} \prod_{k=1}^{m} |x - \mu_k|^{\alpha_k}, \qquad \mathrm{Im}(\mu_k) \neq 0, \qquad k = 1, \ldots, m \tag{3.3}$$

and $\alpha_m = -\alpha_1 - \ldots - \alpha_{m-1}$. Note the discrepancy with the measure (1.1); for convenience we have changed variables $x_j \to x_j/\sqrt{2N}$, the resulting multiplicative constants cancelling each other out.

Our calculation will be guided by that of Krasovsky [35] who treated a similar partition function, but only for the macroscopic regime $d_N = 1$ and $\eta = 0$. In that case the weight function acquires Fisher-Hartwig singularities inside the spectral interval $(-1, 1)$. In contrast, our weight (3.3) posesses singularities in the complex plane that merge towards the point x_0 on the spectral axis at rate d_N as $N \to \infty$. Since this merging process occurs sufficiently slowly (i.e. $d_N = o(N)$), these singularities will not play a crucial role in the calculation.

A special feature of the weight function (3.3) is the *cyclic condition*

$$\sum_{k=1}^{m} \alpha_k = 0. \tag{3.4}$$

This holds because the second term in (2.6) is independent of τ. Our first step is to express the partition function (3.2) in a form suitable for the computation of asymptotics.

3.1. Orthogonal polynomials and differential identity

The multiple integral in (3.2) is intimately connected to the theory of orthogonal polynomials. Let

$$p_n(x) = \chi_n(x^n + \beta_n x^{n-1} + \gamma_n x^{n-2} + \ldots), \quad n = 0, 1, 2, \ldots,$$

be orthogonal polynomials with respect to weight function $w(x)$: $\int_{-\infty}^{\infty} p_m(x)p_n(x)w(x)\,dx = \delta_{m,n}$. When the α_j's are real and each $\alpha_j > -1/2$ we have $w(x) \geq 0$ and the existence of the polynomials $p_n(x)$ is well known [14]. Then, as in [35], the coefficients χ_n, β_n and γ_n and the polynomials $p_n(x)$ are defined for any $\{\alpha_j\}_{j=1}^{m} \in \mathbb{C}^m$ via analytic continuation, provided each $\mathrm{Re}(\alpha_j) > -1/2$.

Now, the partition function (3.2) can be written in terms of the coefficients $\{\chi_j\}_{j=1}^{N}$ (see *e.g.* [41])

$$\varphi_N(\alpha_1, \ldots, \alpha_{m-1}) = \frac{N!}{C} \prod_{j=0}^{N-1} \chi_j^{-2}. \tag{3.5}$$

Thus, in principle, our problem is reduced to computing the asymptotics of the orthogonal polynomials and related quantities with respect to the weight $w(x)$. The crucial point observed in [35] is that by taking the logarithmic derivative on both sides of (3.5) with respect to any of the α_j's, the right-hand side can be written as a sum involving only $O(m)$ terms, rather than N. To state the resulting *differential identity* we also need the following 2×2 matrix involving the orthogonal polynomials and their Cauchy transforms:

$$Y(z) = \begin{pmatrix} \chi_N^{-1} p_N(z) & \chi_N^{-1} \displaystyle\int_{-\infty}^{\infty} \frac{p_N(x)}{x - z} \frac{w(x)dx}{2\pi i} \\ -2\pi i \chi_{N-1} p_{N-1}(z) & -\chi_{N-1} \displaystyle\int_{-\infty}^{\infty} \frac{p_{N-1}(x)}{x - z} w(x)dx \end{pmatrix}. \tag{3.6}$$

Lemma 3.1. *For each $k = 1, \ldots, m$, let μ_k in (3.3) be any complex parameters satisfying $\mathrm{Im}(\mu_k) \neq 0$ and define $\alpha_{m+k} = \alpha_k$, $\mu_{m+k} = \overline{\mu_k}$. Denoting by $'$ differentiation with respect to α_j, the following formula holds for any $j = 1, \ldots, m$.*

$$\begin{aligned} (\log \varphi_N)' = &-N(\log \chi_N \chi_{N-1})' - 2\left(\frac{\chi_{N-1}}{\chi_N}\right)^2 \left(\log \frac{\chi_{N-1}}{\chi_N}\right)' + 2(\gamma_N' - \beta_N \beta_N') \\ &+ \frac{1}{2}\sum_{k=1}^{2m} \alpha_k (Y_{11}(\mu_k)' Y_{22}(\mu_k) - Y_{21}(\mu_k)' Y_{12}(\mu_k) + (\log \chi_N \chi_{N-1})' Y_{11}(\mu_k) Y_{22}(\mu_k)). \end{aligned} \tag{3.7}$$

Proof. The proof follows from simple modifications of the arguments given in Sec. 3 of [35]. In fact, further simplifications occur due to the cyclic condition $\sum_{k=1}^{m} \alpha_k = 0$ and the fact that the singularities μ_k have non-zero imaginary part ($k = 1, \ldots, m$). □

Note that χ_N and the coefficients β_N and γ_N can be computed from the relations:

$$\begin{aligned} Y_{11}(z) &= z^N + \beta_N z^{N-1} + \gamma_N z^{N-2} + \ldots \\ \chi_{N-1}^2 &= \lim_{z \to \infty} \frac{i Y_{21}(z)}{2\pi z^{N-1}} \end{aligned} \tag{3.8}$$

Therefore, our plan will be to compute the asymptotics of $Y(z)$ and then, by making use of identities (3.8), evaluate the right-hand side of (3.7) to the desired accuracy in the limit as $N \to \infty$. We will find that the error terms in the asymptotics are uniform in the variables $\{\alpha_k\}_{k=1}^{m-1}$ belonging to a compact subset of

$$\Omega = \{(\alpha_1, \ldots, \alpha_{m-1}) \mid \mathrm{Re}(\alpha_k) > -1/2, \qquad k = 1, \ldots, m-1\}. \tag{3.9}$$

This uniformity property then allows us to integrate the identity (3.7) recursively with respect to $\{\alpha_k\}_{k=1}^{m-1}$ and obtain asymptotics for the characteristic function (3.2). The asymptotics of $Y(z)$ in the limit $N \to \infty$ can be obtained by using an appropriate Riemann-Hilbert problem. Although this technique is nowadays standard, for the reader's convenience we will briefly summarise the necessary ingredients of the corresponding calculation.

3.2. The Riemann-Hilbert problem for $Y(z)$

The relationship between orthogonal polynomials and Riemann-Hilbert problems was established for general weights in [21] where it was shown that $Y(z)$ solves the following problem:

1. $Y(z)$ is analytic in $\mathbb{C} \setminus \mathbb{R}$.
2. On the real line there is a jump discontinuity
$$Y_+(x) = Y_-(x) \begin{pmatrix} 1 & w(x) \\ 0 & 1 \end{pmatrix}, \quad x \in \mathbb{R}, \tag{3.10}$$
where $Y_+(x)$ and $Y_-(x)$ denote the limiting values of $Y(z)$ as z approaches the point $x \in \mathbb{R}$ from above $(+)$ or below $(-)$.
3. Near $z = \infty$, we have the following asymptotic behaviour
$$Y(z) = \left(I + O\left(\frac{1}{z}\right)\right) z^{N\sigma_3}. \tag{3.11}$$

Here σ_3 is the third Pauli matrix and serves as a convenient notational tool. By definition of the matrix exponential, the notation in (3.11) has the meaning
$$z^{N\sigma_3} = \begin{pmatrix} z^N & 0 \\ 0 & z^{-N} \end{pmatrix}. \tag{3.12}$$

One can verify directly that $Y(z)$ of (3.6) does indeed solve this Riemann-Hilbert problem, while the uniqueness of this solution can be deduced from the observation that $\det Y(z) \equiv 1$, in conjunction with the Liouville theorem. Further details regarding existence and uniqueness of the problem can be found in [14].

In order to obtain asymptotics as $N \to \infty$, we will perform a sequence of transformations to our initial Riemann-Hilbert problem known as the *Deift-Zhou steepest descent* (see e.g. [14] and [13]). The purpose of these transformations is to identify a 'limiting' problem that can be solved with elementary functions, giving the leading order asymptotics to $Y(z)$. For the reader's convenience, we briefly describe the key points underlying these transformations:

1. The first transformation $Y \to T$ normalizes the unsatisfactory asymptotic behaviour in the third condition, equation (3.11). This comes with the cost that the entries of the jump matrix for $T(z)$ on the interval $(-1, 1)$ are now oscillating in N and do not have a limit as $N \to \infty$.
2. The second transformation $T \to S$ aims to remove these oscillations by splitting the contour $(-1, 1)$ into lens shaped contours where now the jump matrices are exponentially close to the identity. For our particular *mesoscopic* problem, we need the lenses to pass below the singularities for each $k = 1, \ldots, m$, so that their distance from $(-1, 1)$ is of order $O(d_N^{-1})$ (see Figure 1).

3. Now it turns out that the jump matrices for S tend to the identity as $N \to \infty$, except on the contour $(-1,1)$. But the jump across $(-1,1)$ is of a special form that can be solved exactly in terms of elementary functions. This solution, denoted $P_\infty(z)$, gives the leading order contribution to the asymptotics in the required regions of the complex plane.

In Sec. 3.5 we will show that the asymptotics obtained in this way lead directly to Theorem 2.2. However, to complete the proof, one has to show that the conclusion of (3), namely that $S(z) \sim P_\infty(z)$ as $N \to \infty$, is really correct. This may be regarded as the most technical part of the Deift-Zhou method. The main problem is that although the jump matrix for $S(z)$ converges to that of $P_\infty(z)$, this convergence is not uniform near the edges $z = \pm 1$. To remedy this, local solutions known as *parametrices* have to be constructed near these points, and then matched to leading order with the so-called outer parametrix $P_\infty(z)$. These final technical issues will be addressed in Appendix A.

3.3. T and S transformations of the Riemann-Hilbert problem

The T transformation is performed in the usual way. First we define the g-function:

$$g(z) = \int_{-1}^{1} \log(z-s)\rho(s)\,ds, \qquad z \in \mathbb{C} \setminus (-\infty, 1], \tag{3.13}$$

where throughout we take the principal branch of the logarithm. Here and below $\rho(s) = (2/\pi)\sqrt{1-s^2}$ denotes the limiting density of eigenvalues. The $Y \to T$ transformation is then given by the formula

$$Y(z\sqrt{2N}) = (2N)^{N\sigma_3} e^{Nl\sigma_3/2} T(z) e^{N(g(z)-l/2)\sigma_3} \tag{3.14}$$

where $l = -1 - 2\log(2)$. Notice that we have rescaled the Riemann-Hilbert problem so that the singularities of the corresponding weight function are of order $O(1)$ as $N \to \infty$, so that from now on we deal with singularities of the form

$$z_k = \frac{\mu_k}{\sqrt{2N}} = x_0 + \frac{\tau_k + i\eta}{d_N}. \tag{3.15}$$

The resulting jump matrix for $T(z)$ can now be computed from the standard properties of the g-function:

$$\begin{aligned} &g_+(x) + g_-(x) - 2x^2 - l = 0, \qquad x \in (-1,1),\\ &g_+(x) + g_-(x) - 2x^2 - l < 0, \qquad x \in \mathbb{R} \setminus [-1,1],\\ &g_+(x) - g_-(x) = \begin{cases} 2\pi i & x \leq -1 \\ 2\pi i \displaystyle\int_x^1 \rho(s)ds & x \in [-1,1] \\ 0 & x \geq 1. \end{cases} \end{aligned} \tag{3.16}$$

In addition, since $g(z) \sim \log(z)$ as $z \to \infty$, we have $e^{Ng(z)\sigma_3} \sim z^{N\sigma_3}$. Thus one easily verifies that $T(z)$ is normalized at $z = \infty$. We now have the following Riemann-Hilbert problem for $T(z)$:

1. $T(z)$ is analytic in $\mathbb{C} \setminus \mathbb{R}$.

2. We have the jump condition

$$T_+(x) = T_-(x)\begin{pmatrix} e^{-N(g_+(x)-g_-(x))} & \prod_{k=1}^{m}|x-z_k|^{\alpha_k} \\ 0 & e^{N(g_+(x)-g_-(x))}\end{pmatrix}, \quad x \in (-1,1), \tag{3.17}$$

$$T_+(x) = T_-(x)\begin{pmatrix} 1 & \prod_{k=1}^{m}|x-z_k|^{\alpha_k}e^{N(g_+(x)+g_-(x)-2x^2-l)} \\ 0 & 1\end{pmatrix}, \quad x \in \mathbb{R}\setminus[-1,1]. \tag{3.18}$$

3. $T(z) = I + O(z^{-1})$ as $z \to \infty$.

We see that although the problem for $T(z)$ is normalized at ∞, the jump matrix (3.17) on $(-1,1)$ has oscillatory diagonal entries that not have a limit as $N \to \infty$. The Deift-Zhou steepest descent procedure remedies this situation by splitting the contour $(-1,1)$ into 'lenses' in the complex plane (see Figure 1), transforming the unwanted oscillations into exponentially decaying matrix elements.

This procedure is facilitated by the factorization of the jump matrix on $(-1,1)$:

$$\begin{pmatrix} e^{-Nh(x)} & \omega(x) \\ 0 & e^{Nh(x)}\end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \omega(x)^{-1}e^{Nh(x)} & 1\end{pmatrix}\begin{pmatrix} 0 & \omega(x)^{-1} \\ -\omega(x)^{-1} & 0\end{pmatrix}\begin{pmatrix} 1 & 0 \\ \omega(x)^{-1}e^{-Nh(x)} & 1\end{pmatrix}$$

where

$$\omega(x) = \prod_{k=1}^{m}|x-z_k|^{\alpha_k} \tag{3.19}$$

$$h(x) = g_+(x) - g_-(x) = -2\pi i\int_1^x \rho(y)dy \tag{3.20}$$

The latter objects (3.19) and (3.20) possess analytic continuations into the lens shaped regions depicted in Figure 1. For the weight $\omega(x)$ we have

$$\omega(z) = \prod_{k=1}^{m-1}\left[\frac{(z-x_0-\tau_k/d_N)^2+(\eta/d_N)^2}{(z-x_0)^2+(\eta/d_N)^2}\right]^{\alpha_k/2}, \tag{3.21}$$

where throughout we take the principal branch of the roots. This function is analytic for all z such that the inequality

$$(\mathrm{Re}(z)-\mathrm{Re}(z_k))^2 > (\mathrm{Im}(z_k))^2 - (\mathrm{Im}(z))^2 \tag{3.22}$$

is satisfied for every $k = 1,\dots,m$. One easily verifies that for $x_0 \in (-1+\delta, 1-\delta)$, the inequality (3.22) holds for any z chosen from the interior region bounded by the lips $\Sigma_{\pm 1}$ and the discs $z \in \partial B_{\pm 1}(\delta)$ of sufficiently small radius (see Figure 1). Finally let $h(z)$ denote the analytic continuation of (3.20) to $\mathbb{C}\setminus((-\infty,-1]\cup[1,\infty))$. We are now ready to define the $T \to S$ transformation. Let

$$S(z) = \begin{cases} T(z), & \text{for } z \text{ outside the lenses,} \\ T(z)\begin{pmatrix} 1 & 0 \\ -\omega(z)^{-1}e^{-Nh(z)} & 1\end{pmatrix}, & \text{for } z \text{ in the upper part of the lenses,} \\ T(z)\begin{pmatrix} 1 & 0 \\ \omega(z)^{-1}e^{Nh(z)} & 1\end{pmatrix}, & \text{for } z \text{ in the lower part of the lenses.}\end{cases} \tag{3.23}$$

Now we get the following Riemann-Hilbert problem for $S(z)$:

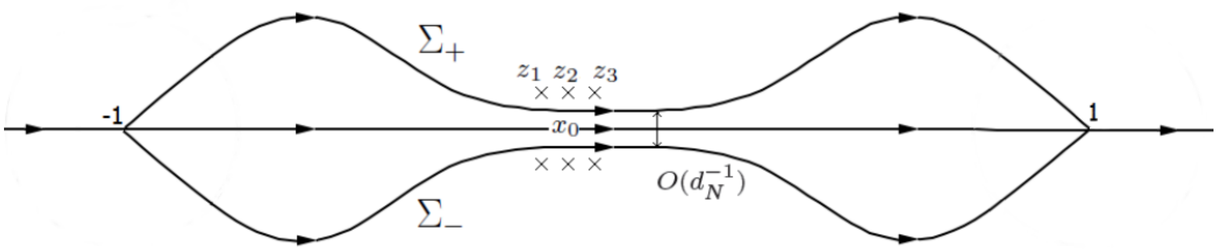

Fig 1. *The contour Σ for the S Riemann-Hilbert problem with $m = 3$. The crosses depict the 3 singularities and their complex conjugates, of distance $O(d_N^{-1})$ from the point $x_0 \in (-1,1)$. The lenses $\Sigma_\pm$ pass between the real line and the singularities into the points ± 1.*

1. $S(z)$ is analytic in $\mathbb{C} \setminus \Sigma$ where $\Sigma = \Sigma_+ \cup \mathbb{R} \cup \Sigma_-$.
2. $S(z)$ has the following jumps on Σ

$$S_+(x) = S_-(x) \begin{pmatrix} 1 & 0 \\ \omega(x)^{-1} e^{\mp N h(x)} & 1 \end{pmatrix}, \qquad x \in \Sigma_\pm,$$

$$S_+(x) = S_-(x) \begin{pmatrix} 0 & \omega(x) \\ -\omega(x)^{-1} & 0 \end{pmatrix}, \qquad x \in (-1,1),$$

$$S_+(x) = S_-(x) \begin{pmatrix} 1 & \omega(x) e^{N(g_+(x)+g_-(x)-2x^2-l)} \\ 0 & 1 \end{pmatrix}, \qquad x \in \mathbb{R} \setminus [-1,1].$$

3. $S(z) = I + O(z^{-1})$ as $z \to \infty$.

At this point in the asymptotic analysis, it becomes clear that the mesoscopic regime under consideration becomes important. In order to obtain asymptotics, it is essential that the jump matrix for $S(z)$ approaches the identity as $N \to \infty$ for $z \in \Sigma_\pm$. In the Appendix (see Prop. A.4) we will see that $|e^{\mp N h(z)}| = O(e^{-c_1 \frac{N}{d_N}})$ as $N \to \infty$ uniformly on $\Sigma_\pm \setminus (B_1(\delta) \cup B_{-1}(\delta))$. Notice that such a bound fails when one approaches the critical situation $d_N = N$ corresponding to the *local* or *microscopic* regime. It is precisely at this scale that one would *not* expect the appearance of a Gaussian process in the limit $N \to \infty$.

Therefore, in the *mesoscopic* regime it is reasonable to expect that in the limit $N \to \infty$ we may neglect the jumps on $\Sigma_\pm \cup (\mathbb{R} \setminus [-1,1])$ and approximate $S(z)$ by a Riemann-Hilbert problem with jumps only on the interval $(-1,1)$. This approximation will be valid only in the region $U_\infty = \mathbb{C} \setminus (B_1(\delta) \cup B_{-1}(\delta))$ and will give rise to an error that is quantified in Appendix A.

3.4. Limiting Riemann-Hilbert problem: Parametrix in U_∞

Before we perform the final transformation $S \to R$ of the Riemann-Hilbert problem, we must construct parametrices in the appropriate regions of the complex plane. We saw in the last section how the jump matrices for $S(z)$ converge to the identity as $N \to \infty$, except on $[-1,1]$. Therefore, outside the lenses and the discs, we expect the solution to the following problem to give a good approximation to $S(z)$ for large N.

1. $P_{\infty}(z)$ is analytic in $\mathbb{C} \setminus [-1, 1]$.

2. We have the jump condition

$$P_{\infty,+}(x) = P_{\infty,-}(x) \begin{pmatrix} 0 & \omega(x) \\ -\omega(x)^{-1} & 0 \end{pmatrix}, \qquad x \in (-1, 1). \tag{3.24}$$

3. $P_{\infty}(z) = I + O(z^{-1})$ as $z \to \infty$.

This problem has the advantage that it has a completely explicit solution. The solution, as obtained in [36], is given by

$$P_{\infty}(z) = \frac{1}{2}(\mathcal{D}_{\infty})^{\sigma_3} \begin{pmatrix} a + a^{-1} & -i(a - a^{-1}) \\ i(a - a^{-1}) & a + a^{-1} \end{pmatrix} \mathcal{D}(z)^{-\sigma_3}, \qquad a(z) = \frac{(z-1)^{1/4}}{(z+1)^{1/4}}, \tag{3.25}$$

where $\mathcal{D}(z)$ is the Szegö function

$$\mathcal{D}(z) = \exp\left(\frac{\sqrt{z+1}\sqrt{z-1}}{2\pi} \int_{-1}^{1} \frac{\log \omega(x)}{\sqrt{1-x^2}} \frac{dx}{z-x}\right) \tag{3.26}$$

and

$$\mathcal{D}_{\infty} = \lim_{z \to \infty} \mathcal{D}(z) = \exp\left(\frac{1}{2\pi} \int_{-1}^{1} \frac{\log \omega(x)}{\sqrt{1-x^2}}\, dx\right). \tag{3.27}$$

Recalling the definition of the weight $\omega(x)$ in (3.19), the integrals in (3.26) can be calculated explicitly by extending the procedure outlined in [35] to the case of complex singularities.

As we shall see in the next subsection, the Szegö function $\mathcal{D}(z)$ will turn out to be the key ingredient in deriving the logarithmic covariance structure in (1.13).

3.5. Asymptotics of the polynomials and proof of Theorem 2.2

We are now ready to present the leading order asymptotics $N \to \infty$ of the Y-matrix in (3.6), leaving the technical matters of estimation of errors and the final transformation of the Riemann-Hilbert problem to Appendix A. Our aim in this subsection is to prove Theorem 2.2 using these asymptotics.

Tracing back the transformations $S \to T \to Y$, we find that

$$Y(z\sqrt{2N}) = (2N)^{N\sigma_3/2} e^{Nl\sigma_3/2} S(z) e^{N(g(z)-l/2)\sigma_3} \tag{3.28}$$

According to (3.7), we need the asymptotics for $Y(z)$ in two different regions of the complex plane, near $z = \infty$ in the first line of (3.7) and at $z = z_k$ in the second line. In the following Proposition, let $\mathcal{A}$ denote the bounded subset of $\mathbb{C}$ enclosed by the lenses $\Sigma_{\pm}$ and the discs $\partial B_{\pm 1}(\delta)$.

Proposition 3.2. *Consider the Riemann-Hilbert problems $S(z)$ and $P_{\infty}(z)$ from Sections 3.3 and 3.4 respectively. Then the following asymptotics hold as $N \to \infty$*

$$S(z) = \left(I + \frac{\tilde{R}_1(z)}{N} + O\left(\frac{1}{N d_N}\right) + O\left(\log(d_N)\, e^{-c_1 \frac{N}{d_N}}\right)\right) P_{\infty}(z), \tag{3.29}$$

uniformly for all $z \in \mathbb{C} \setminus \mathcal{A}$. The function $\tilde{R}_1(z)$ has an asymptotic expansion of the form $\tilde{R}_1(z) = (A/z + B/z^2 + O(z^{-3}))$ as $z \to \infty$ where c_1 is a positive constant depending only on δ and η and

$$A = \begin{pmatrix} 0 & i/24 \\ i/24 & 0 \end{pmatrix}, \qquad B = \begin{pmatrix} -1/48 & 0 \\ 0 & 1/48 \end{pmatrix}. \tag{3.30}$$

Proof. See Appendix A. □

Remark 3.3. The error terms in (3.29) are uniform in the parameters $\{\alpha_k\}_{k=1}^{m-1}$ belonging to Ω (cf. (3.9)), $\{\tau_k\}_{k=1}^{m-1}$ belonging to a compact subset of $\mathbb{R}$ and x_0 belonging to a compact subset of $(-1+\delta, 1-\delta)$. Furthermore, every such error term is an analytic function in the variables $\{\alpha_k\}_{k=1}^{m-1}$ whose derivatives with respect to α_j have the same order in N and have the same uniformity property described above. Hence, in the remainder of this Section it will be implicit that the error terms involved are of this form.

Now inserting the above asymptotics (3.29) into the differential identity (3.7), we obtain

Proposition 3.4. *Let φ_N denote the characteristic function of the stochastic process $W_N^{(\eta)}(\tau)$ defined in (3.2). Then in the limit $N \to \infty$, we have*

$$\begin{aligned}\varphi_N(\alpha_1, \ldots, \alpha_{m-1}) = \exp\Bigg(& N \sum_{k=1}^{m-1} \alpha_k (\mathrm{Re}(g(z_k)) - \mathrm{Re}(g(z_m)))\\ &+ \sum_{k,j=1}^{m-1} \frac{\alpha_k \alpha_j}{2} \left(\phi_0^{(\eta)}(\tau_k) + \phi_0^{(\eta)}(\tau_j) - \phi_0^{(\eta)}(\tau_k - \tau_j)\right)\\ &+ O(d_N^{-1}) + O\left(N \log(d_N) \exp\left(-c_1 \frac{N}{d_N}\right)\right)\Bigg),\end{aligned} \tag{3.31}$$

where $g(z)$ is defined in (3.13) and $\phi_0^{(\eta)}(\tau)$ in (1.11). The asympotics in (3.31) hold uniformly in the same sense described in Remark 3.3.

Remark 3.5. Notice that the asymptotics in (3.31) consist of both *global* error terms, which become large when $d_N \sim 1$ and *local* error terms, which become large when $d_N \sim N$. Throughout the following proof, we will write e_N for the local error term of order

$$e_N = \log(d_N) \exp\left(-c_1 \frac{N}{d_N}\right). \tag{3.32}$$

Proof. We remind the reader that the prime $'$ always denotes differentiation with respect to α_j. We begin by considering the second line of (3.7). Taking into account $\alpha_m = -(\alpha_1 + \ldots + \alpha_{m-1})$, we insert (3.29) into (3.28) and make use of the explicit formula (3.25) for $P_\infty(z)$. Straightforward calculation then gives

$$\begin{aligned}&Y_{11}(\sqrt{2N}z_k)' Y_{22}(\sqrt{2N}z_k) - Y_{21}(\sqrt{2N}z_k)' Y_{12}(\sqrt{2N}z_k)\\ &= (P_\infty(z_k))'_{11}(P_\infty(z_k))_{22} - (P_\infty(z_k))'_{21}(P_\infty(z_k))_{12} + O(N^{-1}) + O(e_N). \qquad (3.33)\\ &= C(z_m, z_k) - C(z_j, z_k) + O(d_N^{-1}) + O(e_N) \qquad (3.34)\end{aligned}$$

where we introduced

$$C(\mu, z) = \frac{\sqrt{z+1}\sqrt{z-1}}{2\pi} \int_{-1}^{1} \frac{\log|x-\mu|}{\sqrt{1-x^2}} \frac{dx}{z-x}, \tag{3.35}$$

and (3.34) was obtained from (3.33) using the estimate $\mathcal{D}_\infty = 1 + O(d_N^{-1})$. Since $C(z_j, \overline{z_k}) = \overline{C(z_j, z_k)}$, we find from (3.34) that

$$\begin{aligned}&\frac{1}{2}\sum_{k=1}^{2m} \alpha_k \left(Y_{11}(\sqrt{2N}z_k)' Y_{22}(\sqrt{2N}z_k) - Y_{21}(\sqrt{2N}z_k)' Y_{12}(\sqrt{2N}z_k)\right) \qquad (3.36)\\ &= \sum_{k=1}^{m} \alpha_k \left(\mathrm{Re}(C(z_m, z_k)) - \mathrm{Re}(C(z_j, z_k))\right) + O(d_N^{-1}) + O(e_N) \qquad (3.37)\\ &= \sum_{k=1}^{m-1} \alpha_k \left(\phi_0^{(\eta)}(\tau_k) + \phi_0^{(\eta)}(\tau_j) - \phi_0^{(\eta)}(\tau_k - \tau_j)\right) + O(d_N^{-1}) + O(e_N) \qquad (3.38)\end{aligned}$$

To obtain (3.38) from (3.37), we used the formula (B.6) to compute the asymptotics of $\mathrm{Re}(C(z_j,z_k))$ and used that $\alpha_m=-(\alpha_1+\ldots+\alpha_{m-1})$.

Now let us compute the asymptotics of the coefficients β_N, γ_N and χ_{N-1} defined in (3.8) and appearing in the first line of (3.7). As usual, these quantities are all obtained by expanding all z-dependent quantities appearing in (3.28) in powers of $1/z$. Firstly, the Szegö function (3.26) satisfies $\mathcal{D}(z)=\mathcal{D}_\infty(1+\mathcal{D}_1/z+(\mathcal{D}_1^2/2+\mathcal{D}_2)/z^2+O(z^{-3}))$ as $z\to\infty$, where

$$\begin{aligned}\mathcal{D}_1&=-\frac{1}{2}\sum_{k=1}^m\alpha_k\,\mathrm{Re}\left(\frac{1}{z_k+\sqrt{z_k+1}\sqrt{z_k-1}}\right),\\ \mathcal{D}_2&=-\frac{1}{8}\sum_{k=1}^m\alpha_k\,\mathrm{Re}\left(\frac{1}{(z_k+\sqrt{z_k+1}\sqrt{z_k-1})^2}\right),\end{aligned}\tag{3.39}$$

and secondly, use of the definitions (3.25) and (3.13) shows that for $z\to\infty$

$$g(z)=\log(z)-\frac{1}{8z^2}+O(z^{-4}),\qquad a(z)=1-\frac{1}{2z}+\frac{1}{8z^2}+O(z^{-3})\tag{3.40}$$

Then expanding (3.29) at $z=\infty$, we can compare with (3.8) and obtain

$$\begin{aligned}\beta_N&=\sqrt{2N}\left(-\mathcal{D}_1+\frac{A_{11}}{N}+O\left(\frac{1}{Nd_N}\right)+O(e_N)\right)\\ \gamma_N&=2N\left(1/8-N/8+\mathcal{D}_1^2/2-\mathcal{D}_2+\frac{B_{11}-A_{11}\mathcal{D}_1-iA_{12}/2}{N}+O\left(\frac{1}{Nd_N}\right)+O(e_N)\right)\\ \chi_{N-1}^2&=\frac{2^{N-1}}{\sqrt{\pi}(N-1)!}\left(\frac{1}{\mathcal{D}_\infty^2}+\frac{1}{N}\left(\frac{1}{12\mathcal{D}_\infty^2}+2iA_{21}\right)+O\left(\frac{1}{Nd_N}\right)+O(e_N)\right)\end{aligned}$$

A similar computation shows that the asymptotics of χ_N^2 are given by

$$\chi_N^2=\frac{2^N}{\sqrt{\pi}N!}\left(\frac{1}{\tilde{\mathcal{D}}_\infty^2}+\frac{1}{N}\left(\frac{1}{12\tilde{\mathcal{D}}_\infty^2}+2iA_{12}\right)+O\left(\frac{1}{Nd_N}\right)+O(e_N)\right)\tag{3.41}$$

where $\tilde{\mathcal{D}}_\infty$ denotes the quantity (3.27) with rescaled singularities $\tilde{z}_k=\sqrt{2N/(2N+2)}z_k$. This rescaling is necessary when estimating χ_N^2, because without it one obtains asymptotics with respect to the weight $w(x)=\prod_j|x-\sqrt{2N+2}z_k|^{\alpha_k}$. Cumbersome though routine manipulations with the above asymptotics yield

$$\begin{aligned}-N(\log\chi_N\chi_{N-1})'&=2N(C(z_j,\infty)-C(z_m,\infty))+O(d_N^{-1})+O(Ne_N),\\ 2(\gamma_N'-\beta_N\beta_N')&=-4N\mathcal{D}_2'+O(d_N^{-1})+O(Ne_N),\end{aligned}\tag{3.42}$$

and

$$\begin{aligned}(\log\chi_N\chi_{N-1})'Y_{11}(\sqrt{2N}z_k)Y_{22}(\sqrt{2N}z_k)&=O(d_N^{-1})+O(e_N),\\ 2\left(\frac{\chi_{N-1}}{\chi_N}\right)^2\left(\log\frac{\chi_{N-1}}{\chi_N}\right)'&=O(d_N^{-1})+O(e_N),\end{aligned}\tag{3.43}$$

where we introduced

$$C(\mu,\infty)=\lim_{z\to\infty}C(\mu,z)=\frac{1}{2\pi}\int_{-1}^1\frac{\log|x-\mu|}{\sqrt{1-x^2}}\,dx\tag{3.44}$$

$$=\frac{1}{2}\log|z+\sqrt{z+1}\sqrt{z-1}|-\frac{1}{2}\log(2).\tag{3.45}$$

Using the explicit formulae (3.45) and (3.39), we get

$$2(C(z_j,\infty)-C(z_m,\infty))-4\mathcal{D}_2'=\mathrm{Re}(g(z_j))-\mathrm{Re}(g(z_m)) \tag{3.46}$$

where we exploited the convenient identity (see *e.g.* the derivation of Eq. 7.89 in [14])

$$\log|z+\sqrt{z+1}\sqrt{z-1}|+\frac{1}{2}\mathrm{Re}\left(\frac{1}{(z+\sqrt{z+1}\sqrt{z-1})^2}\right)=\mathrm{Re}(g(z)). \tag{3.47}$$

Now inserting (3.42), (3.38) and (3.43) into (3.7), we obtain

$$\begin{aligned}&\frac{\partial}{\partial\alpha_j}\log\varphi_N(\alpha_1,\dots,\alpha_{m-1})=N(\mathrm{Re}(g(z_j))-\mathrm{Re}(g(z_m)))\\&+\sum_{k=1}^{m-1}\alpha_k\left(\phi_0^{(\eta)}(\tau_k)+\phi_0^{(\eta)}(\tau_j)-\phi_0^{(\eta)}(\tau_k-\tau_j)\right)+O(d_N^{-1})+O(Ne_N).\end{aligned} \tag{3.48}$$

Note that the error terms in (3.48) hold *uniformly* in the parameters $(\alpha_k)_{k=1}^{m-1}$ (see Remark 3.3), so that we may integrate both sides of (3.48) according to the procedure discussed in Sect. 5 of [35], arriving at the asymptotics (3.31). □

Proof of Theorems 2.2 and 2.3. Bearing in mind Remark 3.3, we differentiate (3.31) with respect to the parameters $(\alpha_k)_{k=1}^{m-1}$ and evaluate near the origin, leading to

$$\mathbb{E}\{W_N^{(\eta)}(\tau)\}=N(\mathrm{Re}(g(z_k))-\mathrm{Re}(g(z_m)))+O(d_N^{-1})+O(Ne_N) \tag{3.49}$$

$$\mathrm{Cov}\{W_N^{(\eta)}(\tau),W_N^{(\eta)}(\upsilon)\}=\phi_0^{(\eta)}(\tau)+\phi_0^{(\eta)}(\upsilon)-\phi_0^{(\eta)}(\tau-\upsilon)+O(d_N^{-1})+O(Ne_N) \tag{3.50}$$

where the error terms are uniform in τ and υ varying in a compact subset of $\mathbb{R}$. Then defining the centered process $\tilde{W}_N^{(\eta)}(\tau)=W_N^{(\eta)}(\tau)-\mathbb{E}\{W_N^{(\eta)}(\tau)\}$ we immediately find from (3.49) and (3.31) that in the mesoscopic regime (2.7), we have

$$\lim_{N\to\infty}\mathbb{E}\left\{e^{i\sum_{k=1}^m s_k\tilde{W}_N^{(\eta)}(\tau_k)}\right\}=\exp\left(-\frac{1}{2}\sum_{k=1}^m\sum_{j=1}^m s_ks_j\left(\phi_0(\tau_k)+\phi_0(\tau_j)-\phi_0(\tau_k-\tau_j)\right)\right) \tag{3.51}$$

where $(s_k)_{k=1}^m\in\mathbb{R}^m$. Theorem 2.2 follows immediately. To complete the proof of Theorem 2.3, it suffices to note that the error terms in (3.50) are uniform, so that the sequence $(\mathbb{E}\{(\tilde{W}_N(\tau))^2\})_{N=1}^\infty$ is uniformly bounded. □

4. Convergence to white noise in the spectral representation

The main achievement of the previous section was to prove that for any mesoscopic scales of the form (2.7), the process $\tilde{W}_N^{(\eta)}(\tau)$ converges in the sense of finite-dimensional distributions to the regularized fractional Brownian motion $B_0^{(\eta)}(\tau)$. We also proved Theorem 2.3 which extends this convergence to an appropriate function space.

In this section we will study $\tilde{W}_N^{(\eta)}(\tau)$ from a different point of view, namely by means of the Fourier coefficients $b_N(s)$ appearing in the spectral decomposition (2.13). We remind the reader of the definition

$$b_N(s)=\frac{1}{\sqrt{s}}\mathrm{Tr}\left(e^{-isd_N(\mathcal{H}-x_0I)}\right),\qquad s>0. \tag{4.1}$$

A useful and interesting feature of the integral representations (2.13) and its $N\to\infty$ limit (1.9) is that they are suggestive of a corresponding limiting law satisfied by the coefficients $b_N(s)$. Namely, we expect that $b_N(s)$ should 'converge' to the white noise measure $B_c(ds)/\sqrt{2}$. The precise mode of the convergence we consider is described in Theorem 2.4 and it is our goal in this Section to prove this result.

By its very definition, the white noise measure $B_c(ds)$ cannot be understood in a pointwise sense and must be regularized by integrating against a test function. We will consider test functions $\xi \in C_0^\infty(\mathbb{R}_+)$, i.e. ξ is a smooth function with compact support on $\mathbb{R}_+$. Then we have the correspondence

$$c_N(\xi) = \int_0^\infty \xi(s) b_N(s)\, ds = \sum_{j=1}^N f(d_N(x_j - x_0)) =: X_N(f) \tag{4.2}$$

where

$$f(x) = \int_0^\infty \frac{\xi(s)}{\sqrt{s}} e^{-isx}\, ds. \tag{4.3}$$

By our assumptions on ξ, it follows that f belongs to the Schwartz space of rapidly decaying smooth functions, i.e. $f \in S(\mathbb{R})$ where

$$S(\mathbb{R}) = \left\{ f \in C^\infty(\mathbb{R}) : \sup_{x\in\mathbb{R}} \left| x^\gamma \frac{d^\beta f(x)}{dx^\beta} \right| < \infty \quad \gamma, \beta = 0, 1, 2, \ldots \right\}. \tag{4.4}$$

In the following three subsections we will obtain results for the mean, variance and distribution of the random variable (4.2) as $N \to \infty$.

4.1. Mean

We begin by proving that centering is not required in Theorem 2.4.

Proposition 4.1. *On any mesoscopic scales of the form $d_N = N^\alpha$ with any $\alpha \in (0, 1)$, we have*

$$\mathbb{E}\{c_N(\xi)\} = O(d_N^{-1}), \qquad N \to \infty. \tag{4.5}$$

Proof. We write the expectation above as an integral over the normalized density of states $\rho_N(x)$,

$$\mathbb{E}\{c_N(\xi)\} = N \int_{-\infty}^\infty f(d_N(x - x_0))\rho_N(x)\, dx \tag{4.6}$$

where

$$\rho_N(x) = \frac{1}{N} \mathbb{E}\left\{ \sum_{j=1}^N \delta(x - x_j) \right\}. \tag{4.7}$$

Firstly, note that the tails of the integral (4.6) can be removed using the rapid decay of f. For any $\epsilon > 0$, we have

$$\mathbb{E}\{c_N(\xi)\} = N \int_{x_0-\epsilon}^{x_0+\epsilon} f(d_N(x - x_0))\rho_N(x)\, dx + O(N d_N^{-\infty}), \tag{4.8}$$

where here and elsewhere, the notation $O(N d_N^{-\infty})$ refers to a quantity that is $O(N d_N^{-\gamma})$ for any $\gamma > 0$. Such a contribution tends to zero for the power law scales $d_N = N^\alpha$ with any $\alpha \in (0, 1)$. Then for small enough ϵ we have the uniform estimate (see [42], Chapter 5.2)

$$\rho_N(x) = \frac{2}{\pi}\sqrt{1 - x^2} + O(N^{-1}), \qquad x \in (x_0 - \epsilon, x_0 + \epsilon) \tag{4.9}$$

After inserting (4.9) into (4.8) we find that

$$\mathbb{E}\{c_N(\xi)\} = \frac{2N}{\pi} \int_{x_0-\epsilon}^{x_0+\epsilon} f(d_N(x - x_0))\sqrt{1 - x^2}\, dx + E_N + O(N d_N^{-\infty}) \tag{4.10}$$

where the error term $E_N = O(d_N^{-1})$, since

$$|E_N| \leq C\left| \int_{x_0-\epsilon}^{x_0+\epsilon} f(d_N(x-x_0))\, dx \right| \leq \frac{C}{d_N} \int_{-\infty}^{\infty} |f(x)|\, dx. \tag{4.11}$$

Similarly, we can replace the integration limits in (4.10) with ± 1 using the Schwartz property of f. We have

$$\mathbb{E}\{c_N(\xi)\} = \frac{2N}{\pi} \int_{-1}^{1} f(d_N(x-x_0))\sqrt{1-x^2}\, dx + O(d_N^{-1}) \tag{4.12}$$

Next we substitute f with the definition (4.3) and interchange the order of integration (justified by the rapid decay of $\xi(s)$) so that,

$$\begin{aligned}\mathbb{E}\{c_N(\xi)\} &= \frac{2N}{\pi} \int_0^{\infty} \xi(s)s^{-1/2}e^{isd_Nx_0} \int_{-1}^{1} e^{-isd_Nx}\sqrt{1-x^2}\, dx\, ds + O(d_N^{-1}) \\ &= 2N \int_0^{\infty} \xi(s)s^{-3/2}J_1(d_Ns)e^{isd_Nx_0}\, ds + O(d_N^{-1})\end{aligned} \tag{4.13}$$

where $J_1(z)$ is the Bessel function of index 1. To finish the proof, note that $J_1(d_Ns)$ has an asymptotic expansion (for any fixed $\gamma \in \mathbb{N}$ and $s > 0$) as $N \to \infty$,

$$\begin{aligned}\sqrt{\frac{\pi}{2}}J_1(d_Ns) &= \cos(d_Ns - 3\pi/4)\sum_{k=0}^{\gamma-1} \frac{C_k}{d_N^{2k+1/2}s^{2k+1/2}} \\ &\quad + \sin(d_Ns - 3\pi/4)\sum_{k=0}^{\gamma-1} \frac{D_k}{d_N^{2k+3/2}s^{2k+3/2}} + E_N(s)\end{aligned} \tag{4.14}$$

where the error term satisfies the bound $|E_N(s)| \leq |C_\gamma d_N^{-2\gamma-1/2}s^{-2\gamma-1/2}|$ and C_k, D_k are constants depending only on k. Such asymptotics can be found in e.g. [22] or [34].

Inserting (4.14) into (4.13) we see that the contribution from each term in the sum in (4.14) is an oscillatory integral of order $O(Nd_N^{-\infty})$, as follows from repeated integration by parts. The final error term $E_N(s)$ is integrable with respect to $\xi(s)$ and gives rise to an error of order $O(Nd_N^{-2\gamma})$. Since $\gamma > 0$ was arbitrary, we conclude that the term proportional to N in (4.12) is in fact asymptotically smaller than the error term. This completes the proof of the proposition. □

4.2. Covariance

Having studied the expectation of $b_N(s)$ in the previous subsection, we now consider the fluctuations. In the introduction it was remarked, in accordance with the expected white noise limit for $b_N(s)$, that we should have $\lim_{N\to\infty} \mathbb{E}\{b_N(s_1)\overline{b_N(s_2)}\} = \delta(s_1 - s_2)$. In this subsection we will make this assertion precise by proving that

$$\lim_{N\to\infty} \mathbb{E}\{c_N(\xi_1)\overline{c_N(\xi_2)}\} = \int_0^{\infty} \xi_1(s)\overline{\xi_2(s)}\, ds \tag{4.15}$$

for all smooth functions ξ_1, ξ_2 with compact support on $\mathbb{R}_+$.

It turns out that there is an exact finite-N formula for the covariance (see Eq. (4.2.38) in [42]):

$$\mathbb{E}\{\tilde{X}_N(f_1)\tilde{X}_N(f_2)\} = \frac{1}{8}\int_{\mathbb{R}^2} \Delta f_1(d_Nx)\Delta f_2(d_Nx)K_N^2(x_1,x_2)\, dx_1\, dx_2 \tag{4.16}$$

where f_1 and f_2 are defined in terms of ξ_1 and ξ_2 as in formula (4.3) and we introduced the notation $\Delta f(x) = f(x_1) - f(x_2)$ for any f. The function $K_N(x_1, x_2)$ is the kernel of the GUE ensemble (see *e.g.* [41], [42]) having the explicit formula

$$K_N(x,y) = \frac{\psi_N^{(N)}(x_1)\psi_{N-1}^{(N)}(x_2) - \psi_N^{(N)}(x_2)\psi_{N-1}^{(N)}(x_1)}{x_1 - x_2} \tag{4.17}$$

where

$$\psi_l^{(N)}(x) = e^{-Nx^2} P_l^{(N)}(x), \tag{4.18}$$

and $P_l^{(N)}(x)$ are (rescaled) Hermite polynomials, normalized by the condition that $\{\psi_l^{(N)}\}_{l=1}^{\infty}$ forms an orthonormal family on $\mathbb{R}$. By making use of the known Plancherel-Rotach asymptotics for the functions $\psi_l^{(N)}(x)$, we deduce the following covariance formula. After noting the correspondence (4.3), we immediately derive from it the δ-correlations (4.15).

Proposition 4.2. *Let the test functions f_1 and f_2 belong to the Schwartz space $S(\mathbb{R})$ defined in (4.4) and consider the mesoscopic regime $d_N = N^{\alpha}$ with any $\alpha \in (0,1)$. We have*

$$\lim_{N\to\infty} \mathbb{E}\{\tilde{X}_N(f_1)\tilde{X}_N(f_2)\} = \frac{1}{2\pi}\int_{-\infty}^{\infty} |s|\hat{f}_1(s)\hat{f}_2(-s)\,ds. \tag{4.19}$$

where $\hat{f}(s) = (2\pi)^{-1/2}\int_{-\infty}^{\infty} f(x)e^{-isx}\,dx$.

Remark 4.3. Formula (4.19) is already known for C^1 functions with compact support, as in Theorem 5.2.7 (iii) of [42]. It was also proved recently in [18] for a class of Wigner matrices with f a Schwartz test function, but only up to scales $d_N = N^{\alpha}$ with any $0 < \alpha < 1/3$. Our main contribution in this subsection is to adapt the argument given in [42] to our test functions f in (4.3), which cannot be compactly supported due to our assumptions on ξ. We note that our proof holds on the full range $0 < \alpha < 1$ and that the smoothness hypothesis can be relaxed to C^1 functions with rapid decay at $\pm\infty$.

Proof. Here we only consider the contribution to integral (4.16) coming from the square $I_\delta^2 = [-(1-\delta),(1-\delta)]^2$ for some small $\delta > 0$. In Appendix C we will show that the complement of this region can be neglected for small enough δ. We will need the following asymptotic formula for the functions $\psi_{N+k}^{(N)}$ defined in (4.18). Uniformly for $|x| < (1-\delta)$ and $k = O(1)$, we have

$$\psi_{N+k}^{(N)}(x) = \left(\frac{2}{\pi\sqrt{1-x^2}}\right)^{1/2} \cos(N\alpha(x) + (k+1/2)\cos^{-1}(x) - \pi/4) + O(N^{-1}) \tag{4.20}$$

where $\alpha(x) = 2\int_{-1}^{x} dt\,\sqrt{1-t^2}$. Formula (4.20) follows immediately from the classical asymptotic results of Plancherel and Rotach (see Sections 5 in [42] and 8 in [51]).

Now, using the symmetry about the line $x_1 = x_2$, we see that the integral (4.16) restricted to I_δ^2 can be written in the convenient form,

$$\frac{1}{4}\int_{I_\delta^2} \frac{\Delta f_1(d_N x)}{\Delta x}\frac{\Delta f_2(d_N x)}{\Delta x}\mathcal{F}_N(x_1,x_2)\,dx_1\,dx_2 \tag{4.21}$$

where

$$\mathcal{F}_N(x_1,x_2) = \psi_N^{(N)}(x_1)^2\psi_{N-1}^{(N)}(x_2)^2 - \psi_N^{(N)}(x_1)\psi_{N-1}^{(N)}(x_1)\psi_N^{(N)}(x_2)\psi_{N-1}^{(N)}(x_2). \tag{4.22}$$

We insert the Plancherel-Rotach formula (4.20) into (4.21) and denote $\theta(x) = \cos^{-1}(x)$. Using the double angle formula for the cosine, we find that the contribution of (4.20) to the product of squares in (4.22) is

$$\frac{1 + \cos(2N\alpha(x_1) + \theta(x_1)/2 - \pi/4) + \cos(2N\alpha(x_2) - \theta(x_2)/2 - \pi/4)}{\pi^2\sqrt{1-x_2^2}\sqrt{1-x_1^2}} \tag{4.23}$$

$$+\frac{\cos(2N\alpha(x_1) + \theta(x_1)/2 - \pi/4)\cos(2N\alpha(x_2) - \theta(x_2)/2 - \pi/4)}{\pi^2\sqrt{1-x_2^2}\sqrt{1-x_1^2}} + O(N^{-1}). \tag{4.24}$$

Inserting the oscillatory terms in lines (4.23) and (4.24) into (4.21) gives rise to error terms that are $O((N/d_N)^{-\infty})$ as $N \to \infty$ for every $\delta > 0$. This can be shown by repeated integration by parts, using the fact that $\alpha(x)$ is smooth and increasing on the interval I_δ. Combined with a similar calculation applied to the second term in (4.22), we see that the integral (4.21) is equal to

$$\begin{aligned}&\frac{1}{4\pi^2}\int_{I_\delta^2}\frac{\Delta f_1(d_N x)}{\Delta x}\frac{\Delta f_2(d_N x)}{\Delta x}\frac{1-x_1x_2}{\sqrt{1-x_1^2}\sqrt{1-x_2^2}}\,dx_1\,dx_2+O((N/d_N)^{-\infty})=\\&\frac{1}{4\pi^2}\int_{\mathbb{R}^2}\frac{\Delta f_1(x)}{\Delta x}\frac{\Delta f_2(x)}{\Delta x}\frac{1-x_1x_2/d_N^2}{\sqrt{1-x_1^2/d_N^2}\sqrt{1-x_2^2/d_N^2}}\chi_{I_N}(x_1)\chi_{I_N}(x_2)\,dx_1\,dx_2+O((N/d_N)^{-\infty})\end{aligned}\tag{4.25}$$

where $\chi_{I_N}(x_1)$ is the indicator function on the set $I_N = (-(1-\delta)d_N, (1-\delta)d_N)$.

Now Lebesgue's dominated convergence theorem can be applied to take the limit under the integral in (4.25). Indeed, it is easy to see that the integrand in (4.25) is bounded by the integrable function

$$\left(\frac{2}{\delta^2}-1\right)\left|\frac{\Delta f_1(x)}{\Delta x}\right|\left|\frac{\Delta f_2(x)}{\Delta x}\right|\tag{4.26}$$

for any $N \in \mathbb{N}$, $(x_1, x_2) \in \mathbb{R}^2$ and $0 < \delta < 1$. We finally see that for all $0 < \delta < 1$, we have

$$\lim_{N\to\infty}\frac{1}{4}\int_{I_\delta^2}\frac{\Delta f_1(d_N x)}{\Delta x}\frac{\Delta f_2(d_N x)}{\Delta x}\mathcal{F}_N(x_1,x_2)\,dx_1\,dx_2=\frac{1}{4\pi^2}\int_{\mathbb{R}^2}\frac{\Delta f_1(x)}{\Delta x}\frac{\Delta f_2(x)}{\Delta x}\,dx_1\,dx_2.\tag{4.27}$$

Rewriting f_1 and f_2 in terms of their Fourier transforms and applying the Plancherel theorem gives the identity

$$\frac{1}{4\pi^2}\int_{\mathbb{R}^2}\frac{f_1(x_1)-f_1(x_2)}{x_1-x_2}\frac{f_2(x_1)-f_2(x_2)}{x_1-x_2}\,dx_1\,dx_2=\frac{1}{2\pi}\int_{\mathbb{R}}|s|\hat{f}_1(s)\hat{f}_2(-s)\,ds,\tag{4.28}$$

which is precisely the right-hand side of (4.19). To complete the proof, we just need to show that the integral (4.16) restricted to the complement of the square I_δ^2 can be neglected in the limit $N \to \infty$. Namely, we prove in the Appendix that

$$\lim_{N\to\infty}\int_{(I_\delta^2)^c}\Delta f_1(d_N x)\Delta f_2(d_N x)K_N^2(x_1,x_2)\,dx_1\,dx_2=O(\delta),\qquad \delta\to 0,\tag{4.29}$$

and so complete the proof of the Proposition by choosing $\delta > 0$ sufficiently small. □

4.3. Convergence in distribution

The aim of this subsection is to study the full distribution of the coefficients $b_N(s)$ and ultimately to prove Theorem 2.4. First we need a preliminary result regarding the stochastic process $\tilde{W}_N^{(\eta)}(\tau)$. It will be convenient to consider the *increments*

$$\begin{aligned}\Delta_p[\tilde{W}_N^{(\eta)}](\tau):&=\tilde{W}_N^{(\eta)}(\tau)-\tilde{W}_N^{(\eta)}(\tau+p)\\&=\frac{1}{2}\int_0^\infty\frac{e^{-\eta s}}{\sqrt{s}}\left\{[1-e^{-ips}]e^{-i\tau s}\tilde{b}_N(s)+[1-e^{ips}]e^{i\tau s}\overline{\tilde{b}_N(s)}\right\}ds,\end{aligned}\tag{4.30}$$

where $\tilde{b}_N(s) = b_N(s) - \mathbb{E}\{b_N(s)\}$.

Similarly, the corresponding limiting object is given by the following stationary Gaussian process

$$\begin{aligned}\Delta_p[B_0^{(\eta)}](\tau):&=B_0^{(\eta)}(\tau)-B_0^{(\eta)}(\tau+p)\\&=\frac{1}{2\sqrt{2}}\int_0^\infty\frac{e^{-\eta s}}{\sqrt{s}}\left\{[1-e^{-ips}]e^{-i\tau s}B_c(ds)+[1-e^{ips}]e^{i\tau s}\overline{B_c(ds)}\right\}.\end{aligned}\tag{4.31}$$

Proposition 4.4. *Let $p \in \mathbb{R}$. For any $h \in S(\mathbb{R})$ and on any power law scales $d_N = N^\alpha$ with $\alpha \in (0,1)$, we have the convergence in distribution*

$$\int_{-\infty}^{\infty} h(\tau)\Delta_p[\tilde{W}_N^{(\eta)}](\tau)\,d\tau \overset{d}{\Longrightarrow} \int_{-\infty}^{\infty} h(\tau)\Delta_p[B_0^{(\eta)}](\tau)\,d\tau, \qquad N \to \infty. \tag{4.32}$$

Proof. The proof will be analogous to our proof of Theorem 2.3, the main difference being we must have good enough control of the tails in the above integrals. This will be taken care of by the rapid decay of h. To proceed, we fix some (arbitrary) $M \in \mathbb{R}$ and $\delta_0 > 0$ and decompose the left-hand side of (4.32) as

$$\begin{aligned}&\int_{-M}^{M} h(\tau)\Delta_p[\tilde{W}_N^{(\eta)}](\tau)\,d\tau \\ &+\int_{|\tau|\in[M,\delta_0 d_N]} h(\tau)\Delta_p[\tilde{W}_N^{(\eta)}](\tau)\,d\tau + \int_{|\tau|\in[\delta_0 d_N,\infty)} h(\tau)\Delta_p[\tilde{W}_N^{(\eta)}](\tau)\,d\tau\end{aligned} \tag{4.33}$$

and label each of the integrals in (4.33) with $\mathcal{I}_1, \mathcal{I}_2$ and $\mathcal{I}_3$. Let us begin with the first integral, $\mathcal{I}_1$. By Theorem 2.2 and the Cramér-Wold device, the finite-dimensional distributions of $\Delta_p[\tilde{W}_N^{(\eta)}](\tau)$ converge in law to those of $\Delta_p[B_0^{(\eta)}](\tau)$. Furthermore, by the uniform estimate (3.50) we have that there is a constant $C > 0$ such that $\mathbb{E}\{(\Delta_p[B_0^{(\eta)}(\tau)])^2\} \leq C$ for all $\tau \in [-M, M]$ and for all N. Therefore the hypotheses of Theorem 3 in [28] are satisfied and we conclude that the first integral in (4.33) converges in distribution to the right-hand side of (4.32) in the limit $N \to \infty$ followed by $M \to \infty$. To complete the proof, it suffices to show that the second and third integrals in (4.33) converge in probability to 0 in the same limit.

For notational convenience we just consider the contributions to $\mathcal{I}_2$ and $\mathcal{I}_3$ where $\tau > 0$ as the situation $\tau < 0$ is almost identical. By Chebyshev's inequality and Cauchy-Schwarz, we have

$$\mathbb{P}\{|\mathcal{I}_2| > \epsilon\} \leq \epsilon^{-2} \int_{M}^{\delta_0 d_N} |h(\tau)|d\tau \int_{M}^{\delta_0 d_N} |h(\tau)|\,\mathbb{E}\{\Delta_p[\tilde{W}_N^{(\eta)}](\tau)^2\}\,d\tau \tag{4.34}$$

We will now argue that the variance term in (4.34) is uniformly bounded. Since $|\tau| \leq \delta_0 d_N$, by choosing δ_0 small enough we see that $|x_0 + \tau/d_N| < 1 - \delta$ for some $\delta > 0$ independent of N. Hence the singularities of the logarithm in (2.6) remain inside the bulk region $(-1+\delta, 1-\delta)$ for all N and we may apply the methods of Section 3 with $m = 2$ and weight (cf. (3.19))

$$\omega(z) = \left[\frac{(z - x_0(\tau, N) - p/d_N)^2 + (\eta/d_N)^2}{(z - x_0(\tau, N))^2 + (\eta/d_N)^2}\right]^{\alpha/2}, \qquad x_0(\tau, N) = x_0 + \tau/d_N. \tag{4.35}$$

The only difference in the analysis of the Riemann-Hilbert problem with this weight is that the new reference point $x_0(\tau, N)$ can vary with N in the small fixed neighbourhood $[x_0 - \delta_0, x_0 + \delta_0]$. However, all the estimates we obtain are uniform for x_0 varying in compact subsets of $(-1+\delta, 1-\delta)$ so that the variance bound (3.50) (with $\upsilon = \tau$) remains valid. This implies that for some N-indepedent $C > 0$,

$$\mathbb{P}\{|\mathcal{I}_2| > \epsilon\} \leq \epsilon^{-2} C \left(\int_{M}^{\delta_0 d_N} |h(\tau)|\,d\tau\right)^2 \to 0, \tag{4.36}$$

in the limit $N \to \infty$ followed by $M \to \infty$.

To bound the integral $\mathcal{I}_3$ we again apply Chebyshev's inequality and exploit the rapid decay of h. We have

$$\mathbb{P}\{|\mathcal{I}_3| > \epsilon\} \leq \epsilon^{-2} \int_{\delta_0 d_N}^{\infty}\int_{\delta_0 d_N}^{\infty} \mathbb{E}\{h(\tau_1)\Delta_p[\tilde{W}_N^{(\eta)}](\tau_1)\overline{h(\tau_2)}\Delta_p[\tilde{W}_N^{(\eta)}](\tau_2)\}\,d\tau_1\,d\tau_2 \tag{4.37}$$

$$= \epsilon^{-2} \int_{\delta_0 d_N}^{\infty}\int_{\delta_0 d_N}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(\tau_1)\overline{h(\tau_2)} \prod_{j=1}^{2}(q(x_1,\tau_j) - q(x_2,\tau_j))\,K_N^2(x_1,x_2)\,dx_1dx_2d\tau_1d\tau_2 \tag{4.38}$$

where we computed the expectation using the identity (4.16) and

$$q(x,\tau) = -\log\left|x - x_0 - \frac{\tau + i\eta}{d_N}\right| + \log\left|x - x_0 - \frac{\tau + p + i\eta}{d_N}\right|. \tag{4.39}$$

Now, since h is a Schwartz test function, we know that for any $\gamma > 0$ and $u > 0$, we have $|h(ud_N)| \leq (d_N u)^{-\gamma}$ for N large enough. Then using the inequalities $|q(x,\tau)| \leq C_{p,\eta}$ for some finite constant depending only on p and η, $K_N^2(x_1,x_2) \leq N^2\rho_N(x_1)\rho_N(x_2)$ and substituting $\tau_j = ud_N$ we obtain

$$\mathbb{P}(|\mathcal{I}_3| > \epsilon) \leq 4\epsilon^{-2}C_{p,\eta}^2 N^2 d_N^{-2\gamma+2}\left(\int_{\delta_0}^{\infty} u^{-\gamma}du\right)^2. \tag{4.40}$$

Then provided d_N takes the form $d_N = N^{\alpha}$ with $\alpha \in (0,1)$ we can always choose $\gamma > 0$ large enough such that the right-hand side of (4.40) tends to 0 as $N \to \infty$. $\square$

We can now translate the result (4.32) into a statement about the Fourier coefficients $b_N(s)$, allowing us to prove Theorem 2.4. For the convenience of the reader, we repeat the statement of the latter result here.

Theorem 4.5. *Let $\xi_1,\ldots,\xi_m$ be smooth functions compactly supported on $\mathbb{R}_+$. Then the vector $(c_N(\xi_1),\ldots,c_N(\xi_m))$ converges in distribution to a centered complex Gaussian vector Z with relation matrix $C = \mathbb{E}\{ZZ^{\mathrm{T}}\} = 0$ and covariance matrix $\Gamma = \mathbb{E}\{ZZ^{\dagger}\}$ given by*

$$\Gamma_{j,k} = \int_0^{\infty} \xi_j(s)\overline{\xi_k(s)}\,ds, \qquad j,k = 1,\ldots,m. \tag{4.41}$$

Proof. Define functions h_k in terms of their Fourier transform as

$$\int_{-\infty}^{\infty} h_k(\tau)e^{-i\tau s}\,d\tau = \frac{\sqrt{s}}{1 - e^{-ips}}e^{\eta s}\xi_k(s) \qquad k = 1,\ldots,m. \tag{4.42}$$

Then for sufficiently small p, the right-hand side of (4.42) is smooth and compactly supported. Therefore, its Fourier transform h_k is a Schwartz function, i.e. $h_k \in S(\mathbb{R})$. Next, note that with $c_N(\xi)$ as in (4.2), we have the identity

$$c_N(\xi_k) - \mathbb{E}(c_N(\xi_k)) = 2\int_{-\infty}^{\infty} h_k(\tau)\Delta_p[\tilde{W}_N^{(\eta)}](\tau)\,d\tau \tag{4.43}$$

which holds almost surely and follows after inserting the representation (4.30) and interchanging the order of integration, justified by the rapid decay of ξ_k and h_k. Now we apply Proposition 4.4 with $h(\tau) = \sum_{k=1}^m \alpha_k h_k(\tau)$ where $\alpha_k \in \mathbb{C}$. Since $\mathbb{E}(c_N(\xi_k)) = O(d_N^{-1})$, we get the convergence in distribution

$$\sum_{k=1}^m \alpha_k c_N(\xi_k) \overset{d}{\Longrightarrow} 2\sum_{k=1}^m \alpha_k \int_{-\infty}^{\infty} h_k(\tau)\Delta_p[B_0^{(\eta)}](\tau)\,d\tau, \qquad N \to \infty. \tag{4.44}$$

By the Cramér-Wold device, this implies the convergence in distribution

$$(c_N(\xi_1),\ldots,c_N(\xi_k)) \overset{d}{\Longrightarrow} (Z(h_1),\ldots,Z(h_m)) \tag{4.45}$$

where

$$Z(h_k) = 2\int_{-\infty}^{\infty} h_k(\tau)\Delta_p[B_0^{(\eta)}](\tau)\,d\tau. \tag{4.46}$$

Since $\Delta_p[B_0^{(\eta)}](\tau)$ is a Gaussian process, one easily sees that $(Z(h_1),\ldots,Z(h_m))$ is a mean zero complex Gaussian vector. Then by a simple computation using the integral representation (4.31) and basic properties of the white noise measure $B_c(ds)$, we find the covariance structure

$$\Gamma_{j,k} = \mathbb{E}\{Z(h_j)\overline{Z(h_k)}\} = \int_0^{\infty} \xi_j(s)\overline{\xi_k(s)}\,ds, \tag{4.47}$$

and $C_{j,k} = \mathbb{E}\{Z(h_j)Z(h_k)\} = 0$ for all $j,k = 1,\ldots,m$. $\square$

5. Macroscopic regime

The main goal of this section is to prove Theorem 2.1. Namely, we will show that the process $\tilde{D}_N(x)$ (2.5) converges in probability law as $N \to \infty$ to the generalized Gaussian process $F(x)$ given by (1.7). The convergence is interpreted in the Sobolev space $V^{(-a)}$, *i.e.* the assertion of Theorem 2.1 is that for any bounded continuous functional q on $V^{(-a)}$, we have

$$\lim_{N\to\infty} \mathbb{E}\{q(\tilde{D}_N)\} = \mathbb{E}\{q(F)\}. \tag{5.1}$$

Our proof is an adaptation for the GUE matrices $\mathcal{H}$ of the proof of a similar result for the CUE matrices given in [30]. First, we will prove that the finite-dimensional distributions of $\tilde{D}_N(x)$ converge to those of $F(x)$ and then establish that the sequence $\tilde{D}_N(x)$ is tight in $V^{(-a)}$. This will imply the convergence in probability law in $V^{(-a)}$ as in (5.1). As explained in section 2.1, for the GUE matrices there are additional analytical complications compared with the case of CUE matrices.

We start with a deterministic result, writing down the Chebyshev-Fourier series for $\tilde{D}_N(x)$.

Lemma 5.1. *Let $\mathcal{H}$ be a Hermitian matrix of size $N \times N$ with eigenvalues $x_1, \dots, x_N$. Then*

$$-\log|\det(\mathcal{H} - xI)| = N\log 2 + \sum_{k=0}^{\infty} c_k(D_N)T_k(x)$$

where the convergence is pointwise for any $x \in [-1,1]\backslash\{x_1,\dots,x_N\}$ and the Chebyshev-Fourier coefficients $c_k(D_N)$ are given for any $k > 0$ by the formula

$$c_k(D_N) = \sum_{j=1}^{N} \frac{2}{k}T_k(x_j) + \sum_{j=1}^{N} r_k^+(x_j) + \sum_{j=1}^{N} r_k^-(x_j) \tag{5.2}$$

and

$$c_0(D_N) = -\sum_{j=1}^{N} r_0^+(x_j) - \sum_{j=1}^{N} r_0^-(x_j) \tag{5.3}$$

where for $k > 0$

$$r_k^{\pm}(x) = \left[(2/k)(-T_k(x) + (x \mp \sqrt{x^2-1})^k\right]\chi_{(\pm 1,\pm\infty)}(x) \tag{5.4}$$

and

$$r_0^{\pm}(x) = \log|x \mp \sqrt{x^2-1}|\chi_{(\pm 1,\pm\infty)}(x) \tag{5.5}$$

In the above formulae, $\chi_J(x)$ is the indicator function on the set J.

Proof. This follows immediately from Lemma 3.1 in [27]. □

It follows from this Lemma that for our random matrices $\mathcal{H}$, with probability one,

$$\tilde{D}_N(x) = \sum_{k=0}^{\infty} c_k(\tilde{D}_N)T_k(x), \quad \text{where} \quad c_k(\tilde{D}_N) = c_k(D_N) - \mathbb{E}\{c_k(D_N)\}.$$

5.1. Convergence of finite-dimensional distributions

The main goal of this subsection is to establish the following:

Proposition 5.2. *Fix $M \in \mathbb{N}$ and let $X_1, \dots, X_M$ be independent Gaussian random variables with mean zero and variance one. Then for any $(t_k)_{k=1}^M \in \mathbb{R}^M$ we have the convergence in distribution*

$$\sum_{k=0}^{M} c_k(\tilde{D}_N)t_k \overset{d}{\Longrightarrow} \sum_{k=1}^{M} \frac{X_k}{\sqrt{k}}t_k, \qquad N \to \infty. \tag{5.6}$$

Proof. We begin by inserting Eq. (5.2) into the left-hand side of (5.6). Then from [31] or [42], we know that the sum

$$\sum_{k=1}^{M} t_k \left(\sum_{j=1}^{N} \frac{2}{k} T_k(x_j) - \mathbb{E}\left\{ \sum_{j=1}^{N} \frac{2}{k} T_k(x_j) \right\} \right) \tag{5.7}$$

converges in distribution to the right-hand side of (5.6) as $N \to \infty$. The main technical part of our proof of (5.6) consists in showing that the other terms appearing in (5.2) and (5.3) do not contribute in the limit $N \to \infty$. All such terms that appear are of the form

$$A^{\pm}_{k,N} = \sum_{j=1}^{N} r^{\pm}_k(x_j) \tag{5.8}$$

and by definition of the test function $r^{\pm}_k(x)$, they are non-zero only when an eigenvalue x_j lies outside the bulk of the limiting spectrum $[-1, 1]$. Intuitively this is a rare event and we show below that in fact $\mathbb{E}|A^{\pm}_{k,N}| \to 0$ as $N \to \infty$. We note in passing that the regularity of the test functions $r^{\pm}_k(x)$ lies outside the best known $C^{1/2+\epsilon}$ threshold in [50], due to the singularities at the spectral edges.

Let us focus our attention on the case $\mathbb{E}\{|A^{+}_{k,N}|\}$, since the estimation of $\mathbb{E}\{|A^{-}_{k,N}|\}$ follows exactly the same pattern. First, one sees from the explicit formula (5.4) and the elementary inequality $(x - \sqrt{x^2-1})^k \leq T_k(x) \leq (x + \sqrt{x^2-1})^k$, $x \geq 1$ that $-r^{+}_k(x)$ is non-negative for all $x \in \mathbb{R}$. Therefore $\mathbb{E}\{|A^{+}_{k,N}|\} = -\mathbb{E}\{A^{+}_{k,N}\}$.

In terms of the normalized eigenvalue density, we have

$$\mathbb{E}\{A^{+}_{k,N}\} = N \int_1^{\infty} r^{+}_k(x) \rho_N(x) dx. \tag{5.9}$$

To proceed, we split the integral as

$$\mathbb{E}\{A^{+}_{k,N}\} = N \int_1^{1+\delta_N} r^{+}_k(x) \rho_N(x)\, dx + N \int_{1+\delta_N}^{\infty} r^{+}_k(x) \rho_N(x)\, dx \tag{5.10}$$

where we choose $\delta_N = N^{-7/12}$. The first integral in (5.10) is over a shrinking neighbourhood of the spectral edge $x = 1$. An estimate that holds uniformly in this region can be given in terms of the Airy function $\mathrm{Ai}(x)$ and its derivatives. In particular, Eq. 4.4 of [17] (see also the *Proof of Lemma 2.2* in [29]) shows that as $N \to \infty$

$$\begin{aligned} N\rho_N(x) = &\left(\frac{\Phi'(x)}{4\Phi(x)} - \frac{\gamma'(x)}{\gamma(x)} \right) [2\mathrm{Ai}(N^{2/3}\Phi(x))\mathrm{Ai}'(N^{2/3}\Phi(x))] \\ &+ N^{2/3}\Phi'(x)[(\mathrm{Ai}'(N^{2/3}\Phi(x)))^2 - N^{2/3}\Phi(x)(\mathrm{Ai}(N^{2/3}\Phi(x)))^2] + O\left(\frac{1}{N(\sqrt{x-1})} \right) \end{aligned} \tag{5.11}$$

where

$$\gamma(x) = \left(\frac{x-1}{x+1} \right)^{1/4} \tag{5.12}$$

and

$$\Phi(x) = \begin{cases} -\left(3\int_x^1 \sqrt{1-y^2}\, dy\right)^{2/3}, & |x| \leq 1 \\ \left(3\int_1^x \sqrt{y^2-1}\, dy\right)^{2/3}, & |x| > 1 \end{cases} \tag{5.13}$$

Since $\Phi(x) \geq 0$ for $x \geq 1$, the functions $\mathrm{Ai}(N^{2/3}\Phi(x))$ and $\mathrm{Ai}'(N^{2/3}\Phi(x))$ are uniformly bounded on $[1, \infty)$. Furthermore, $\left(\frac{\Phi'(x)}{4\Phi(x)} - \frac{\gamma'(x)}{\gamma(x)} \right)$ and $\Phi'(x)$ are bounded near $x = 1$. Inserting (5.11) into the first integral in (5.10), we obtain the bound

$$N \int_1^{1+\delta_N} r^{+}_k(x) \rho_N(x)\, dx = c_1 N^{2/3} \int_1^{1+\delta_N} r^{+}_k(x)\, dx + O\left(\frac{1}{N} \right), \tag{5.14}$$

where c_1 is an N-independent constant. In (5.14) we used that $r_k^+(x)(x-1)^{-1/2}$ is bounded near $x = 1$ to estimate the contribution of the error term in (5.11). A simple computation shows that $\int_1^{1+\delta_N} r_k^+(x)\,dx = O(\delta_N^{3/2})$ as $N \to \infty$ for $k \geq 0$. Inserting the latter into (5.14) yields the bound

$$N \int_1^{1+\delta_N} r_k^+(x)\rho_N(x)\,dx = O(N^{2/3}\delta_N^{3/2}) = O(N^{-5/24}). \tag{5.15}$$

Now consider the second integral in (5.10). We will prove below that it is exponentially small as $N \to \infty$. Using the fact that (for $k \geq 1$) $-r_k^+(x) \leq T_k(x)$ and applying Lemma C.1, we obtain

$$-N \int_{1+\delta_N}^{\infty} r_k^+(x)\rho_N(x)\,dx \tag{5.16}$$

$$\leq N\delta_N \int_1^{\infty} T_k(1+u\delta_N)\rho_N(1+u\delta_N)\,du \tag{5.17}$$

$$\leq B^{-1} \int_1^{\infty} u^{-1} T_k(1+u\delta_N) e^{-buN^{1/8}}\,du \tag{5.18}$$

where $B, b > 0$ are absolute constants. Then *e.g.* expanding $T_k(1+u\delta_N)$ in powers of $(u\delta_N)$ and integrating (5.18) term by term, we can apply the standard Laplace method and find that (5.18) is $O(e^{-cN^{1/8}})$ for some $c > 0$. If $k = 0$ in the integral (5.16), one can use the inequality $|r_0^+(1+x)| \leq \sqrt{2x}$, $x > 0$ and then apply the Laplace method as before yielding a similar error bound. This completes the proof of the Proposition. □

5.2. Tightness

The final ingredient required for proving the weak convergence in (5.1) is to show that the sequence $\tilde{D}_N$ is tight in $V^{(-a)}$. In direct analogy to the proof given in Theorem 2.5 of [30] for the Circular Unitary Ensemble, we will exploit the convenient fact that for $-\infty < a < b < \infty$, the closed unit ball in $V^{(b)}$ is compact in $V^{(a)}$. Then by Chebyshev's inequality, tightness follows if we can bound the variance

$$\mathbb{E}\|\tilde{D}_N\|_{(-b)}^2 = \sum_{k=0}^{\infty} \mathbb{E}\{c_k(\tilde{D}_N)^2\}(1+k^2)^{-b} \tag{5.19}$$

uniformly in N. Such a uniform bound will follow for any $b > 1/2$ provided we show that $\mathbb{E}\{c_k(\tilde{D}_N)^2)\} \leq C$ for some constant C independent of k and N. We begin by writing the Chebyshev-Fourier coefficient as

$$c_k(\tilde{D}_N) = \sum_{j=1}^{N} h_k(x_j) - \mathbb{E}\left\{\sum_{j=1}^{N} h_k(x_j)\right\} \tag{5.20}$$

where

$$\begin{aligned} h_k(x) = (2/k)T_k(x)\chi_{[-1,1]}(x) - (2/k)(x - \sqrt{x^2-1})^k \chi_{(1,\infty)}(x) \\ - (2/k)(x + \sqrt{x^2-1})^k \chi_{(-1,-\infty)}(x) \end{aligned} \tag{5.21}$$

Then by formula (4.16) we have

$$\mathbb{E}\{c_k(\tilde{D}_N)^2\} = \frac{1}{8}\int_{\mathbb{R}^2} (h_k(x_1) - h_k(x_2))^2 K_N(x_1,x_2)^2\,dx_1\,dx_2 \tag{5.22}$$

where $K_N(x,y)$ is the GUE kernel defined in Eq. (4.17).

First we consider the contribution to the integral (5.22) coming from the region $[-1,1]^2$, namely the integral

$$\frac{1}{2k^2}\int_{[-1,1]^2}\left(\frac{\Delta T_k(x)}{\Delta x}\right)^2 \mathcal{F}_N(x_1,x_2)\,dx_1\,dx_2 \tag{5.23}$$

where $\mathcal{F}_N(x_1,x_2)$ is defined by (4.22) and, as in Section 4, for a function f, we denote by Δf the difference $\Delta f(x) = f(x_1) - f(x_2)$. By the Plancherel-Rotach asymptotics of Hermite polynomials, we have the bound (as follows from *e.g.* parts (iii) and (v) of Theorem 2.2 in [13])

$$|\mathcal{F}_N(x_1,x_2)| \leq \frac{K_1}{\sqrt{1-x_1^2}\sqrt{1-x_2^2}} \tag{5.24}$$

uniformly for $(x_1,x_2)\in[-1,1]^2$. This implies that the modulus of (5.23) is bounded by

$$\frac{K_1}{2k^2}\int_{[-1,1]^2}\left(\frac{\Delta T_k(x)}{\Delta x}\right)^2 \frac{1}{\sqrt{1-x_1^2}\sqrt{1-x_2^2}}\,dx_1\,dx_2 = K_1\pi^2/8. \tag{5.25}$$

The equality in (5.25) is a simple exercise involving standard properties of Chebyshev polynomials and we omit the derivation.

Finally consider the contribution to the integral (5.22) from outside the square $[-1,1]^2$. For simplicity, consider just the region $1 < x_1 < \infty$ and $-1 < x_2 < 1$, all others being analogous. Since $h_k(x)$ is uniformly bounded in k and x on the whole real line, we have

$$\int_{-1}^{1}\int_{1}^{\infty}(h_k(x_1)-h_k(x_2))^2 K_N(x_1,x_2)^2\,dx_1\,dx_2 \tag{5.26}$$

$$\leq \int_{-\infty}^{\infty}\int_{1}^{\infty} K_N(x_1,x_2)^2\,dx_1\,dx_2 \tag{5.27}$$

$$= \int_{1}^{\infty} N\rho_N(x_1)\,dx_1 = \int_{1}^{1+\delta} N\rho_N(x_1)\,dx_1 + O(Ne^{-c_\delta N}) \tag{5.28}$$

where $\delta > 0$ is a constant and $c_\delta > 0$. The last equality in (5.28) follows from Theorem 5.2.3 (iii) in [42]. Now we can insert the formula (5.11) which holds uniformly on $[1, 1+\delta]$. The first term in (5.11) is bounded in N and x_1 and so its integral over $[1, 1+\delta]$ is bounded in N. The third term gives an error of order $1/N$. The contribution from the middle term can be explicitly integrated using the substitution $u = N^{2/3}\Phi(x_2)$:

$$\int_{1}^{1+\delta} N^{2/3}\Phi'(x_2)\left(\mathrm{Ai}'^2(N^{2/3}\Phi(x_2)) - N^{2/3}\Phi(x_2)\mathrm{Ai}^2(N^{2/3}\Phi(x_2))\right)\,dx_2 \tag{5.29}$$

$$= \int_{0}^{N^{2/3}\Phi(1+\delta)}[\mathrm{Ai}'^2(u) - u\mathrm{Ai}^2(u)]\,du \tag{5.30}$$

$$= -\left[\frac{2}{3}(u^2\mathrm{Ai}^2(u) - u\mathrm{Ai}'^2(u)) - \frac{1}{3}\mathrm{Ai}(u)\mathrm{Ai}'(u)\right]_0^{N^{2/3}\Phi(1+\delta)} \tag{5.31}$$

$$= \mathrm{Ai}(0)\mathrm{Ai}'(0)/3 + O(e^{-d_\delta N}) \tag{5.32}$$

where $d_\delta > 0$. A completely analogous argument proves that the integral over the region $\{1 < x_1 < \infty, 1 < x_2 < \infty\}$ is also uniformly bounded in k and N, in addition to the remaining 6 regions that make up B^c. This completes the proof that $\tilde{D}_N$ is tight in $V^{(-a)}$ for any $a > 1/2$ and hence completes the proof of Theorem 2.1.

Appendix A: Proof of Proposition 3.2

The purpose of this Appendix is to give the technical details required to show that the matrix $P_\infty(z)$ in Sec. 3.4 gives a good approximation to the matrix $S(z)$ in Sec. 3.3 for large N, as described by Proposition 3.2. Although we can mostly follow the now standard techniques described in [13], we must take special care with the estimates because the system of contours in Figure 1 can come arbitrarily close to the real axis as $N\to\infty$.

Remark A.1. In this Appendix there are many estimates holding uniformly in the parameters $\{\tau_k\}_{k=1}^{m-1}$, $\{\alpha_k\}_{k=1}^{m-1}$ and x_0 that appear in the partition function (3.2). We will use the big-oh notation $\mathcal{O}$ (distinguished from the usual O) for an error term that defines an analytic function of the parameters $\{\alpha_k\}_{k=1}^{m-1}$ on Ω (cf. (3.9)) satisfying uniformity in the following parameters

- τ_k varying in a compact subset of $\mathbb{R}$ for $k=1,\ldots,m-1$,
- α_k varying in a compact subset of Ω for $k=1,\ldots,m-1$,
- x_0 varying in a compact subset of $(-1+\delta,1-\delta)$.

Construction of the Parametrices at $z=\pm1$

The parametrices at $z=\pm1$ consist of a matrix valued function $P_{\pm1}(z)$ defined in the discs $B_{\pm1}(\delta)$ (cf. Figure 1) satisfying the following properties:

1. $P_{\pm1}(z)$ is analytic in $B_{\pm1}(\delta)\setminus\Sigma$.
2. $P_{\pm1}(z)$ satisfies the same jump conditions as $S(z)$ on $\Sigma\cap B_{\pm\delta}$.
3. The following matching condition is satisfied on the boundary $\partial B_{\pm1}(\delta)$
$$P_{\pm1}(z)P_\infty(z)^{-1}=I+O(N^{-1}),\qquad z\in\partial B_{\pm1}(\delta), \tag{A.1}$$
as $N\to\infty$.

The functions $P_1(z)$ and $P_{-1}(z)$ can be obtained in precisely the same way as in [35], which was itself based on the construction in [13] corresponding to weights $\omega(z)\equiv1$. In our situation, the only difference is that our weight $\omega(z)$ and the Szegö function $\mathcal{D}(z)$ are N-dependent, so that one has to be careful with the matching condition (A.1). From Eq. (76) in [35], we have

$$P_{\pm1}(z)P_\infty(z)^{-1}=P_\infty(z)\omega(z)^{\sigma_3/2}\tilde{P}_\infty(z)^{-1}\tilde{P}_{\pm1}(z)\tilde{P}_\infty(z)^{-1}\tilde{P}_\infty(z)\omega(z)^{-\sigma_3/2}P_\infty(z)^{-1}, \tag{A.2}$$

where $\tilde{P}_{\pm1}(z)$ and $\tilde{P}_\infty(z)$ are the quantities $P_{\pm1}(z)$ and $P_\infty(z)$ with $\omega(z)\equiv1$. For our purposes we will not need the explicit expression for $\tilde{P}_{\pm1}(z)$, which can be found in *e.g.* [13] or [35]. Our main goal here is to check that the matching condition (A.1) is still satisfied.

Lemma A.2. *Let $P_{\pm1}(z)$ denote the parametrix defined in (A.2). Then we have as $N\to\infty$*

$$P_{\pm1}(z)P_\infty(z)^{-1}=I+\frac{\tilde{\Delta}_1^{(\pm1)}(z)}{N}+\mathcal{O}\left(\frac{1}{Nd_N}\right),\qquad z\in\partial B_{\pm1}(\delta) \tag{A.3}$$

where the estimate is uniform for $z\in\partial B_{\pm1}(\delta)$. The first correction term $\tilde{\Delta}_1^{(\pm1)}(z)$ depends only on z and is analytic except for a second order pole at $z=\pm1$.

Proof. Prop. 7.7 of [13] implies that there is a uniform asymptotic expansion

$$\tilde{P}_{\pm1}(z)\tilde{P}_\infty(z)^{-1}\sim I+\sum_{k=1}^{\infty}\frac{\tilde{\Delta}_k^{(\pm1)}(z)}{N^k},\qquad z\in\partial B_{\pm1}(z) \tag{A.4}$$

where $\tilde{\Delta}_k^{(\pm 1)}(z)$ are independent of N (and independent of $\omega(z)$), and have meromorphic continuations inside the disc $\partial B_{\pm 1}(\delta)$ with a pole of order $(3k+1)/2$ at $z = \pm 1$. Inserting (A.4) back into (A.2), we find that

$$P_{\pm 1}(z)P_\infty(z)^{-1} - I \sim \sum_{k=1}^{\infty} \frac{Q(z)\tilde{\Delta}_k^{(\pm 1)}(z)Q(z)^{-1}}{N^k}, \qquad z \in \partial B_{\pm 1}(\delta) \tag{A.5}$$

where $Q(z) = P_\infty(z)\omega(z)^{\sigma_3/2}\tilde{P}_\infty(z)^{-1}$. To prove the Lemma it is sufficient to show that

$$Q(z) = I + \mathcal{O}(d_N^{-1}), \qquad z \in B_{\pm 1}(\delta). \tag{A.6}$$

First note that

$$\omega(z) = 1 + \mathcal{O}(d_N^{-1}), \qquad z \in \partial B_{\pm 1}(\delta) \cup [-1, 1] \tag{A.7}$$

as follows immediately from the representation (3.21). Then the proof is complete if we can check that

$$\frac{\sqrt{z-1}\sqrt{z+1}}{2\pi} \int_{-1}^{1} \frac{\log \omega(x)}{\sqrt{1-x^2}(z-x)}\, dx = \mathcal{O}(d_N^{-1}), \qquad z \in \partial B_{\pm 1}(\delta) \tag{A.8}$$

because this would imply the corresponding estimate for the Szegö function $\mathcal{D}(z) = 1 + \mathcal{O}(d_N^{-1})$ (cf. (3.26)) so that $P_\infty(z) = \tilde{P}_\infty(z) + \mathcal{O}(d_N^{-1})$. We will prove (A.8) below only for $z \in \partial B_1(\delta)$, the case $z \in \partial B_{-1}(\delta)$ being identical. If $(z-x)^{-1}$ is bounded, the result follows immediately from (A.7), therefore we consider only the contribution to the integral (A.8) from a small neighbourhood $[1-\delta-\epsilon_0, 1-\delta+\epsilon_0]$ and the points $z \in \partial B_1(\delta)$ such that $0 < |z-(1-\delta)| < \epsilon_0/2$. First consider $\mathrm{Im}(z) > 0$ and let $\mathcal{C}$ denote the clockwise oriented semi-circle in the upper-half plane connecting the points $1-\delta-\epsilon_0$ and $1-\delta+\epsilon_0$. Then by the residue theorem and analyticity of $\omega(x)$, (A.8) is equal to

$$i\sqrt{z+1}\sqrt{z-1}\frac{\log \omega(z)}{\sqrt{1-z^2}} + \frac{\sqrt{z-1}\sqrt{z+1}}{2\pi} \int_{\mathcal{C}} \frac{\log \omega(x)}{\sqrt{1-x^2}(x-z)}\, dx \tag{A.9}$$

where we take the principal branch of the square root. Now both terms in (A.9) are clearly $O(d_N^{-1})$, as follows from (A.7) and the fact that $(x-z)^{-1}$ is uniformly bounded in (A.9). A similar calculation applies when $\mathrm{Im}(z) < 0$. This completes the proof of the Lemma. $\square$

Final transformation

We will now define the final transformation of the Riemann-Hilbert problem, $S \to R$. As usual, we set

$$R(z) = \begin{cases} S(z)P_\infty(z)^{-1}, & z \in U_\infty \setminus \Sigma \\ S(z)P_{\pm 1}(z)^{-1}, & z \in B_{\pm 1}(\delta) \setminus \Sigma \end{cases} \tag{A.10}$$

From the Riemann-Hilbert problem for $S(z)$, it is easily shown that $R(z)$ has jumps only on $\partial B_{\pm 1}(\delta)$, $\mathbb{R} \setminus [-1-\delta, 1+\delta]$ and the parts of $\Sigma_\pm$ outside of $B_1(\delta) \cup B_{-1}(\delta)$ (denoted here by $\Gamma_\pm$). In what follows, we will denote the disjoint union of these contours as Σ_R, which we plot in Figure 2. The function $R(z)$ satisfies the following:

1. $R(z)$ is analytic in $\mathbb{C} \setminus \Sigma_R$.

2. $R(z)$ satisfies the jump condition $R_+(s) = R_-(s)J(s)$ where

$$J(s) = P_\infty(s) \begin{pmatrix} 1 & \omega(s)e^{N(g_+(s)+g_-(s)-2s^2-l)} \\ 0 & 1 \end{pmatrix} P_\infty(s)^{-1}, s \in \mathbb{R} \setminus [-1-\delta, 1+\delta] \tag{A.11}$$

$$J(s) = P_\infty(s) \begin{pmatrix} 1 & 0 \\ \omega(s)^{-1}e^{\mp Nh(s)} & 1 \end{pmatrix} P_\infty(s)^{-1}, \qquad s \in \Gamma_\pm \tag{A.12}$$

$$J(s) = P_{\pm 1}(s)P_\infty(s)^{-1} \qquad s \in \partial B_{\pm 1} \tag{A.13}$$

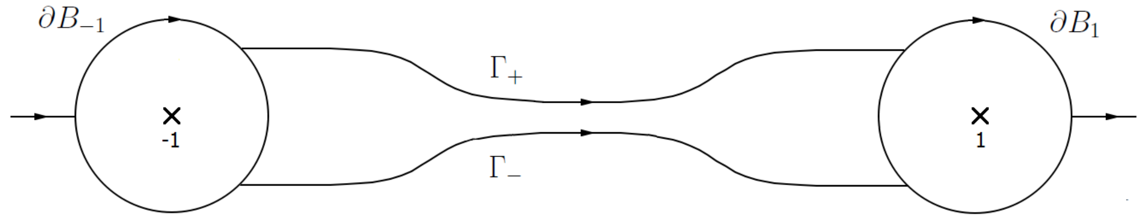

FIG 2. *The contour Σ_R for the $R(z)$ Riemann-Hilbert problem. The parts of the lenses $\Gamma = \Sigma \setminus \partial B_{\pm 1}(\delta)$ near x_0 are of distance $O(d_N^{-1})$ from the real line. The circles $\partial B_{\pm 1}(\delta)$ are of radius δ.*

3. $R(z) = I + O(z^{-1})$ as $z \to \infty$.

Estimating the jump matrix $\boldsymbol{\Delta(s)}$

Before we estimate the jump matrix we need to understand the behaviour of $P_\infty(z)$ (cf. (3.25)) on the contours $\Gamma_\pm$.

Lemma A.3. *The Szegö function $\mathcal{D}(s)$ in (3.26) and its inverse $\mathcal{D}(s)^{-1}$ are uniformly bounded on the contours $\Gamma_\pm$. In fact we have*

$$\log \mathcal{D}(s) = \mathcal{O}(1), \qquad N \to \infty, \tag{A.14}$$

uniformly for $s \in \Gamma_\pm$.

Proof. It suffices to prove that

$$\int_{-1}^{1} \frac{\log \omega(x)}{(s-x)\sqrt{1-x^2}}\, dx = \mathcal{O}(1). \tag{A.15}$$

By definition of the weight (3.21), we have the elementary inequality

$$|\log(\omega(x))| \leq \frac{1}{2}\sum_{k=1}^{m-1} |\alpha_k| \Big| \log\left(1 + g_{\tau,\eta,N}(x,x_0)\right) \Big| \tag{A.16}$$

where

$$g_{\tau,\eta,N}(x,x_0) = \frac{(\tau/d_N)^2 - 2(x-x_0)\tau/d_N}{(x-x_0)^2 + (\eta/d_N)^2} \tag{A.17}$$

Now, clearly if $x \leq x^* = x_0 + \tau/(2d_N)$, we have $g_{\tau,\eta,N}(x,x_0) \geq 0$, so that $\log(1 + g_{\tau,\eta,N}(x,x_0)) \leq g_{\tau,\eta,N}(x,x_0)$. If $x > x^*$, we symmetrise about the point x^* exploiting the symmetry $|\log(1+g_{\tau,\eta,N}(x^* - x, x_0)| = |\log(1 + g_{\tau,\eta,N}(x^* + x, x_0)|$ to obtain

$$|\log(1 + g_{\tau,\eta,N}(x,x_0))| \leq |g_{\tau,\eta,N}(x,x_0)| + |g_{\tau,\eta,N}(2x^* - x, x_0))|. \tag{A.18}$$

We will focus only on the region $x \in [x_0 - \epsilon, x_0 + \epsilon]$ as this gives the dominant contribution to the integral (A.15). For $s \in \Gamma_\pm$ and $x \in [x_0 - \epsilon, x_0 + \epsilon]$ we have $|s - x|^{-1} \leq ((x - x_0)^2 + (\eta/2d_N)^2)^{-1/2}$ and $(1 - x^2)^{-1/2} = \mathcal{O}(1)$. Then the contribution to (A.15) from the first term on the r.h.s. of (A.18) is bounded by

$$\int_{x_0-\epsilon}^{x_0+\epsilon} \frac{|g_{\tau,\eta,N}(x,x_0)|}{\sqrt{(x-x_0)^2 + (\eta/2d_N)^2}}\, dx \leq \int_{-1}^{1} \frac{|(\tau/d_N)^2 - 2x\tau/d_N|}{(x^2 + (\eta/2d_N)^2)^{3/2}}\, dx \tag{A.19}$$

$$= \frac{8|\tau|}{\eta}\left(\frac{\sqrt{\tau^2/\eta^2+1}\sqrt{(2d_N/\eta)^2+1} - 1}{\sqrt{(2d_N/\eta)^2+1}}\right) = \mathcal{O}(1) \tag{A.20}$$

where we changed variables $x \to x - x_0$ and extended the limits of integration back to $[-1, 1]$. The resulting integral on the right-hand side of (A.19) can be evaluated exactly in *e.g.* Maple.

For the second term in (A.18), we use the estimate $((x-x_0)^2+(\eta/d_N)^2)^{-1/2} \leq c((x_0-x+\tau/d_N)^2 + (\eta/d_N)^2)^{-1/2}$ (where c depends on η and τ only) to get

$$\int_{x_0-\epsilon}^{x_0+\epsilon} \frac{|g_{\tau,\eta,N}(2x^* - x, x_0)|}{\sqrt{(x-x_0)^2 + (\eta/(2d_N))^2}}\, dx \leq c \int_{x_0-\epsilon}^{x_0+\epsilon} \frac{|(\tau/d_N)^2 - 2(x_0 - x + \tau/d_N)\tau/d_N|}{((x_0 - x + \tau/d_N)^2 + (\eta/(2d_N))^2)^{3/2}}\, dx \tag{A.21}$$

$$= c \int_{-\epsilon+\tau/d_N}^{\epsilon+\tau/d_N} \frac{|(\tau/d_N)^2 - 2u\tau/d_N|}{(u^2 + (\eta/(2d_N))^2)^{3/2}}\, du = \mathcal{O}(1) \tag{A.22}$$

where we used that the last integral is bounded by the r.h.s. of (A.19). □

Proposition A.4. *Let $\Delta(s) = J(s) - I$ where $J(s)$ is the jump matrix for $R(z)$ defined on the contour Σ_R. We have the following bounds*

- *On the discs*

$$|\Delta(s)| = \mathcal{O}(N^{-1}), \qquad s \in \partial B_{\pm 1}(\delta). \tag{A.23}$$

- *On the upper and lower lips*

$$|\Delta(s)| = \mathcal{O}\left(\exp\left(-c_1 \frac{N}{d_N}\right)\right), \qquad s \in \Gamma_\pm. \tag{A.24}$$

- *On the real line*

$$|\Delta(s)| = \mathcal{O}\left(\exp\left(-c_2 N\right)\right), \qquad s \in \mathbb{R} \setminus [-1-\delta, 1+\delta]. \tag{A.25}$$

Here, $c_1 > 0$ and $c_2 > 0$ are constants depending only on δ and η.

Proof. The bound (A.23) follows immediately from Lemma A.2, while (A.25) follows from the fact that $P_\infty(s)$ is uniformly bounded in $\mathbb{R}\setminus[-1-\delta, 1+\delta]$ combined with the inequalities (3.16). It remains to settle (A.24). On the contours $\Gamma_\pm$, we have the explicit expression

$$\Delta(s) = e^{\mp N h(s)} \begin{pmatrix} P_\infty(s)_{12} P_\infty(s)_{22} & -(P_\infty(s)_{12})^2 \\ (P_\infty(s)_{22})^2 & -P_\infty(s)_{12} P_\infty(s)_{22} \end{pmatrix}, \qquad s \in \Gamma_\pm. \tag{A.26}$$

where $h(s)$ was defined in (3.20). By Lemma A.3 we see that $P_\infty(s)$ is uniformly bounded on $\Gamma_\pm$. Therefore, the only danger is that $\operatorname{Re} h(s)$ vanishes too quickly as $N \to \infty$. However, a careful examination of the function (3.20) shows that $\operatorname{Re} h(z)$ vanishes at the same rate that the contours $\Gamma_\pm$ collapse onto the real axis. Indeed, an elementary calculation using Taylor's theorem shows that we have the inequalities

$$\begin{aligned} &\operatorname{Re}(h(s)) > c_1/d_N, \qquad s \in \Gamma_+, \\ &\operatorname{Re}(h(s)) < -c_1/d_N, \qquad s \in \Gamma_-, \end{aligned} \tag{A.27}$$

where $c_1 = 4\eta\sqrt{1-(1-\delta)^2}$. This concludes the proof of (A.24). □

Estimating the R-matrix and the proof of Proposition 3.2

Finally we are in a position to prove Proposition 3.2. The proof follows from the standard method described in [13]. However, in our case extra care must be taken with the estimates because our contour Σ_R depends explicitly on N, see *e.g.* [6] for another example of N-dependent contours.

Proposition A.5. *The matrix $R(z)$ satisfies the following estimate*

$$R(z) = I + \mathcal{O}\left(\frac{1}{N}\right) + \mathcal{O}\left(\log(d_N)\exp\left(-c_1\frac{N}{d_N}\right)\right), \qquad N \to \infty \tag{A.28}$$

uniformly for $z \in \mathbb{C} \setminus \Sigma_R$.

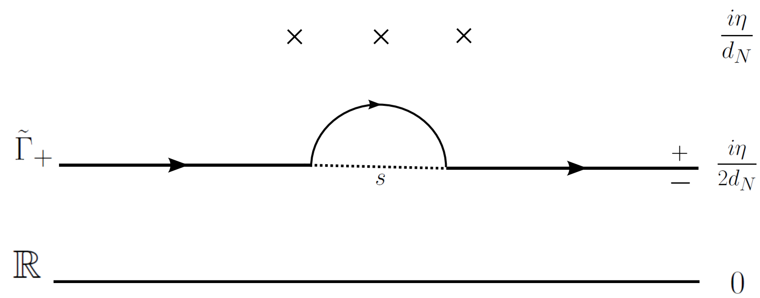

FIG 3. *The deformed contour $\tilde{\Gamma}_+$. The semi-circle of radius $\eta/(4d_N)$ is sufficiently small that it does not touch the singularities (crosses), whose imaginary parts are η/d_N.*

Proof. Since for every N, Σ_R is a finite union of smooth contours, standard theory (see *e.g.* [14, 36, 35]) gives

$$R(z) = I + \frac{1}{2\pi i}\int_{\Sigma_R} \frac{\Delta(s)\nu(s)}{s-z}\,ds \tag{A.29}$$

where $\Delta(s)$ is as in Proposition A.4 and $\nu(s)$ is the unique solution to the singular integral equation $\nu(s) = I + C_-[\nu\Delta](s)$. Here, C_- is the Cauchy operator on $L^2(\Sigma_R)$, defined by

$$C_-[f](s) = \frac{1}{2\pi i}\int_{\Sigma_R} \frac{f(x)}{x - s_-}\,dx, \qquad f \in L^2(\Sigma_R) \tag{A.30}$$

where s_- denotes the limiting value of the integral as the point $s \in \Sigma_R$ is approached from the minus side of the contour.

We begin by solving the equation for $\nu(s)$ in a perturbation series (see *e.g.* [5])

$$\nu(s) = I + \sum_{k=1}^{\infty} \nu_k(s), \qquad \nu_k(s) = C_-[\nu_{k-1}\Delta](s), \tag{A.31}$$

and $\nu_0 = I$. We need to show that this series is absolutely and uniformly convergent for any $s \in \Sigma_R$. Let $s \in \Gamma_+$ and deform Γ_+ to a new contour $\tilde{\Gamma}_+$ differing only by a small semi-circle of radius $\eta/(4d_N)$ centered at s, as depicted in Figure 3. Denote by $\tilde{\Sigma}_R$ the contour Σ_R with Γ_+ replaced with $\tilde{\Gamma}_+$. By the Cauchy theorem, we have

$$\nu_1(s) = \frac{1}{2\pi i}\int_{\Sigma_R} \frac{\Delta(x)}{x - s_-}\,dx = \frac{1}{2\pi i}\int_{\tilde{\Sigma}_R} \frac{\Delta^{(0)}(x)}{x - s}\,dx \tag{A.32}$$

where $\Delta^{(0)}$ is the analytic continuation of Δ to $\tilde{\Sigma}_R$ and satisfies the same bounds as in Proposition A.4. Now we estimate, splitting the integral into a contribution from the discs $\partial B_{\pm 1}(\delta)$, the real line $\mathbb{R} \setminus [-1-\delta, 1+\delta]$ (both of which are at most $\mathcal{O}(N^{-1})$) and the contribution from $\tilde{\Gamma}_\pm$:

$$\begin{aligned}
|\nu_1(s)| &\leq c_3/N + \frac{1}{2\pi}\int_{\tilde{\Gamma}_\pm} \frac{|\Delta^{(0)}(x)|}{|x-s|}\,dx, \qquad s \in \Gamma_+ \\
&\leq c_3/N + \frac{1}{2\pi} e^{-c_1 N/d_N} \int_{\tilde{\Gamma}_\pm} \frac{1}{|x-s|}\,dx, \qquad s \in \Gamma_+ \\
&\leq c_3/N + c_2 \log(d_N) e^{-c_1 N/d_N}, \qquad s \in \Gamma_+
\end{aligned} \tag{A.33}$$

where c_3 and c_2 are constants depending only on δ and η, with a similar bound if $s \in \Gamma_-$. If $s \in \Sigma_R \setminus (\Gamma_+ \cup \Gamma_-)$ then the same bound holds with $c_2 = 0$. Applying this procedure inductively, we obtain

$$|\nu_j(s)| \leq K_1 N^{-j} + K_2 \left(\log(d_N) e^{-c_1 N/d_N}\right)^j, \qquad s \in \Sigma_R, \tag{A.34}$$

where we can choose $K_2 = 0$ if $s \in \Sigma_R \setminus (\Gamma_+ \cup \Gamma_-)$. The bound (A.34) implies that the series (A.31) is absolutely convergent. Inserting (A.31) back into (A.29) we arrive at

$$R(z) = I + \sum_{j=1}^{\infty} R_j(z), \qquad R_j(z) = \frac{1}{2\pi i}\int_{\Sigma_R} \frac{\nu_{j-1}(s)\Delta(s)}{s-z}\,ds, \qquad j = 1,2,3,\dots \tag{A.35}$$

Now we bound the terms in the sum (A.35). First consider the case that $\mathrm{dist}(z,\Sigma_R) \geq \eta/(4d_N)$. Then estimates entirely analogous to (A.33) yield

$$|R_j(z)| \leq K_1 N^{-j} + K_2\left(\log(d_N)e^{-c_2 N/d_N}\right)^j, \qquad j = 1,2,3,\dots \tag{A.36}$$

On the other hand, if $0 < \mathrm{dist}(z,\Sigma_R) < \eta/(4d_N)$, one can again deform the contour with a semi-circle of radius $\eta/(4d_N)$ and obtain the same bound (A.36) after essentially repeating the steps (A.32) and (A.33). $\square$

Remark A.6. To complete the proof of Proposition 3.2 we will derive the explicit form of the $O(1/N)$ term in (A.28). Thus we need to compute the function $R_1(z)$ defined in (A.35). By Proposition A.4 and Lemma A.2 we have

$$R_1(z) = \frac{\tilde{R}_1(z)}{N} + \mathcal{O}\left(\frac{1}{Nd_N}\right) + \mathcal{O}\left(d_N \exp\left(-c_1\frac{N}{d_N}\right)\right) \tag{A.37}$$

where

$$\tilde{R}_1(z) = \frac{1}{2\pi i}\int_{\partial B_1(\delta)} \frac{\Delta_1^{(+1)}(s)}{s-z}\,ds + \frac{1}{2\pi i}\int_{\partial B_{-1}(\delta)} \frac{\Delta_1^{(-1)}(s)}{s-z}\,ds. \tag{A.38}$$

The functions $\Delta_1^{(\pm 1)}(s)$ are explicitly known, *e.g.* by setting $\omega(z) \equiv 1$ in Eqs. (79, 83), of [35] or by using the results in [13]. Then expanding (A.37) near $z = \infty$ and computing the residues of the function $\Delta_1^{(\pm 1}(s)$ near the poles $s = \pm 1$, we find that

$$\tilde{R}_1(z) = A/z + B/z^2 + O(z^{-3}), \qquad z \to \infty \tag{A.39}$$

where

$$A = \begin{pmatrix} 0 & i/24 \\ i/24 & 0 \end{pmatrix}, \qquad B = \begin{pmatrix} -1/48 & 0 \\ 0 & 1/48 \end{pmatrix} \tag{A.40}$$

Then inserting (A.28) and the first order correction above into the definition (A.10), we arrive at (3.29).

Appendix B: The Szegö function

For a weight $\omega(x)$, the Szegö function is defined by the formula

$$\mathcal{D}(z) = \exp\left(\frac{\sqrt{z+1}\sqrt{z-1}}{2\pi}\int_{-1}^{1}\frac{\log(\omega(x))}{\sqrt{1-x^2}}\frac{dx}{z-x}\right). \tag{B.1}$$

It satisfies the properties

1. $\mathcal{D}(z)$ is non-zero and analytic in $\mathbb{C} \setminus [-1,1]$,
2. $\mathcal{D}_+(x)\mathcal{D}_-(x) = \omega(x)$ for $x \in (-1,1)$,
3. $\lim_{z\to\infty}\mathcal{D}(z) = \mathcal{D}_\infty \neq 0$.

For our problem, we are interested in the weight $\omega(x) = \prod_{k=1}^{m} |x - z_k|^{\alpha_k}$ where $\mathrm{Im}(z_k) \neq 0$ for $k = 1, \ldots, m$. It can easily be seen that the above three properties uniquely specify the Szegö function for this weight. Let $c(z) = z + \sqrt{z-1}\sqrt{z+1}$ be the conformal map from $\mathbb{C} \setminus [-1, 1]$ to the exterior of the unit disk. Then the Szegö function for the weight $|x - \mu|^2$ is

$$\frac{|c(\mu)|}{2}\left(1 - \frac{1}{c(\mu)c(z)}\right)\left(1 - \frac{1}{\overline{c(\mu)}c(z)}\right), \quad \mathrm{Im}(\mu) \neq 0. \tag{B.2}$$

This can be checked by verifying the above three conditions using the properties $c(z) + \frac{1}{c(z)} = 2z$ and $c_+(x)c_-(x) = 1$ for $x \in [-1, 1]$. Thus the Szegö function for $\omega(x)$ is

$$\mathcal{D}(z) = \prod_{k=1}^{m}\left(\frac{|c(z_k)|}{2}\left(1 - \frac{1}{c(z_k)c(z)}\right)\left(1 - \frac{1}{\overline{c(z_k)}c(z)}\right)\right)^{\alpha_k/2}. \tag{B.3}$$

Similar considerations show straightforwardly that the function $C(z, \mu)$ defined in (3.35) is given by

$$C(z, \mu) = \frac{1}{2}\log\left(\frac{|c(\mu)|}{2}\left(1 - \frac{1}{c(\mu)c(z)}\right)\left(1 - \frac{1}{\overline{c(\mu)}c(z)}\right)\right). \tag{B.4}$$

Defining $z_k = x_0 + \frac{\tau_k + i\eta}{d_N}$, one easily gets the asymptotic

$$d_N\frac{|c(z_j)|}{2}\left(1 - \frac{1}{c(z_j)c(z_k)}\right)\left(1 - \frac{1}{\overline{c(z_j)}c(z_k)}\right) = 2\eta + i(\tau_j - \tau_k) + \mathcal{O}(d_N^{-1}) \tag{B.5}$$

which immediately implies that

$$\mathrm{Re}(C(z_j, z_k)) = -\frac{1}{2}\log(d_N) + \frac{1}{4}\log((\tau_j - \tau_k)^2 + 4\eta^2) + \mathcal{O}(d_N^{-1}). \tag{B.6}$$

The uniformity of the error term in the relevant compact sets follows from the uniform expansions of the logarithm and square roots in these regions. From (B.3) we obviously have the expansion

$$\mathcal{D}(z) = \mathcal{D}_\infty\left(1 + \frac{\mathcal{D}_1}{z} + \frac{\mathcal{D}_1^2/2 + \mathcal{D}_2}{z^2}\right) + O(z^{-3}) \tag{B.7}$$

where

$$\mathcal{D}_\infty = \prod_{k=1}^{m-1}\left|\frac{c(z_k)}{c(z_m)}\right|^{\alpha_k/2}. \tag{B.8}$$

and

$$\mathcal{D}_1 = -\frac{1}{2}\sum_{k=1}^{m}\alpha_k \,\mathrm{Re}\left(\frac{1}{c(z_k)}\right), \quad \mathcal{D}_2 = -\frac{1}{8}\sum_{k=1}^{m}\alpha_k \,\mathrm{Re}\left(\frac{1}{c(z_k)^2}\right). \tag{B.9}$$

Appendix C: Proof of equation 4.29

Our first task is to prove that we have the limit

$$\lim_{N\to\infty}\int_{[I_N^c]^2}\frac{\Delta f_1(d_N x)}{\Delta x}\frac{\Delta f_2(d_N x)}{\Delta x}F_N(x_1, x_2)\,dx_1\,dx_2 = 0, \tag{C.1}$$

where I_N^c is the complement of the region $I_N = [-(1 - \delta_N), (1 - \delta_N)]$, $\delta_N = N^{-7/12}$ and we defined $F_N(x, y) = (x - y)^2 K_N^2(x, y)$ in terms of the GUE kernel (4.17). After proving (C.1) we show that δ_N can be replaced with an N-independent $\delta > 0$ costing an error term that can be neglected.

Let $0 < \epsilon < 1$ and consider the following three subsets of $\mathbb{R}^2$,

$$
\begin{aligned}
R_1 &= \{(x_1, x_2) \in \mathbb{R}^2 \mid (|x_1| < \epsilon) \wedge (x_2 > (1+\delta_N))\}, \\
R_2 &= \{(x_1, x_2) \in \mathbb{R}^2 \mid (|x_1| < \epsilon) \wedge (1-\delta_N < x_2 < 1+\delta_N)\}, \\
R_3 &= \{(x_1, x_2) \in \mathbb{R}^2 \mid (x_1 > \epsilon) \wedge (x_2 > \epsilon)\}.
\end{aligned}
$$

It is sufficient to consider only these regions, because together with their reflections in the x_1 and x_2 axes, they cover the entire region $[I_N^c]^2$. In the following we will prove that the contribution from each of these regions to the integral (C.1) tends to zero as $N \to \infty$. Finally we complete the proof of Eq. (4.19) by showing that the difference between the integral (C.1) over $[I_N^c]^2$ and $[I_\delta^c]^2$ converges as $N \to \infty$ to a function that is $O(\delta)$ as $\delta \to 0$.

We start with the contribution of the region R_3 to the integral (C.1). Using the Schwartz property of f_1, f_2 and the inequality $K_N^2(x_1, x_2) \leq N^2 \rho_N(x_1)\rho_N(x_2)$, we have for any $\gamma > 0$

$$
\left| \int_\epsilon^\infty \int_\epsilon^\infty \Delta f_1(d_N x) \Delta f_2(d_N x) K_N^2(x_1, x_2)\, dx_1\, dx_2 \right| \tag{C.2}
$$

$$
\leq N^2 (2\epsilon d_N)^{-2\gamma} \left(\int_\epsilon^\infty \rho_N(x_1)\, dx_1 \right) \left(\int_\epsilon^\infty \rho_N(x_2)\, dx_2 \right) = O(N^2 d_N^{-\infty}) \tag{C.3}
$$

where we used the inequality $|\Delta g_j(d_N x)| \leq |g_j(d_N x_1) + g_j(d_N x_2)| \leq d_N^{-\gamma}(|x_1|^{-\gamma} + |x_2|^{-\gamma}) \leq 2 d_N^{-\gamma}(\epsilon^{-\gamma})$. We conclude that the integral (4.16) restricted to the region R_3 is of order $O(N^{-\infty})$ as $N \to \infty$.

Now let us consider the edge region R_2. We will make use of the following Lemma from [42], which states

Lemma C.1 (Theorem 5.2.3 (ii) [42])**.** *Let $\rho_N(x)$ denote the normalized density of states, as in (4.7). The bound*

$$
\rho_N(1 + sN^{-2/3}) \leq (BN^{1/3}s)^{-1} e^{-bs^{3/2}} \tag{C.4}
$$

holds for N large enough. Here B and b are absolute constants and $s \to \infty$ as $N \to \infty$.

Using this result and again the bound $K_N(x_1, x_2)^2 \leq N^2 \rho_N(x_1)\rho_N(x_2)$, we see that the contribution to the integral (C.1) from the region R_2 is bounded by

$$
\begin{aligned}
& N^2 \int_{-\infty}^\infty \int_{(1+\delta_N)}^\infty |\Delta f_1(d_N x)||\Delta f_2(d_N x)| \rho_N(x_1)\rho_N(x_2)\, dx_1\, dx_2 \\
&= C\delta_N N^2 \int_{-\infty}^\infty \int_1^\infty \rho_N(1 + x_1\delta_N)\rho_N(x_2)\, dx_1\, dx_2 \\
&\leq CBN \int_{-\infty}^\infty \int_1^\infty x_1^{-1} e^{-bx_1^{3/2} N^{1/8}} \rho_N(x_2)\, dx_1\, dx_2 = O(N^{-\infty}), \qquad N \to \infty,
\end{aligned} \tag{C.5}
$$

where we used that f_1, f_2 are uniformly bounded on $\mathbb{R}^2$.

For the region R_1, we need a bound for the absolute value of the functions $\psi_l^{(N)}(x)$.

Lemma C.2 (Szegö, Sect. 10.8 [51])**.** *Let $\psi_l^{(N)}(x)$ denote the orthonormal functions defined in (4.18). Then the following bound holds uniformly in l as $N \to \infty$,*

$$
\sup_{u \in \mathbb{R}} |\psi_l^{(N)}(u)| = O(N^{1/4}). \tag{C.6}
$$

First consider the contribution from the product of squares, i.e. that of $\psi_N^{(N)}(x_1)^2 \psi_{N-1}^{(N)}(x_2)^2$ in $F_N(x_1, x_2)$. Since in the region R_1 we have $x_1 \neq x_2$, the bound $|\Delta f_j(d_N x)/\Delta x| \leq C$, $j = 1, 2$ holds for some N-independent $C > 0$. Then the contribution coming from $\psi_N^{(N)}(x_1)^2\psi_{N-1}^{(N)}(x_2)^2$ is bounded by

$$C\int_{(1-\delta_N)}^{(1+\delta_N)}\int_{-\epsilon}^{\epsilon}\psi_N^{(N)}(x_1)^2\psi_{N-1}^{(N)}(x_2)^2\,dx_1\,dx_2 \tag{C.7}$$

$$\leq C\int_{(1-\delta_N)}^{(1+\delta_N)}\int_{-\infty}^{\infty}\psi_N^{(N)}(x_1)^2\sup_{u\in\mathbb{R}}|\psi_{N-1}^{(N)}(u)|^2\,dx_1\,dx_2 \leq C'N^{-1/12} \tag{C.8}$$

where $C' > 0$ is another constant independent of N. A similar calculation shows that the contribution from the mixed term $\psi_N^{(N)}(x_1)\psi_{N-1}^{(N)}(x_1)\psi_N^{(N)}(x_2)\psi_{N-1}^{(N)}(x_2)$ is also $O(N^{-1/12})$ as $N \to \infty$. We conclude that the contribution of the region R_1 is $O(N^{-1/12})$ as $N \to \infty$. Finally, a completely analogous calculation shows that the contribution to (4.16) coming from all reflections of the regions R_1, R_2 and R_3 in the x_1 and x_2 axes satisfy the same corresponding asymptotic estimates as $N \to \infty$ and therefore may be neglected. Eq. (C.1) is proven.

To complete the argument, we need to show that the difference between the integral (4.16) over I_N^2 and the same integral over $I_\delta = [-(1-\delta),(1-\delta)]^2$ for some N-independent $\delta > 0$, can be neglected in the limit $N \to \infty$. It will be sufficient to consider only the thin strip $|x_1| < \epsilon$ and $(1-\delta) < x_2 < (1-\delta_N)$, because the remaining parts of $I_N^c \setminus I_\delta$ are either reflections of this region or are subsets of the region R_1 treated earlier. Thus, we just have to estimate the integral

$$\int_{(1-\delta)}^{(1-\delta_N)}\int_{-\epsilon}^{\epsilon}\frac{\Delta f_1(d_N x)}{\Delta x}\frac{\Delta f_2(d_N x)}{\Delta x}F_N(x_1,x_2)\,dx_1\,dx_2 \tag{C.9}$$

According to the first Plancherel-Rotach formula of Corollary 5.1.5 in [42], we have the bound $F_N(x_1,x_2) = (1-x_1^2)^{-1/2}(4-x_2^2)^{-1/2}O(1)$ uniformly as $N \to \infty$. Therefore since $x_1 \neq x_2$ in (C.9) and f_1, f_2 are uniformly bounded, we see that (C.9) is bounded in absolute value by

$$C\bigg|\int_{(1-\delta)}^{(1-\delta_N)}\int_{-\epsilon}^{\epsilon}(1-x_1^2)^{-1/2}(1-x_2^2)^{-1/2}\,dx_1\,dx_2\bigg| \tag{C.10}$$

$$\leq C|\left(\cos^{-1}(1-\delta_N)-\cos^{-1}(1-\delta)\right)| \to C|\cos^{-1}(1-\delta)|, \quad N \to \infty, \tag{C.11}$$

where $C > 0$ is some N-independent constant. Hence, by choosing $\delta > 0$ sufficiently small, we can ensure that the integral over this strip is as small as we desire. This proves Eq. (4.29).

Acknowledgements

We would like to thank Paul Bourgade, Arno Kuijlaars and Leonid Pastur for insightful discussions and correspondence relating to our results, Philippe Sosoe and Percy Wong for sharing their preprint [50] with us and Jérémie Unterberger for bringing the paper [52] to our attention. We are particularly grateful to Igor Krasovsky for informative discussions about the Riemann-Hilbert problem, and also to anonymous referees for their helpful suggestions and constructive critique of the first version of the paper. The second author thanks the Isaac Newton Institute, Cambridge, UK, for its support and hospitality during the semester Periodic and Ergodic Spectral Problems.

References

[1] Greg W. Anderson, Alice Guionnet, and Ofer Zeitouni. *An introduction to random matrices*, volume 118 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2010.

[2] Kari Astala, Peter Jones, Antti Kupiainen, and Eero Saksman. Random conformal weldings. *Acta Math.*, 207(2):203–254, 2011.

[3] E. Bacry and J. F. Muzy. Log-infinitely divisible multifractal processes. *Comm. Math. Phys.*, 236(3):449–475, 2003.

[4] Julien Barral and Benoît B. Mandelbrot. Non-degeneracy, moments, dimension, and multifractal analysis for random multiplicative measures (Random multiplicative multifractal measures. II). In *Fractal geometry and applications: a jubilee of Benoît Mandelbrot, Part 2*, volume 72 of *Proc. Sympos. Pure Math.*, pages 17–52. Amer. Math. Soc., Providence, RI, 2004.

[5] P. Bleher. Lectures on Random Matrix Models. the Riemann-Hilbert approach. In J. Harnad, editor, *Random Matrices, Random Processes and Integrable Systems*. Springer, 2008.

[6] Pavel M. Bleher and Vladimir V. Fokin. Exact solution of the six-vertex model with domain wall boundary conditions. Disordered phase. *Comm. Math. Phys.*, 268(1):223–284, 2006.

[7] Alexei Borodin and Vadim Gorin. General beta Jacobi corners process and the Gaussian Free Field. eprint = `arXiv:1305.3627`, 2013.

[8] P. Bourgade, L. Erdös, H.-T. Yau, and J. Yin. Fixed energy universality for generalized Wigner matrices. eprint = `arXiv:1407.5606` , 2014.

[9] A. Boutet de Monvel and A. Khorunzhy. Asymptotic distribution of smoothed eigenvalue density. I. Gaussian random matrices. *Random Oper. Stochastic Equations*, 7(1):1–22, 1999.

[10] A. Boutet de Monvel and A. Khorunzhy. Asymptotic distribution of smoothed eigenvalue density. II. Wigner random matrices. *Random Oper. Stochastic Equations*, 7(2):149–168, 1999.

[11] J. Breuer and M. Duits. Universality of mesoscopic fluctuations for orthogonal polynomial ensembles. eprint = `arXiv:1411.5205`, 2014.

[12] 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 ane sinh-Gordon models. *Phys. Rev. E*, 63:026110 (33pp), 2001.

[13] 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(12):1491–1552, 1999.

[14] P. A. Deift. *Orthogonal polynomials and random matrices: a Riemann-Hilbert approach*, volume 3 of *Courant Lecture Notes in Mathematics*. New York University Courant Institute of Mathematical Sciences, New York, 1999.

[15] Persi Diaconis and Mehrdad Shahshahani. On the eigenvalues of random matrices. *J. Appl. Probab.*, 31A:49–62, 1994. Studies in applied probability.

[16] M. Duits and K. Johansson. On mesoscopic equilibrium for linear statistics in dyson's brownian motion. eprint = `arXiv:1312.4295`, 2013.

[17] N. M. Ercolani and K. D. T.-R. McLaughlin. Asymptotics of the partition function for random matrices via Riemann-Hilbert techniques and applications to graphical enumeration. *Int. Math. Res. Not.*, (14):755–820, 2003.

[18] László Erdős and Antti Knowles. The Altshuler–Shklovskii Formulas for Random Band Matrices I: the Unimodular Case. *Comm. Math. Phys.*, 333(3):1365–1416, 2015.

[19] László Erdős and Antti Knowles. The Altshuler–Shklovskii Formulas for Random Band Matrices II: The General Case. *Ann. Henri Poincaré*, 16(3):709–799, 2015.

[20] E. Faleiro, J. M. G. Gómez, R. A. Molina, L. Muñoz, A. Relaño, and J. Relamosa. Theoretical Derivation of $1/f$ Noise in Quantum Chaos. *Phys. Rev. Lett.*, 93(24):244101[4pages], 2004.

[21] A. S. Fokas, A. R. Its, and A. V. Kitaev. The isomonodromy approach to matrix models in 2D quantum gravity. *Comm. Math. Phys.*, 147(2):395–430, 1992.

[22] Ronald F. Boisvert Frank W. J. Olver, Daniel W. Lozier, Macintosh Charles W. Clark, With 1 CD-ROM (Windows, and UNIX), editors. *NIST handbook of mathematical functions*. U.S. Department of Commerce National Institute of Standards and Technology, Washington, DC, 2010.

[23] Y. V. Fyodorov, G. H. Hiary, and J. P. Keating. Freezing Transition, Characteristic Polynomials of Random Matrices, and the Riemann Zeta-Function. *Phys. Rev. Lett.*, 108:170601, 2012.

[24] Yan V. Fyodorov and Jean-Philippe Bouchaud. Freezing and extreme-value statistics in a random

energy model with logarithmically correlated potential. *J. Phys. A*, 41(37):372001, 12, 2008.

[25] Yan V. Fyodorov and Jonathan P. Keating. Freezing transitions and extreme values: random matrix theory, and disordered landscapes. *Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 372(2007):20120503, 32, 2014.

[26] Yan V. Fyodorov, Pierre Le Doussal, and Alberto Rosso. Counting function fluctuations and extreme value threshold in multifractal patterns: the case study of an ideal $1/f$ noise. *J. Stat. Phys.*, 149(5):898–920, 2012.

[27] Stavros Garoufalidis and Ionel Popescu. Analyticity of the planar limit of a matrix model. *Ann. Henri Poincaré*, 14(3):499–565, 2013.

[28] L. Š. Grinblat. A limit theorem for measurable random processes and its applications. *Proc. Amer. Math. Soc.*, 61(2):371–376 (1977), 1976.

[29] Jonas Gustavsson. Gaussian fluctuations of eigenvalues in the GUE. *Ann. Inst. H. Poincaré Probab. Statist.*, 41(2):151–178, 2005.

[30] C. P. Hughes, J. P. Keating, and Neil O'Connell. On the characteristic polynomial of a random unitary matrix. *Comm. Math. Phys.*, 220(2):429–451, 2001.

[31] Kurt Johansson. On fluctuations of eigenvalues of random Hermitian matrices. *Duke Math. J.*, 91(1):151–204, 1998.

[32] Jean-Pierre Kahane. Sur le chaos multiplicatif. *Ann. Sci. Math. Québec*, 9(2):105–150, 1985.

[33] J. P. Keating and N. C. Snaith. Random matrix theory and $\zeta(1/2 + it)$. *Comm. Math. Phys.*, 214(1):57–89, 2000.

[34] I. Krasikov. Some asymptotics for the Bessel functions with an explicit error term. eprint = `arXiv:1107.2007`, 2011.

[35] I. V. Krasovsky. Correlations of the characteristic polynomials in the Gaussian unitary ensemble or a singular Hankel determinant. *Duke Math. J.*, 139(3):581–619, 2007.

[36] A. B. J. Kuijlaars, K. T.-R. McLaughlin, W. Van Assche, and M. Vanlessen. The Riemann-Hilbert approach to strong asymptotics for orthogonal polynomials on $[-1, 1]$. *Adv. Math.*, 188(2):337–398, 2004.

[37] A. B. J. Kuijlaars and M. Vanlessen. Universality for eigenvalue correlations at the origin of the spectrum. *Comm. Math. Phys.*, 243(1):163–191, 2003.

[38] A. Lytova and L. Pastur. Central limit theorem for linear eigenvalue statistics of random matrices with independent entries. *Ann. Probab.*, 37(5):1778–1840, 2009.

[39] C. Male, G. Le Caër, and R. Delannay. $1/f^{\alpha}$ noise in the fluctuations of the spectra of tridiagonal random matrices from the β-hermite ensemble. *Phys. Rev. E*, 76(4):042101[4pages], 2007.

[40] Benoit B. Mandelbrot and John W. Van Ness. Fractional Brownian motions, fractional noises and applications. *SIAM Rev.*, 10:422–437, 1968.

[41] Madan Lal Mehta. *Random matrices*, volume 142 of *Pure and Applied Mathematics (Amsterdam)*. Elsevier/Academic Press, Amsterdam, third edition, 2004.

[42] Leonid Pastur and Mariya Shcherbina. *Eigenvalue distribution of large random matrices*, volume 171 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2011.

[43] George Oppenheim Paul Doukhan and Murad S. Taqqu, editors. *Theory and applications of long-range dependence*. Birkhäuser Boston Inc., Boston, MA, 2003.

[44] R. Rhodes and V. Vargas. Gaussian multiplicative chaos and applications: a review. eprint = `arXiv:1305.6221`, 2013.

[45] Brian Rider and Bálint Virág. The noise in the circular law and the Gaussian free field. *Int. Math. Res. Not. IMRN*, (2):Art. ID rnm006, 33, 2007.

[46] Gennady Samorodnitsky. Long range dependence. *Found. Trends Stoch. Syst.*, 1(3):163–257, 2006.

[47] F. G. Schmitt. A causal multifractal stochastic equation and its statistical properties. *Eur. Phys. J. B*, 34:85–98, 2003.

[48] Scott Sheffield. Gaussian free fields for mathematicians. *Probab. Theory Related Fields*, 139(3-

4):521–541, 2007.

[49] Alexander Soshnikov. The central limit theorem for local linear statistics in classical compact groups and related combinatorial identities. *Ann. Probab.*, 28(3):1353–1370, 2000.

[50] Philippe Sosoe and Percy Wong. Regularity conditions in the CLT for linear eigenvalue statistics of Wigner matrices. *Adv. Math.*, 249:37–87, 2013.

[51] Gabor Szegö. *Orthogonal Polynomials*. American Mathematical Society, New York, 1939. American Mathematical Society Colloquium Publications, v. 23.

[52] Jérémie Unterberger. Stochastic calculus for fractional Brownian motion with Hurst exponent $H > \frac{1}{4}$: a rough path method by analytic extension. *Ann. Probab.*, 37(2):565–614, 2009.