Mesoscopic fluctuations for unitary invariant ensembles

Gaultier Lambert*

Abstract

Considering a determinantal point process on the real line, we establish a connection between the sine-kernel asymptotics for the correlation kernel and the CLT for mesoscopic linear statistics. This implies universality of mesoscopic fluctuations for a large class of unitary invariant Hermitian ensembles. In particular, this shows that the support of the equilibrium measure need not be connected in order to see Gaussian fluctuations at mesoscopic scales. Our proof is based on the cumulants computations introduced in [45] for the CUE and the sine process and the asymptotic formulae derived by Deift et al. [13]. For varying weights $e^{-N\operatorname{Tr}V(\mathrm{H})}$, in the one-cut regime, we also provide estimates for the variance of linear statistics $\operatorname{Tr} f(\mathrm{H})$ which are valid for a rather general function f. In particular, this implies that the logarithm of the absolute value of the characteristic polynomials of such Hermitian random matrices converges in a suitable regime to a regularized fractional Brownian motion with logarithmic correlations introduced in [17]. For the GUE and Jacobi ensembles, we also discuss how to obtain the necessary sine-kernel asymptotics at mesoscopic scale by elementary means.

1 Introduction and results

A point process on $\mathbb{R}$ is called **determinantal** if its correlation functions (with respect to the Lebesgue measure) exist and are of the form:

$$\rho_k(x_1,\dots,x_k) = \det_{k\times k}\big[K(x_i,x_j)\big], \qquad \forall x_1,\dots,x_k \in \mathbb{R}, \qquad \forall k \in \mathbb{N}, \qquad (1.1)$$

where $K : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is called the correlation kernel. Such processes arise in random matrix theory to describe eigenvalues of the so-called unitary (invariant) ensembles; see

---

*KTH Royal Institute of Technology, Department of Mathematics,
E-mail: glambert@kth.se. Supported by the grant KAW 2010.0063 from the Knut and Alice Wallenberg Foundation.

theorem 1.1 below and section 3 for further details. There are many other interesting examples of determinantal processes such as random tilings or the positions of non-colliding stochastic diffusions that we will not discuss here. We refer to [46, 24, 20, 2, 36] for various introductions to the general theory and more examples. A fundamental feature of determinantal processes is that all the information about the process is encoded in the correlation kernel. For instance, for unitary ensembles, universality of the local correlations in the bulk of the spectrum follows from the uniform convergence of the rescaled correlation kernel to the sine-kernel, [12, 40, 44]. In this work, we show that at mesoscopic scales, the sine-kernel asymptotics still hold and this leads to the following central limit theorem:

Theorem 1.1. *Let $V : \mathbb{R} \to \mathbb{R}$ be real-analytic such that*

$$\liminf_{|x| \to \infty} \frac{V(x)}{\log(x^2 + 1)} = +\infty, \tag{1.2}$$

and consider the probability measure $\mathbb{P}_N^V = \mathrm{Z}_{V,N}^{-1} e^{-N \operatorname{Tr} V(\mathrm{H})} d\mathrm{H}$ on the space of $N \times N$ Hermitian matrices equipped with the Lebesgue measure $d\mathrm{H}$. If $(\lambda_1, \ldots, \lambda_N)$ denote the eigenvalues of a random matrix H distributed according to $\mathbb{P}_N^V$, then for any $x_0 \in J_V$, any $0 < \alpha < 1$, and for any $f \in C^1(\mathbb{R})$ with compact support, we have

$$\sum_{k=1}^{N} f\big(N^\alpha(\lambda_k - x_0)\big) - \mathbb{E}_N^V\left[\sum_{k=1}^{N} f\big(N^\alpha(\lambda_k - x_0)\big)\right] \underset{N \to \infty}{\Longrightarrow} \mathcal{N}\big(0, \|f\|_{H^{1/2}}^2\big). \tag{1.3}$$

Proof. Section 3.2. □

The condition (1.2) guarantees that $\mathrm{Z}_{V,N} < \infty$, so that the measure $\mathbb{P}_N^V$ is well-defined. This also implies that for large N, the eigenvalue process is supported on a deterministic compact set $\overline{J_V}$ with high probability. Hence, the potential V is confining and the condition $x_0 \in J_V$ means that we zoom in around a point x_0 which lies in the bulk of the spectrum.

In theorem 1.1, the parameter $\alpha \in [0, 1]$ is called the **scale**. Since the eigenvalue density is of order N in the bulk, if $\alpha = 0$, the random variable on the LHS of (1.3) depends on the whole spectrum of H and this regime is called global or macroscopic. On the other hand, if $\alpha = 1$, the rescaled point process $\big\{N(\lambda_k - x_0)\big\}_{k=1}^{N}$ converges to the sine process as the size N of the matrix tends to infinity and this regime is called local or microscopic. Any intermediate scale, $0 < \alpha < 1$, is called **mesoscopic**. Note that in this regime, the limit (1.3) is independent of the potential V, the scale α and x_0. Hence, this establishes universality of fluctuations for a large class of Hermitian random matrix ensembles. The variance in formula (1.3) is given by

$$\|f\|_{H^{1/2}}^2 = \int_{\mathbb{R}} \big|\hat{f}(u)\big|^2 |u| du = \frac{1}{4\pi^2} \iint_{\mathbb{R}^2} \left|\frac{f(x) - f(y)}{x - y}\right|^2 dx dy, \tag{1.4}$$

where $\hat{f}(u) = \int f(x) e^{-i2\pi x u} dx$ denotes the Fourier transform of the test function f. Formula (1.4) defines a complete normed subspace of $L^2(\mathbb{R})$ that we denote by $H^{1/2}(\mathbb{R})$. Most of the work on unitary invariant ensembles has focused on the asymptotics of local or global statistics and we briefly review the main results, further references can be found in the textbooks [12, 40]. Under the assumptions of theorem 1.1, there exists a probability density ϱ_V with compact support $\overline{J_V}$ on $\mathbb{R}$ such that for any f continuous and bounded on $\mathbb{R}$, one has $\mathbb{P}_N^V$ almost surely,

$$\frac{1}{N} \sum_{k=1}^{N} f(\lambda_k) \underset{N \to \infty}{\longrightarrow} \int f(x) \varrho_V(x) dx. \tag{1.5}$$

It means that for large N, the eigenvalues of a random matrix sampled according to $\mathbb{P}_N^V$ are distributed according to the **equilibrium density** ϱ_V. Moreover, it is known that the fluctuations around this equilibrium configuration remain bounded as $N \to \infty$. The precise behavior of linear statistics depends on the support of ϱ_V. In the simplest case, there exists $c_0 \in \mathbb{R}$ and $\ell > 0$ so that

$$J_V = \big(c_0 - \ell, c_0 + \ell\big), \tag{1.6}$$

then the potential V is said to satisfy the **one-cut condition** and there is a CLT: for any $f \in C^2 \cap L^\infty(\mathbb{R})$,

$$\sum_{k=1}^N f\big(\frac{\lambda_k - x_0}{\ell}\big) - N\ell \int f(x)\varrho_V\big(x_0 + \ell x\big)dx \underset{N\to\infty}{\Longrightarrow} \mathcal{N}\big(0, \Sigma(f)^2\big), \tag{1.7}$$

where

$$\Sigma(f)^2 := \frac{1}{4\pi^2} \iint_{[-1,1]^2} \left| \frac{f(x) - f(y)}{x - y} \right|^2 \frac{1 - xy}{\sqrt{1 - x^2}\sqrt{1 - y^2}} dxdy. \tag{1.8}$$

The CLT (1.7) was first proved in [23] when V is a polynomial of even degree using a variational method. We refer to [3] for further developments and to [11, 31] for alternative proofs which also are valid for more general determinantal processes. It is known that (1.6) holds when the potential V is strictly convex on $\mathbb{R}$ and, if $\tilde{V}(x) = V(c_0 + \ell x)$, by considering the ensemble $\mathbb{P}_N^{\tilde{V}}$ instead of $\mathbb{P}_N^V$, we can always assume that $c_0 = 0$ and $\ell = 1$. The one-cut condition is crucial to observe a Gaussian process in the limit. If $\overline{J_V} = \mathrm{supp}(\varrho_V)$ is not connected, then for a generic test function f, the behavior of the linear statistic $\sum f(\lambda_k)$ is quasi-periodic in N and, even though this sequence of random variables is tight, it has no limit as $N \to \infty$, [38]. This complicated behavior is explained by the fact that the numbers of eigenvalues in the different components of J_V fluctuate. Nevertheless, along appropriate subsequences, the asymptotic behavior of the random variable $\sum f(\lambda_k)$ can still be described and it is generally not Gaussian, [4]. On the other hand, it is known that at the local scale, the behavior of the eigenvalue process is independent of the equilibrium density and it is described by the sine process in the bulk. Theorem 1.1 shows that mesoscopic fluctuations are universal as well regardless of the geometry of the equilibrium measure. Actually, this results was first derived heuristically by Pastur in [38] based on the *semi–classical* asymptotic formulae derived in [13] for the orthogonal polynomials with respect to the measure $e^{-NV(x)}dx$ on $\mathbb{R}$.

Theorem 1.1 has the following interpretation when viewing the eigenvalue process as a random measure

$$\Xi_N^{x_0,\alpha} := \sum_{k=1}^N \delta_{N^\alpha(\lambda_k - x_0)}. \tag{1.9}$$

Once centered, $\Xi_N^{x_0,\alpha}$ converges in distribution to a random Gaussian process $\mathfrak{G}$ with covariance structure

$$\mathbb{E}\left[\mathfrak{G}(f)\mathfrak{G}(g)\right] = \int_{\mathbb{R}} \hat{f}(u)\hat{g}(-u)|u|du. \tag{1.10}$$

The random process $\mathfrak{G}$ is called the $H^{1/2}$-Gaussian field, see [22, chapter 1]. Its special feature is that it is scale invariant. If $f_\eta(x) = f(\eta x)$, then $\mathfrak{G}(f_\eta) \sim \mathfrak{G}(f)$ for any $\eta > 0$, as can be seen from (1.10). Heuristically, this explains why this process is expected to describe mesoscopic fluctuations of point processes with strong repulsion such as eigenvalues of random matrices, see the discussion in [47]. In some respect, these ensembles behave like the sine process and this is the main idea behind the proof of

theorem 1.2 below. The mesoscopic correlations can also be guessed from formulae (1.7–1.8). Namely, if $x_0 = 0$ and $\ell = 1$, by a change of variables

$$\Sigma(f_\eta)^2 = \frac{1}{4\pi^2} \iint_{[-\eta,\eta]^2} \left| \frac{f(x) - f(y)}{x - y} \right|^2 \frac{1 - xy/\eta^2}{\sqrt{1 - (x/\eta)^2}\sqrt{1 - (y/\eta)^2}} dxdy,$$

and, if f decays sufficiently fast, we obtain

$$\lim_{\eta \to \infty} \Sigma(f_\eta)^2 = \frac{1}{4\pi^2} \iint_{\mathbb{R}^2} \left| \frac{f(x) - f(y)}{x - y} \right|^2 dxdy = \|f\|_{H^{1/2}}^2.$$

It is natural to investigate whether (1.3) holds under the optimal condition $f \in H^{1/2}(\mathbb{R})$. To the author' knowledge, this question remains open even for the Gaussian Unitary Ensemble (GUE). To some extent, this issue is addressed in section 3.3. If the potential V satisfies the one-cut condition (1.6), we provide bounds for the variance of the random variable $\Xi_N^{x_0,\alpha} f$ which allows us to extend the mesoscopic CLT to e.g. all bounded functions of class $H^{1/2}(\mathbb{R})$ with compact support, see proposition 3.5 below. In an appendix, we provide additional variance estimates valid in the global regime. Once more, the goal is to extend a CLT from [31] from polynomials to more general test functions; see theorem A.4. This result is the analogue for one-cut unitary invariant ensembles of that obtained by Sosoe and Wong in [48, Thm 3] for a class of Hermitian Wigner matrices.

Mesoscopic spectral statistics were first considered in [7, 8] for Hermitian and symmetric Wigner matrices. In particular, the authors proved a result analogous to (1.3) for the GUE and GOE using the resolvent of the matrix as a test function. Their results were recently generalize down to the optimal scale and for a wide class of Hermitian Wigner matrices in [26]. One of the pioneering works on the subject which has been of inspiration for this article is Soshnikov's CLT for eigenvalue statistics of Haar distributed random matrices from the compact groups, [45]. In the case of the unitary group, this point process is known as the Circular Unitary Ensemble (CUE) and it is determinantal with a correlation kernel

$$K_{U_N}(x,y) = \frac{\sin\left(N(x - y)/2\right)}{2\pi \sin\left((x - y)/2\right)}, \qquad \forall x, y \in \mathbb{R}/2\pi\mathbb{Z}, \tag{1.11}$$

and one may interpret the counterpart of (1.3) as a continuous analogue of the Strong Szegő theorem. Theorem 1.1 is also closely related to the main result of [17] which establishes that, at mesoscopic scales, the logarithm of the absolute value of the characteristic polynomial of a GUE matrix converges weakly to a $H = 0$ *fractional Brownian motion* (see definition 4.1) which is a Gaussian generalized function whose correlation kernel has a logarithmic singularity on the diagonal. On the one hand, one can infer from [17, Thm 2.4] a CLT for mesoscopic linear statistics of Schwartz-class test functions. On the other hand, we show in section 4 that one can deduce the result of [17] from a mesoscopic CLT provided that it is valid for sufficiently general test functions. As an important consequence of theorem 3.5, this allows us to generalize the work of Fyodorov et al. to the characteristic polynomials of other one-cut matrix models; see theorem 4.2. This elaborates on the intriguing connection between logarithmically correlated Gaussian processes and random matrix theory which has been an active research direction since the work of [21]. For instance, based on the so-called *freezing transition scenario*, Fyodorov et al. produced interesting conjectures for the distributions of the extreme values of these polynomials and, by analogy, for the extreme values the Riemann Zeta function on some large intervals high up along the critical line, [16, 18]. This also indicates that

the characteristic polynomials of random matrices give raise to regularizations of certain *Gaussian Multiplicative Chaos measures* which have recently played a significant role in some physical theories such as turbulence, disordered systems or Liouville quantum gravity; see [15, 50] and references therein.

The proof of theorem 1.1 is based on the so-called Plancherel-Rotach asymptotics for the orthogonal polynomials (OP) with respect to the weight $e^{-NV(x)}$ on $\mathbb{R}$ derived in [13] and the following general result. For any function $\rho : \mathbb{R} \to \mathbb{R}_+$ which is locally integrable, we let

$$J_\rho := \{t \in \mathbb{R} : \rho(t) > 0 \text{ and } \rho(t) \text{ is continuous}\} \tag{1.12}$$

and, for all $x \in \mathbb{R}$, we define

$$F_\rho(x) := \int_0^x \rho(t)dt. \tag{1.13}$$

We also denote by $C_0^k(\mathbb{R})$ the space of compactly supported real-valued functions with k continuous derivatives on $\mathbb{R}$.

Theorem 1.2. *Consider a determinantal process on $\mathbb{R}$ with a correlation kernel K_N which is locally trace-class. For any $x_0 \in \mathbb{R}$, $\alpha \in [0,1]$ and $f \in C_0(\mathbb{R})$, we consider the linear statistic*

$$\Xi_N^{x_0,\alpha} f = \sum f\big(N^\alpha(\lambda_k - x_0)\big),$$

where the sum is over the point configuration $\{\lambda_k\}$ of the process. Assume that there exists a function $\rho : \mathbb{R} \to \mathbb{R}_+$ such that $x_0 \in J_\rho$ and uniformly for all $x, y \in [x_0 - \epsilon_N, x_0 + \epsilon_N]$,

$$K_N(x,y) = \frac{\sin \pi N\big(F_\rho(x) - F_\rho(y)\big)}{\pi(x-y)} + \underset{N \to \infty}{O(1)}, \tag{1.14}$$

Then, if $\lim_{N \to \infty} \epsilon_N N^\alpha = +\infty$ for some $\alpha < 1$, we obtain for any $f \in C_0^1(\mathbb{R})$,

$$\Xi_N^{x_0,\alpha} f - \mathbb{E}\left[\Xi_N^{x_0,\alpha} f\right] \underset{N \to \infty}{\Longrightarrow} \mathcal{N}\left(0, \|f\|_{H^{1/2}}^2\right) \tag{1.15}$$

and for any $f \in C_0(\mathbb{R})$,

$$\Xi_N^{x_0,1} f \underset{N \to \infty}{\Longrightarrow} \Xi_{\rho(x_0)}^{\sin} f. \tag{1.16}$$

Proof. Section 2. □

On the RHS of (1.16), $\Xi_\nu^{\sin}$ denotes the sine process with density $\nu > 0$, i.e. the determinantal process on $\mathbb{R}$ with the correlation kernel

$$K_\nu^{\sin}(\xi, \zeta) = \frac{\sin[\pi\nu(\xi - \zeta)]}{\pi(\xi - \zeta)}. \tag{1.17}$$

In the local regime, (1.16) implies the convergence of the process $\Xi_N^{x_0,1}$ to the sine process. As we already emphasized, this behavior is universal for Hermitian ensembles. In the context of theorem 1.1, it was proved in [39, 13, 32, 33]. Assuming that the kernel K_N is locally trace-class is standard, it means that for any function $f \in L^\infty(\mathbb{R})$ with compact support, the integral operator $h \mapsto \int h(x)K_N(\cdot, x)f(x)dx$ is trace-class on $L^2(\mathbb{R})$ and it implies that the linear statistic $\sum f(\lambda_k)$ has a finite Laplace transform and its cumulants are well-defined, see formula (2.4) below. In the context of theorem 1.1 or, in greater generality, for the orthogonal polynomials discussed in section 3, their correlation kernels define some rank N projection operators on $L^2(\mathbb{R})$. However, note that in theorem 1.2, we do not assume that the kernel K_N is reproducing, nor have finite rank. Thus, our result still applies if the configuration $\{\lambda_k\}$ has a random number of points or infinitely many.

It is obvious that the CUE kernel (1.11) has an asymptotic expansion of the form (1.14) with $\rho = 1/2\pi$ on $\mathbb{R}$, so that Soshnikov's CLT is a special case of theorem 1.2. In fact, our main observation is that, if the correlation kernel satisfies (1.14), then we can still apply Soshnikov's method to prove a central limit theorem. In particular, the fact that the limiting process is Gaussian follows from the Main combinatorial Lemma of [45], theorem 2.7 below. For any determinantal process within the sine process universality class, it is plausible that the asymptotics (1.14) hold at sufficiently small scales, so that theorem 1.2 explains the appearance of the $H^{1/2}$-Gaussian field in this context. However, the general mechanism behind universality of mesoscopic fluctuations is still far from being understood. In particular, it would be interesting to understand further the connection between random matrix theory and logarithmically correlated Gaussian fields as discussed in [17]. Within other symmetry classes and for Dyson's β-ensembles, mesoscopic fluctuations are also conjectured to be described by the $H^{1/2}$-Gaussian field. For instance, this has been rigorously established for the Gaussian β-Ensembles in [5], for random matrices from the special orthogonal and symplectic groups in [45] as well as in number theory when considering mesoscopic linear statistics of the zeros of the Riemann ζ function [6, 42].

In this article, we focus on applications to random matrices, but theorem 1.2 should be useful to investigate mesoscopic fluctuations for other instances of determinantal processes. Based on the Riemann-Hilbert formulation of [14], it is possible to derive very precise asymptotics for the orthogonal polynomials and the Christoffel-Darboux kernels for large families of measures on $\mathbb{R}$. These results combined with theorem 1.2 allow us to prove universality of the mesoscopic correlations for a extensive pool of unitary invariant random matrix ensembles. For the GUE, it is possible to derive the asymptotics (1.14) using only the Plancherel-Rotach asymptotics for the Hermite polynomials, [41], and this leads to a rather elementary proof of theorem 1.1; see section 3.1. Finally, let us mention that for non-varying compactly supported measures, an alternative approach to universality has been developed in [11, 10]. This approach relies on the asymptotics of the OP recurrence coefficients rather than on that of the correlation kernel and is further discussed in remarks 3.1 and 3.14 below. In particular, we discuss applications to the so-called modified Jacobi ensembles in section 3.4 and, by using theorem 1.2, we provide a new and perhaps simpler proof of [10, Thm 1.1].

The rest of the paper is organized as follows. In section 2, we review the cumulant method introduced in [45] to study linear statistics of determinantal processes and we prove theorem 1.2. The proof relies on ideas developed in [25]. In section 3, we begin by a brief introduction to the theory of unitary invariant ensembles focusing on the orthogonal polynomials method. The core of this section is dedicated to show that the asymptotics (1.14) are valid for the correlation kernels of a large class of orthogonal polynomial ensembles. In section 3.1, we start by giving an elementary derivation of the asymptotics of the GUE kernel based on the Plancherel-Rotach asymptotics for the Hermite polynomials, [41]. In section 3.2, we review the Riemann-Hilbert formulation for the orthogonal polynomials and the fundamental results of [13], hence completing the proof of theorem 1.1. In section 3.4, we prove the counterpart of theorem 1.1 for the so-called Modified Jacobi ensembles. In section 3.3, focusing on the one-cut regime, we provide some estimates for the variance of linear statistics. This is necessary in order to extend the mesoscopic CLT to a wider class of test functions and to generalize the result of [17] to the characteristic polynomials of other matrix models. The proof is given in section 4. Finally, in the appendix A, we further generalize the variance estimates obtained in section 3.3 so that there are valid in the global regime for general test functions with at most polynomial growth.

In the sequel, we use the notation $b_N(x) = O_\epsilon(a_N)$ if there exists a set A_ϵ such that the estimate is uniform for all $x \in A_\epsilon$ (similarly for o_ϵ). Moreover, C denotes a positive constant which may change from line to line.

2   Proof of theorem 1.2

We consider a family of determinantal processes on $\mathbb{R}$ with correlation kernels K_N which depend on a parameter $N > 0$ and satisfy (1.14) for a given function $\rho : \mathbb{R} \to \mathbb{R}^+$ and we let $J = J_\rho$ and $F = F_\rho$ according to (1.12), respectively (1.13). We want to study the laws as $N \to \infty$ of the random variables

$$\Xi_N^{x_0, \alpha} f = \sum f\big(N^\alpha(\lambda_k - x_0)\big), \tag{2.1}$$

where $\{\lambda_k\}$ is a configuration of the determinantal process with kernel K_N, $f \in C_0^1(\mathbb{R})$ is a test function such that $\mathrm{supp}(f) \subset [-L, L]$, and $x_0 \in J$. For any real-valued random variable Z with a well-defined Laplace transform, its cumulants $\mathrm{C}^n[Z]$ are defined by the generating function:

$$\log \mathbb{E}\left[e^{tZ}\right] = \sum_{n=1}^\infty \mathrm{C}^n[Z]\frac{t^n}{n!}. \tag{2.2}$$

Observe that $\mathrm{C}^1[Z] = \mathbb{E}[Z]$ and the higher-order cumulants do not depend on $\mathbb{E}[Z]$. In particular, we have $\mathrm{C}^2[Z] = \mathrm{Var}[Z]$ and, if Z is Gaussian, $\mathrm{C}^n[Z] = 0$ for all $n \geq 3$. Hence, to prove the CLT (1.15), it is enough to show that

$$\lim_{N \to \infty} \mathrm{C}^2\left[\Xi_N^{x_0, \alpha} f\right] = \|f\|_{H^{1/2}}^2 \quad \text{and} \quad \lim_{N \to \infty} \mathrm{C}^n\left[\Xi_N^{x_0, \alpha} f\right] = 0 \quad \forall n \geq 3. \tag{2.3}$$

Using formula (1.1), one can compute moments and cumulants of linear statistics of determinantal processes. In particular, it was proved in [45] that, if the correlation kernel is locally trace-class and $f \in C_0(\mathbb{R})$, then for any $n \in \mathbb{N}$,

$$\mathrm{C}^n\left[\sum f(\lambda_k)\right] = \sum_{\ell=1}^n \frac{(-1)^{\ell+1}}{\ell} \sum_{\substack{m_1,\ldots,m_\ell \geq 1 \\ m_1 + \cdots + m_\ell = n}} \frac{n!}{m_1! \cdots m_\ell!} \mathrm{Tr}\left[f^{m_1} K_N \cdots f^{m_\ell} K_N\right], \tag{2.4}$$

where

$$\mathrm{Tr}\left[f^{m_1} K \cdots f^{m_\ell} K\right] = \int_{\mathbb{R}^\ell} f(x_1)^{m_1} K(x_1, x_2) \cdots f(x_\ell)^{m_\ell} K(x_\ell, x_1) d^\ell x. \tag{2.5}$$

Let us denote for for all $\xi, \zeta \in \mathbb{R}$,

$$\widetilde{K}_N(\xi, \zeta) = N^{-\alpha} K_N\big(x_0 + N^{-\alpha}\xi, x_0 + N^{-\alpha}\zeta\big).$$

Then, by (2.1), (2.4) and a change of variables, we get

$$\mathrm{C}^n\left[\Xi_N^{x_0, \alpha} f\right] = \sum_{\ell=1}^n \frac{(-1)^{\ell+1}}{\ell} \sum_{\substack{m_1,\ldots,m_\ell \geq 1 \\ m_1 + \cdots + m_\ell = n}} \frac{n!}{m_1! \cdots m_\ell!} \mathrm{Tr}\left[f^{m_1} \widetilde{K}_N \cdots f^{m_\ell} \widetilde{K}_N\right]. \tag{2.6}$$

It was observed in [25] that, if the correlation kernel K_N satisfy the uniform asymptotics (1.14), then we can relate its cumulants to those of the sine process as $N \to \infty$. In particular, lemma 2.1 which is the main ingredient to prove proposition 2.2 below is a straightforward adaptation of [25, Lemma 2.6].

Lemma 2.1. *Let us consider two families of kernels $(S_N)_{N>0}$ and $(\tilde{S}_N)_{N>0}$ on $\mathbb{R}$. If there exist $\alpha > 0$, $L > 0$, and a function $\Gamma_N : \mathbb{R} \to \mathbb{R}^+$ such that when N is sufficiently large:*

(1) for all $x, y \in [-L, L]$, $\left| \tilde{S}_N(x, y) - S_N(x, y) \right| \leq C_L N^{-\alpha}$.

(2) for all $x, y \in [-L, L]$, $\left| S_N(x, y) \right| \leq \Gamma_N(x - y)$.

(3) $\displaystyle \int_{-2L}^{2L} \Gamma_N(s) ds \leq C \log(LN)$.

Then, for all $\epsilon > 0$, $\ell \in \mathbb{N}$, and for any functions $f_{N,1}, \ldots, f_{N,\ell}$ with support in $[-L, L]$ such that $\sup \left\{ \|f_{N,k}\|_\infty : k = 1, \ldots, \ell \right\} \leq C_\ell$, one has

$$\mathrm{Tr}[f_{N,1} \tilde{S}_N \cdots f_{N,\ell} \tilde{S}_N] = \mathrm{Tr}[f_{N,1} S_N \cdots f_{N,\ell} S_N] + \underset{N \to \infty}{O} \left(N^{-\alpha + \epsilon} \right).$$

Let us define

$$S_N(\xi, \zeta) = \frac{\sin \left[\pi N \left(F(x_0 + \xi N^{-\alpha}) - F(x_0 + \zeta N^{-\alpha}) \right) \right]}{\pi (\xi - \zeta)}$$

and

$$\Gamma_N(\xi - \zeta) = \begin{cases} C_0 N^{1-\alpha} & \text{if } |\xi - \zeta|^{-1} \leq \frac{1}{N} \\ 1/\pi |\xi - \zeta| & \text{if } |\xi - \zeta|^{-1} > \frac{1}{N} \end{cases}. \tag{2.7}$$

When $\epsilon_N N^\alpha \to +\infty$, the asymptotics (1.14) imply that the kernels $\tilde{K}_N$ and S_N satisfy condition (1) of lemma 2.1 for any $L > 0$. Moreover, we claim that the conditions (2) and (3) hold as well so that we obtain for any $m_1, \ldots, m_\ell \in \mathbb{N}$,

$$\mathrm{Tr} \left[f^{m_1} \tilde{K}_N \cdots f^{m_\ell} \tilde{K}_N \right] = \mathrm{Tr} \left[f^{m_1} S_N \cdots f^{m_\ell} S_N \right] + \underset{N \to \infty}{O} \left(N^{-\alpha + \epsilon} \right). \tag{2.8}$$

By (2.7), it is straightforward to check that for any $C_0 > 0$,

$$\int_{-2L}^{2L} \Gamma_N(s) ds \leq \log(LN) + O(1),$$

so that condition (3) holds. To check condition (2), note that by definition of J, for any $0 < \epsilon_0 < 1/2$, there exists $\delta_0 > 0$ so that the density ρ is continuous on $[x_0 - \delta_0, x_0 + \delta_0]$ and for all $|x - x_0| < \delta_0$,

$$1 - \epsilon_0 \leq \frac{\rho(x)}{\rho(x_0)} \leq 1 + \epsilon_0. \tag{2.9}$$

If $N^\alpha > L/\delta_0$ and $C_0 \geq \rho(x_0)(1 + \epsilon_0)$, this implies that for all $\xi, \zeta \in [-L, L]$,

$$\left| F(x_0 + \xi N^{-\alpha}) - F(x_0 + \zeta N^{-\alpha}) \right| = N^{-\alpha} \left| \int_\xi^\zeta \rho(x_0 + s N^{-\alpha}) ds \right| \leq C_0 N^{-\alpha} |\xi - \zeta|.$$

Thus, if we use the trivial bound $|\sin x| \leq |x| \vee 1$, by (2.7), we conclude that $\left| S_N(\xi, \zeta) \right| \leq \Gamma_N(\xi - \zeta)$. The map F is continuous non-decreasing, so it has a generalized inverse:

$$G(x) = \inf \left\{ t \in \mathbb{R} : F(t) \geq x \right\}. \tag{2.10}$$

In the sequel, we will assume that δ_0 is sufficiently small, so that (2.9) holds and the map G is continuously differentiable on $[F(x_0) - \delta_0, F(x_0) + \delta_0]$ with

$$G'(x) = \frac{1}{F'(G(x))} = \frac{1}{\rho(G(x))}. \tag{2.11}$$

Finally, recall that the sine process Ξ_ν^{sin} is the determinantal process on $\mathbb{R}$ with a correlation kernel K_ν^{sin} given by (1.17).

Proposition 2.2. *Let $f \in C_0(\mathbb{R})$ and $\alpha \in (0, 1]$. We have for any $n \geq 1$,*

$$\lim_{N \to \infty} \mathrm{C}^n \left[\Xi_N^{\alpha, x_0} f \right] = \lim_{N \to \infty} \mathrm{C}^n \left[\Xi_{\nu_N}^{\sin} f_N \right],$$

where $\nu_N = N^{1-\alpha} \rho(x_0)$ and

$$f_N(x) = f \left(N^\alpha \left\{ G \left(F(x_0) + \rho(x_0) \frac{x}{N^\alpha} \right) - x_0 \right\} \right). \tag{2.12}$$

Remark 2.3. Observe that for any $0 < \epsilon_0 < 1/2$, by (2.9) and (2.11), if $N^\alpha > 2\rho(x_0)L/\delta_0$, then for all $x \in [-2L, 2L]$,

$$\frac{1}{1 + \epsilon_0} \leq \rho(x_0) G' \left(F(x_0) + \rho(x_0) \frac{x}{N^\alpha} \right) \leq \frac{1}{1 - \epsilon_0}.$$

If we integrate this estimate, since $f \in C_0([-L, L])$, this implies that the function $f_N \in C_0(\mathbb{R})$ has support in $[-L_0, L_0]$ where $L_0 = L(1 + \epsilon_0)$.

Proof of proposition 2.2. We fix $m_1, \ldots, m_\ell \in \mathbb{N}$ and we suppose that N^α is sufficiently large. We can make the change of variables

$$y_k = \frac{N^\alpha}{\rho(x_0)} \left\{ F(x_0 + N^{-\alpha} x_k) - F(x_0) \right\} \tag{2.13}$$

in the formula

$$\mathrm{Tr} \left[f^{m_1} S_N \cdots f^{m_\ell} S_N \right] = \int_{-L}^{L} \cdots \int_{-L}^{L} \prod_{k=1}^{\ell} f(x_k)^{m_k} \frac{\sin \left[\pi N^{1-\alpha} \rho(x_0)(y_k - y_{k+1}) \right]}{\pi(x_k - x_{k+1})} d^\ell x.$$

If we let

$$g(y) = G(F(x_0) + \rho(x_0)y) - x_0, \tag{2.14}$$

and f_N be given by (2.12), according to remark 2.3, this leads to

$$\mathrm{Tr} \left[f^{m_1} S_N \cdots f^{m_\ell} S_N \right] = \int_{-L_0}^{L_0} \cdots \int_{-L_0}^{L_0} \prod_{k=1}^{\ell} f_N(y_k)^{m_k} \widetilde{S}_N(y_k, y_{k+1}) d^\ell y$$

$$= \mathrm{Tr} \left[f_N^{m_1} \tilde{S}_N \cdots f_N^{m_\ell} \tilde{S}_N \right], \tag{2.15}$$

where

$$\widetilde{S}_N(y, z) := \frac{g'(yN^{-\alpha}) \sin \left[\pi \rho(x_0) N^{1-\alpha} (y - z) \right]}{\pi N^\alpha \left\{ g(yN^{-\alpha}) - g(zN^{-\alpha}) \right\}}.$$

For any $0 < \alpha \leq 1$, a Taylor expansion in (2.14) yields for all $y, z \in [-L_0, L_0]$,

$$g' \left(yN^{-\alpha} \right)^{-1} N^\alpha \left\{ g \left(yN^{-\alpha} \right) - g \left(zN^{-\alpha} \right) \right\} = (y - z) \left\{ 1 + O \left(|y - z| N^{-\alpha} \right) \right\}.$$

This implies that uniformly for all $y, z \in [-L_0, L_0]$,

$$\widetilde{S}_N(x, y) = \frac{\sin \left[\pi \nu_N (y - z) \right]}{y - z} + O \left(N^{-\alpha} \right),$$

where $\nu_N = N^{1-\alpha} \rho(x_0)$. Thus, the kernels $\widetilde{S}_N$ and $K_{\nu_N}^{\sin}$ also satisfy condition (1) of lemma 2.1. Moreover, if Γ_N is given by (2.7) with $C_0 = \rho(x_0)$, the kernel $K_{\nu_N}^{\sin}$ also satisfies condition (2). Therefore, since the functions $f_{N,k} = f_N^{m_k}$ have support in $[-L_0, L_0]$ and

$$\sup \left\{ \|f_{N,k}\|_\infty : k = 1, \ldots, \ell \right\} \leq \|f\|_\infty^{m_1 \vee \cdots \vee m_\ell},$$

by lemma 2.1, we obtain

$$\text{Tr}\left[f_N^{m_1}\widetilde{S}_N\cdots f_N^{m_\ell}\widetilde{S}_N\right] = \text{Tr}\left[f_N^{m_1}K_{\nu_N}^{\sin}\cdots f_N^{m_\ell}K_{\nu_N}^{\sin}\right] + O\left(N^{-\alpha+\epsilon}\right). \tag{2.16}$$

If we combine formulae (2.8), (2.15) and (2.16), we have proved that for any $m_1,\ldots,m_\ell \in \mathbb{N}$,

$$\text{Tr}\left[f^{m_1}\widetilde{K}_N\cdots f^{m_\ell}\widetilde{K}_N\right] = \text{Tr}\left[f_N^{m_1}K_{\nu_N}^{\sin}\cdots f_N^{m_\ell}K_{\nu_N}^{\sin}\right] + O\left(N^{-\alpha+\epsilon}\right). \tag{2.17}$$

Since, by formula (2.6), the cumulants of the random variable $\Xi_N^{x_0,\alpha}f$ are linear combinations of such traces, we conclude by (2.17) that for any $n \geq 1$,

$$\text{C}^n\left[\Xi_N^{\alpha,x_0}f\right] = \text{C}^n\left[\Xi_{\nu_N}^{\sin}f_N\right] + O\left(N^{-\alpha+\epsilon}\right). \tag{2.18}$$

$\square$

Remark 2.4. In the physics literature, the change of variables (2.13) is known as *unfolding the spectrum* since it corresponds to rescaling the eigenvalue process so that it has a constant density ν_N in a mesoscopic range around the point $x_0 \in J_\rho$. Note that in formula (1.14), if the density ρ is smooth and $\rho(x_0) \neq 0$, a Taylor expansion of the function F_ρ shows that we recover the classical sine-kernel asymptotics in the regime $\alpha > 1/2$. Namely, if $\xi,\zeta \in [-L,L]$ and we make the change of variables $x = x_0 + \xi/N^\alpha$ and $y = x_0 + \zeta/N^\alpha$, then

$$\frac{\sin \pi N\left(F_\rho(x)-F_\rho(y)\right)}{\pi(x-y)} = \frac{\sin\left[\pi N^{1-\alpha}\rho(x_0)(\xi-\zeta)\right]}{\pi(x-y)} + \underset{N\to\infty}{O_L}\left(\left|\frac{\xi-\zeta}{x-y}\right|N^{1-2\alpha}\right) \tag{2.19}$$

$$N^{-\alpha}K_N(x,y) = \frac{\sin\left[\pi N^{1-\alpha}\rho(x_0)(\xi-\zeta)\right]}{\pi(\xi-\zeta)} + \underset{N\to\infty}{O_L}\left(N^{1-2\alpha}\right).$$

Hence, at sufficiently small scales, the fact that the eigenvalues are not uniformly distributed is not relevant and, if the integrated density of states F_ρ is smooth on J_ρ, we can deduce proposition 2.2 directly from lemma 2.1 without making the change of variables (2.13).

First, let us apply proposition 2.2 to obtain local correlations. In the regime $\alpha = 1$, for any $n \geq 1$,

$$\lim_{N\to\infty}\text{C}^n\left[\Xi_N^{1,x_0}f\right] = \lim_{N\to\infty}\text{C}^n\left[\Xi_{\rho(x_0)}^{\sin}f_N\right]. \tag{2.20}$$

By (2.11), a Taylor expansion of the map G yields for all $|x| < L_0$,

$$\lim_{N\to\infty}N^\alpha\left\{G\left(F(x_0) + \rho(x_0)\frac{x}{N^\alpha}\right) - x_0\right\} = x. \tag{2.21}$$

By remark 2.3, the function f_N has support in $[-L_0,L_0]$ and by continuity of f, the limit (2.21) implies that $\lim_{N\to\infty} f_N(x) = f(x)$ for all $x \in \mathbb{R}$. Hence, by the dominated convergence theorem, we get

$$\lim_{N\to\infty}\text{C}^n\left[\Xi_{\rho(x_0)}^{\sin}f_N\right] = \text{C}^n\left[\Xi_{\rho(x_0)}^{\sin}f\right].$$

By (2.20), this proves that $\lim_{N\to\infty}\text{C}^n\left[\Xi_N^{1,x_0}f\right] = \text{C}^n\left[\Xi_{\rho(x_0)}^{\sin}f\right]$ for any $f \in C_0(\mathbb{R})$ and (1.16) follows from the fact that compactly supported linear statistics of the sine process are characterized by their cumulants.

We now turn to the proof of (1.15) in the mesoscopic regime, $0 < \alpha < 1$. The argument is different because, in formula (2.18), the density of the sine-process $\nu_N \to \infty$ as $N \to \infty$. A relevant result in this regime is a CLT due to Soshnikov for the sine process:

Theorem 2.5 (Thm 4, [47])**.** *For any function $f \in H^{1/2}(\mathbb{R})$, as $\nu \to \infty$,*

$$\Xi_\nu^{\sin} f - \mathbb{E}\left[\Xi_\nu^{\sin} f\right] \;\Rightarrow\; \mathcal{N}\left(0, \|f\|_{H^{1/2}}\right).$$

The proof is based on Fourier analysis and a combinatorial argument given in the article [45]. Although the original proof is given for Schwartz functions, using a density argument, it is not difficult to extend Soshnikov's CLT to all test functions in the Sobolev space $H^{1/2}(\mathbb{R})$. In order to deduce theorem 1.2 from proposition 2.2, we see that it suffices to extend the proof of theorem 2.5 to deal with test functions f_N of the form (2.12). In order to proceed, we need to recall two key lemmas from [45]. For any tuple $m \in \mathbb{N}^\ell$, we define

$$\Upsilon_n(u_1, \ldots, u_n) = \sum_{\ell=1}^n \frac{(-1)^{\ell+1}}{\ell} \sum_{\substack{m_1,\ldots,m_\ell \geq 1 \\ m_1 + \cdots + m_\ell = n}} \frac{n!}{m_1! \cdots m_\ell!} \max_{1 \leq i \leq \ell}\{u_1 + \cdots + u_{m_1 + \cdots + m_i}\}. \quad (2.22)$$

Lemma 2.6 ([45])**.** *There exists a constant $C_n > 0$ which depends only on $n \geq 2$ such that for any $\nu > 0$ and any function $f \in L^1(\mathbb{R})$,*

$$\left| \mathbb{C}^n\left[\Xi_\nu^{sin} f\right] + 2 \int_{\mathbb{R}_0^n} \Re\left\{\prod_i \hat{f}(u_i)\right\} \Upsilon_n(u_1, \ldots, u_n) d^{n-1}u \right|$$

$$\leq C_n \int_{\mathcal{A}_\nu^n} \left| \prod_i \hat{f}(u_i) \right| (|u_1| + \cdots |u_n|) d^{n-1}u,$$

where $\mathbb{R}_0^n = \left\{u \in \mathbb{R}^n : u_1 + \cdots + u_n = 0\right\}$ and $\mathcal{A}_\nu^n = \left\{u \in \mathbb{R}_0^n : |u_1| + \cdots + |u_n| > \nu\right\}$.

Lemma 2.7 (Main Combinatorial lemma, [45])**.** *For any $u \in \mathbb{R}_0^n$,*

$$\sum_{\sigma \in \mathbb{S}_n} \Upsilon_n\left(u_{\sigma(1)}, \ldots, u_{\sigma(n)}\right) = \begin{cases} |u_1| & \text{if } n = 2 \\ 0 & \text{if } n > 2 \end{cases}.$$

If $g \in C_0^1(\mathbb{R})$, we define

$$\|g\|_{H^1}^2 = \int_{\mathbb{R}} |\hat{g}(u)|^2 |u|^2 du = \frac{1}{4\pi^2} \int_{\mathbb{R}} |g'(x)|^2 dx. \quad (2.23)$$

We will also need the following result.

Lemma 2.8. *If $f \in C_0^1(\mathbb{R})$ and the function f_N is given by (2.12), then*

$$\lim_{N \to \infty} \|f_N - f\|_{H^1} = 0.$$

Proof. Since $G \in C^1\left([F(x_0) - \delta_0, F(x_0) + \delta_0]\right)$, by remark 2.3, if $N^\alpha > 2\rho(x_0)L/\delta_0$, the functions f_N are continuously differentiable on $\mathbb{R}$ and

$$f_N'(x) = \rho(x_0)G'\left(F(x_0) + \rho(x_0)\frac{x}{N^\alpha}\right) f'\left(N^\alpha\left\{G\left(F(x_0) + \rho(x_0)\frac{x}{N^\alpha}\right) - x_0\right\}\right).$$

Then, by the triangle inequality,

$$\left|f_N'(x) - f'(x)\right| \leq \|f'\|_\infty \left|\rho(x_0)G'\left(F(x_0) + \rho(x_0)\frac{x}{N^\alpha}\right) - 1\right| \quad (2.24)$$

$$+ \left|f'\left(N^\alpha\left\{G\left(F(x_0) + \rho(x_0)\frac{x}{N^\alpha}\right) - x_0\right\}\right) - f'(x)\right|.$$

First note that, by (2.21) and the continuity of f',

$$\lim_{N \to \infty} \left|f'\left(N^\alpha\left\{G\left(F(x_0) + \rho(x_0)\frac{x}{N^\alpha}\right) - x_0\right\}\right) - f'(x)\right| = 0. \quad (2.25)$$

Second, by remark 2.3, we have

$$\lim_{N \to \infty} \left| \rho(x_0) G' \left(F(x_0) + \rho(x_0) \frac{x}{N^\alpha} \right) - 1 \right| = 0. \tag{2.26}$$

In the end, by the dominate convergence and the estimates (2.24–2.26), we conclude that

$$\lim_{N \to \infty} \|f_N - f\|_{H^1} = \lim_{N \to \infty} \frac{1}{4\pi^2} \int_{-L_0}^{L_0} |f'_N(x) - f'(x)| = 0. \qquad \square$$

Observe that, if $g \in C_0^1(\mathbb{R})$, according to (1.4) and (2.23),

$$\|g\|_{H^{1/2}}^2 \le \|\hat{g}\|_\infty^2 + \|g\|_{H^1}^2$$
$$\le \|g\|_{L^1}^2 + \|g\|_{H^1}^2.$$

By (2.21) and the dominated convergence theorem, we get $\lim_{N \to \infty} \|f_N - f\|_{L^1} = 0$. Thus, by lemma 2.8, we obtain for any $f \in C_0^1(\mathbb{R})$,

$$\lim_{N \to \infty} \|f_N - f\|_{H^{1/2}} = 0. \tag{2.27}$$

For now, let us also assume that, with $\nu_N = N^\alpha \rho(x_0)$ and $\mathcal{A}_\nu^n$ defined in lemma 2.6, one has

$$\lim_{N \to \infty} \int_{\mathcal{A}_{\nu_N}^n} \left| \prod_i \hat{f}_N(u_i) \right| (|u_1| + \cdots |u_n|) d^{n-1} u = 0. \tag{2.28}$$

Then one gets for any $n \ge 2$,

$$\lim_{N \to \infty} \mathrm{C}^n \left[\Xi_{\nu_N}^{\sin} f_N \right] = 2 \lim_{N \to \infty} \int_{\mathbb{R}_0^n} \Re \left\{ \prod_i \hat{f}_N(u_i) \right\} \Upsilon_n(u_1, \ldots, u_n) d^{n-1} u.$$

Since f is real-valued and $\Upsilon_2(u, -u) = |u|/2$, by lemma 2.7, this implies that

$$\lim_{N \to \infty} \mathrm{C}^n \left[\Xi_{\nu_N}^{\sin} f_N \right] = \begin{cases} \lim_{N \to \infty} \int_{\mathbb{R}} \left| \hat{f}_N(u) \right|^2 |u| du & \text{if } n = 2 \\ 0 & \text{if } n > 2 \end{cases}.$$

By proposition 2.2 and (2.27), we conclude that for any $f \in C_0^1(\mathbb{R})$,

$$\lim_{N \to \infty} \mathrm{C}^n \left[\Xi_N^{\alpha, x_0} f \right] = \begin{cases} \|f\|_{H^{1/2}}^2 & \text{if } n = 2 \\ 0 & \text{if } n > 2 \end{cases}.$$

A special case of the limit (2.28) was computed in [25, proposition 4.13]. The proof relies on lemma 2.8 and it is straightforward to generalize the argument of [25] to obtain (2.28). Hence, by (2.3), the CLT (1.15) holds for any $x_0 \in J$ and $f \in C_0^1(\mathbb{R})$.

3 Unitary invariant ensembles

The most well-known probability measure on the space of $N \times N$ Hermitian matrices is the Gaussian Unitary Ensemble:

$$\mathbb{P}_N^{\mathrm{GUE}} = \mathrm{Z}_N^{-1} e^{-2N \operatorname{Tr} \mathrm{H}^2} d\mathrm{H}. \tag{3.1}$$

In this section, we will consider various generalizations of the GUE of the form:

$$\mathbb{P}_N^\omega = \mathrm{Z}_{\omega, N}^{-1} e^{\operatorname{Tr} \log \omega(\mathrm{H})} d\mathrm{H}, \tag{3.2}$$

where the function $\omega : \mathbb{R} \to [0, +\infty)$ is upper-semicontinuous and such that for all $k \geq 0$,

$$\int |x|^k \omega(x) dx < \infty. \tag{3.3}$$

This condition implies that the partition function $Z_{\omega,N} < \infty$ so that the measure is well-defined. For scaling reasons, the weight ω might also depend on the dimension N even though we will not emphasize this to keep our notations as simple as possible. The matrix $\log \omega(\mathrm{H})$ is defined by functional calculus and the measure (3.2) is invariant under the transformation $\mathrm{H} \mapsto \mathrm{UHU}^*$ for any $\mathrm{U} \in \mathcal{U}(N)$. Hence, the name unitary invariant ensemble. In particular, if we use the spectral decomposition of H, under $\mathbb{P}_N^\omega$, the eigenvectors are independent of the spectrum Λ and $\Lambda = \{\lambda_1, \ldots, \lambda_N\}$ has a joint density on $\mathbb{R}^N$ which is given by

$$\mathcal{P}_N^\omega(x_1, \ldots, x_N) = Z_{\omega,N}^{-1} \det_{N \times N} \left[x_j^{k-1} \right] \det_{N \times N} \left[x_j^{k-1} \omega(x_j) \right]. \tag{3.4}$$

In order to analyse the behavior of the eigenvalues, in [37], Gaudin and Metha introduced a method based on the orthogonal polynomials with respect to the measure $\omega(x)dx$ on $\mathbb{R}$. The condition (3.3) guarantees that these polynomials exist and we define for any $k \geq 0$,

$$\pi_k(x) = x^k + \alpha_k x^{k-1} + \cdots \qquad \text{and} \qquad \int \pi_k(x) \pi_j(x) \omega(x) dx = \gamma_k^{-2} \mathbb{1}_{k=j}. \tag{3.5}$$

Then, it follows from formula (3.4) that the eigenvalues density is

$$\mathcal{P}_N^\omega(x_1, \ldots, x_N) = \frac{1}{N!} \det_{N \times N} \left[K_N^\omega(x_j, x_k) \right], \tag{3.6}$$

where

$$K_N^\omega(x, y) = \gamma_{N-1}^2 \frac{\pi_N(x)\pi_{N-1}(y) - \pi_{N-1}(x)\pi_N(y)}{x - y} \sqrt{\omega(x)\omega(y)}. \tag{3.7}$$

Formulae (3.6–3.7) implies that the eigenvalue point process Λ is determinantal with correlation kernel K_N^ω in the sense of (1.1). These facts are well-known and we refer to e.g. [12, 27] for an introduction to the subject. By theorem 1.2, this reduces the question of universality of mesoscopic fluctuations for the ensembles (3.2) to obtain precise asymptotics for the OPs with respect to the measure $\omega(x)dx$.

Remark 3.1. Beyond the context of random matrix theory, one may consider the determinantal process (3.6) associated with a more general measure, e.g. certain discrete measures corresponds to random tilings models, [27]. These processes are known as orthogonal polynomial ensembles and significant research developments have focused on proving the sine-process universality at the local scale, see [35, 43] and reference therein. At mesoscopic scales, another universality result just appeared in [10] from which the authors deduced a weaker version of theorem 3.12 below. Instead of working with the correlation kernel of the process, Breuer and Duits reformulate the cumulant problem in terms of the so-called Jacobi matrix of the measure $\omega(x)dx$ and this reduces the question of universality to controlling the asymptotics of the recurrence coefficients which define the OPs. The drawback of this method is that, for technical reasons, it fails when the reference measure varies with the dimension N, like in the context of theorem 1.1 where $\omega(x) = e^{-NV(x)}$. However, it requires only the asymptotics of the recurrence coefficients and applies to discrete or singular measures where the asymptotics of the correlation kernel can be difficult to derive.

Under general conditions and provided that the weight ω is suitably normalized as $N \to \infty$, see [9, 19], it is known that there is a Law of Large numbers:

$$\frac{1}{N} \sum_{k=1}^{N} f(\lambda_k) \underset{N \to \infty}{\longrightarrow} \int f(x) d\mu_\omega(x) dx, \qquad \mathbb{P}_N^\omega\text{–almost surely,} \qquad (3.8)$$

where μ_ω is called the equilibrium measure and it has compact support. In the following, we will suppose that it is absolutely continuous: $d\mu_\omega = \varrho_\omega(x) dx$. The equilibrium density ϱ_ω plays a fundamental role in the non-linear steepest descent introduced in [14, 12] and therefore appears in the asymptotics of the Christoffel-Darboux kernel. Namely, we will show below, based on the results of [13, 29, 30], that for a large class of weight ω, we have the *global asymptotics*:

$$K_N^\omega(x,y) = \frac{\sin\left[\pi N\left(F_\omega(x) - F_\omega(y)\right)\right]}{\pi(x - y)} + \underset{N \to \infty}{O(1)}, \qquad \text{(K)}$$

where the error is uniform for all $x, y \in I$, for any closed interval $I \subset \mathrm{supp}(\varrho_\omega)$, and $F_\omega = F_{\varrho_\omega}$ according to formula (1.13). Note that one shall interpret the RHS of (K) according to (3.8),

$$F_\omega(x) - F_\omega(y) \simeq \frac{\#\ \text{eigenvalues in } [x,y]}{N} \qquad \text{for any } x < y,$$

which is why the function F_ω is usually called the *integrated density of states*.

This section is organized as follows. First, in section 3.1, we provide an elementary proof of formula (K) for the GUE kernel (with weight $\omega(x) = e^{-2Nx^2}$) based on the classical Plancherel–Rotach asymptotics, [41]. In section 3.2, we give the proof of theorem 1.1 using the Riemann-Hilbert formulation for the OPs and the results of Deift et al., [13]. Then, in section 3.3, we provide estimates for the variance of linear statistics valid for the so-called one-cut ensembles. As mentioned in the introduction, the main goal is to use these estimates to extend the scope of the mesoscopic CLT to a wider class of test functions, see theorem 3.5. Finally, in section 3.4, we briefly discuss the case of non-varying weights ω, focusing on the so-called modified Jacobi ensembles. Although the technique of section 3.2 applies as well in this setting, we provide another elementary proof of (K) which is inspired from the case of the Chebyshev's polynomials.

3.1 The Gaussian unitary ensemble

The GUE (3.1) was introduced by E. Wigner as a model to describe scattering resonances of Heavy nuclei. In addition to being unitary invariant, the entries of a GUE matrix are independent Gaussian random variables, which makes the GUE a central model in random matrix theory. So, we dedicate this section to provide an elementary proof of the mesoscopic CLT. By theorem 1.2, it suffices to obtain the following asymptotics for the GUE correlation kernel:

Theorem 3.2. *Let* $\varrho_{\mathrm{sc}}(x) = \frac{2\sqrt{1-x^2}}{\pi} \mathbb{1}_{[-1,1]}(x)$ *and* $F_{\mathrm{sc}}(x) = \int_0^x \varrho_{\mathrm{sc}}(u) du$. *For any* $\epsilon > 0$, *we have for all* $|x|, |y| \le 1 - \epsilon$,

$$K_N^{\mathrm{GUE}}(x,y) = \frac{\sin\left[\pi N\left(F_{\mathrm{sc}}(x) - F_{\mathrm{sc}}(y)\right)\right]}{\pi(x-y)} + \underset{N \to \infty}{O_\epsilon}(1). \qquad (3.9)$$

In particular, theorem 3.2 implies that $K_N^{\mathrm{GUE}}(x,x) = N\varrho_{\mathrm{sc}}(x) + O(1)$ for all $|x| < 1$, and we recover that, with our normalization, the GUE equilibrium measure is the Wigner

semicircle law ϱ_{sc}. We shall show that formula (3.9) follows from the basic properties of the Hermite functions and the classical Plancherel–Rotach asymptotics. First of all, let us observe that the GUE weight satisfies $\omega(x) = \omega_G(\sqrt{2N}x)$ where $\omega_G(x) = e^{-x^2}$ and the OPs with respect to the Gaussian weight ω_G are the classical Hermite polynomials, for all $k \geq 0$,

$$\pi_k(x) = e^{x^2}\left(-\frac{1}{2}\frac{d}{dx}\right)^k e^{-x^2} \quad \text{and} \quad \gamma_k = \sqrt{\frac{2^k}{\sqrt{\pi}k!}}. \tag{3.10}$$

Therefore, if we let $\phi_k(x) = \sqrt{\omega_G(x)}\gamma_k\pi_k(x)$, according to formula (3.7), the correlation kernel of the GUE eigenvalue process satisfies $K_N^{\mathrm{GUE}}(x,y) = \sqrt{2N}K_N^{\omega_G}(\sqrt{2N}x,\sqrt{2N}y)$ where

$$K_N^{\omega_G}(x,y) = \sqrt{\frac{N}{2}}\frac{\phi_N(x)\phi_{N-1}(y) - \phi_{N-1}(x)\phi_N(y)}{x-y}.$$

That is, if we let $\widetilde{\phi}_k(x) = \left(\frac{N}{2}\right)^{1/4}\phi_k(\sqrt{2N}x)$, then the GUE kernel is given by

$$K_N^{\mathrm{GUE}}(x,y) = \frac{\widetilde{\phi}_N(x)\widetilde{\phi}_{N-1}(y) - \widetilde{\phi}_{N-1}(x)\widetilde{\phi}_N(y)}{x-y}. \tag{3.11}$$

The functions ϕ_k are usually called the Hermite (wave) functions, they form an orthonormal basis of $L^2(\mathbb{R})$, and they have well-known asymptotic expansions:

Proposition 3.3 ([40], Proposition 5.1.3). *Let for all $|x| < 1$ and $N > 0$,*

$$H(x) = \theta(x) - x\sqrt{1-x^2} \quad \text{and} \quad \Psi_N(x) = NH(x) - \frac{\arcsin(x)}{2}, \tag{3.12}$$

where $\theta(x) = \arccos(x)$. For any $\epsilon > 0$ and $\kappa \geq 0$, we have for all $|x| \leq 1 - \epsilon$,

$$\widetilde{\phi}_{N-\kappa}(x) = \frac{\cos\left[\Psi_N(x) - \kappa\theta(x)\right]}{\sqrt{\pi}(1-x^2)^{1/4}} + \underset{N\to\infty}{O_{\epsilon,\kappa}}\left(N^{-1}\right). \tag{3.13}$$

We note that the asymptotics (3.13) were first obtained in [41] using an integral formula for the Hermite polynomials (3.10) and the steepest descent method. Interestingly, Plancherel and Roctach not only obtained the leading order, but the full asymptotic expansion; see [41, formula 7]. In order to obtain the uniform asymptotics of theorem 3.2, we shall use the representation of [1, section 3.5.2] for the kernel (3.11):

$$K_N^{\mathrm{GUE}}(x,y) = \int_0^1 \left(\widetilde{\phi}_N(x)\widetilde{\phi}'_{N-1}(x+\xi t) - \widetilde{\phi}_{N-1}(x)\widetilde{\phi}'_N(x+\xi t)\right)dt, \tag{3.14}$$

where $\xi = y - x$, and the following lemma.

Lemma 3.4. *For any $\epsilon > 0$ and $\kappa \geq 0$, we have for all $|x| \leq 1 - \epsilon$,*

$$\widetilde{\phi}'_{N-\kappa}(y) = -\sqrt{\frac{2}{\pi}}\frac{\Psi'_N(y)\sin\left[\Psi_N(y) - \kappa\theta(x)\right]}{(1-y^2)^{1/4}} + \underset{N\to\infty}{O_{\epsilon,\kappa}}\left(1\right). \tag{3.15}$$

Proof. It is well-known that the Hermite polynomials are an Appell sequence, so that for all $k \geq 0$,

$$\phi'_k(x) = \frac{k\gamma_k}{\gamma_{k-1}}\phi_{k-1}(x) - x\phi_k(x) = \sqrt{2k}\phi_{k-1}(x) - x\phi_k(x)$$

$$\widetilde{\phi}'_k(x) = 2\sqrt{kN}\widetilde{\phi}_{k-1}(x) - 2Nx\widetilde{\phi}_k(x).$$

Thus, $\widetilde{\phi}'_{N-\kappa}(x) = 2N\left(\{1 + O(\kappa/N)\}\widetilde{\phi}_{N-\kappa-1}(x) - x\widetilde{\phi}_{N-\kappa}(x)\right)$ and the asymptotics of proposition 3.3 imply that

$$\widetilde{\phi}'_{N-\kappa}(x) = \frac{2N}{\sqrt{\pi}(1-x^2)^{1/4}}\left\{\cos\left[\Psi_N(x) - (\kappa+1)\theta(x)\right] - x\cos\left[\Psi_N(x) - \kappa\theta(x)\right]\right\} + \underset{N\to\infty}{O_{\epsilon,\kappa}}\left(1\right).$$

Then, applying the trigonometric identity $\cos[\alpha - \theta(x)] - x\cos[\alpha] = \sqrt{1-x^2}\sin[\alpha]$ which is valid for all $\alpha \in \mathbb{R}$ and $x \in [-1,1]$, we obtain for all $|y| \leq 1 - \epsilon$,

$$\widetilde{\phi}_{N-\kappa}(y) = \frac{2N\sqrt{1-y^2}}{\sqrt{\pi}(1-y^2)^{1/4}}\sin\left[\Psi_N(y) - \kappa\theta(y)\right] + \underset{N\to\infty}{O_{\epsilon,\kappa}}(1).$$

To complete the proof, it remains to check that the function $H \in C^1(-1,1)$ with

$$H'(x) = -2\sqrt{1-x^2}, \tag{3.16}$$

and note that for all $|x| \leq 1 - \epsilon$,

$$\Psi'_N(x) = NH'(x) - \frac{1}{2\sqrt{1-x^2}} = -2N\sqrt{1-x^2} + \underset{N\to\infty}{O_\epsilon}(1). \tag{3.17}$$

$\square$

Proof of theorem 3.2. Fix $0 < \epsilon < 1$. Using the asymptotics (3.13), (3.15), and some elementary trigonometry, we obtain for all $|x|, |z| \leq 1 - \epsilon$,

$$\widetilde{\phi}_N(x)\widetilde{\phi}'_{N-1}(z) - \widetilde{\phi}_{N-1}(x)\widetilde{\phi}'_N(z)$$

$$= \Psi'_N(z)\frac{\sin\left[\Psi_N(z)\right]\cos\left[\Psi_N(x) - \theta(x)\right] - \cos\left[\Psi_N(x)\right]\sin\left[\Psi_N(z) - \theta(z)\right]}{\pi(1-x^2)^{1/4}(1-z^2)^{1/4}} + \underset{N\to\infty}{O_\epsilon}(1)$$

$$= \Psi'_N(z)\frac{\sin\left[\Psi_N(z)\right]\sin\left[\Psi_N(x)\right]\sqrt{1-x^2} + \cos\left[\Psi_N(x)\right]\cos\left[\Psi_N(z)\right]\sqrt{1-z^2}}{\pi(1-x^2)^{1/4}(1-z^2)^{1/4}}$$

$$\qquad\qquad + \Psi'_N(z)\frac{\sin\left[\Psi_N(z)\right]\cos\left[\Psi_N(x)\right](x-z)}{\pi(1-x^2)^{1/4}(1-z^2)^{1/4}} + \underset{N\to\infty}{O_\epsilon}(1)$$

$$= \frac{\Psi'_N(z)\cos[\Psi_N(z) - \Psi_N(x)]}{\pi} + \mho(x,z)\Psi'_N(z)\sin\left[\Psi_N(z)\right]\sin\left[\Psi_N(x)\right] \tag{3.18}$$

$$+ \mho(z,x)\Psi'_N(z)\cos\left[\Psi_N(z)\right]\cos\left[\Psi_N(x)\right] + \Im(x,z)\Psi'_N(z)\sin\left[\Psi_N(z)\right]\cos\left[\Psi_N(x)\right] + \underset{N\to\infty}{O_\epsilon}(1),$$

where

$$\mho(x,z) = \frac{1}{\pi}\left(\left(\frac{1-x^2}{1-z^2}\right)^{1/4} - 1\right) \quad\text{and}\quad \Im(x,z) = \frac{x-z}{\pi(1-x^2)^{1/4}(1-z^2)^{1/4}}.$$

In particular, since $\mho(x,x) = \Im(x,x) = 0$, this implies that

$$\widetilde{\phi}_N(x)\widetilde{\phi}'_{N-1}(x) - \widetilde{\phi}_{N-1}(x)\widetilde{\phi}'_N(x) = \frac{\Psi'_N(x)}{\pi} + \underset{N\to\infty}{O_\epsilon}(1).$$

Thus, by formulae (3.14) with $\xi = 0$ and (3.17), this shows that

$$K_N^{\mathrm{GUE}}(x,x) = \widetilde{\phi}_N(x)\widetilde{\phi}'_{N-1}(x) - \widetilde{\phi}_{N-1}(x)\widetilde{\phi}'_N(x) = N\varrho_{\mathrm{sc}}(x) + \underset{N\to\infty}{O_\epsilon}(1), \tag{3.19}$$

which is consistent with (3.9). For now, we suppose that $\xi \neq 0$. Then, we claim that

$$\int_0^1 \Psi'_N(x + \xi t)\cos[\Psi_N(x + \xi t) - \Psi_N(x)]dt = -\frac{\sin[\Psi_N(x+\xi) - \Psi_N(x)]}{\xi},$$

and that for any function $\mathfrak{Q} \in C^1(J)$ such that $\mathfrak{Q}(x) = 0$ and $\sup_{|z| \leq 1-\epsilon}\left|\mathfrak{Q}'(z)\right| \leq C_\epsilon$, we have

$$\left|\int_0^1 \mathfrak{Q}(x + \xi t)\Psi'_N(x + \xi t)\sin\left[\Psi_N(x + \xi t)\right]dt\right| \leq 2C_\epsilon. \tag{3.20}$$

So, if we apply this estimate to the functions $z \mapsto \mho(x, z)$ and $z \mapsto \mathfrak{Z}(x, z)$ (the same holds as well for $z \mapsto \mho(z, x)$ if we replace $\sin$ by $\cos$ in (3.20)), taking $z = x + \xi t$ in formula (3.18), we obtain for all $|x|, |x + \xi| \leq 1 - \epsilon$,

$$
\begin{aligned}
K_N^{\mathrm{GUE}}(x, y) &= \int_0^1 \left(\widetilde{\phi}_N(x) \widetilde{\phi}'_{N-1}(x + \xi t) - \widetilde{\phi}_{N-1}(x) \widetilde{\phi}'_N(x + \xi t) \right) dt \\
&= \frac{\sin[\Psi_N(x) - \Psi_N(x + \xi)]}{\pi \xi} + O_\epsilon(1).
\end{aligned}
$$

To conclude, it remains to see that by (3.12) and (3.16), if $\xi = y - x$, then

$$
\begin{aligned}
\Psi_N(x) - \Psi_N(x + \xi) &= N\big(H(x) - H(y)\big) + O_\epsilon(\xi) \\
&= -N\pi \int_y^x \varrho_{\mathrm{sc}}(du) + O_\epsilon(\xi).
\end{aligned}
$$

Finally, to complete the proof, note that we may prove (3.20) by integration by parts. By assumption $|\mathfrak{Q}(x + \xi)| \leq C_\epsilon |\xi|$ so that

$$
\begin{aligned}
\int_0^1 \mathfrak{Q}(x + \xi t) \Psi'_N(x + \xi t) \sin \big[\Psi_N(x + \xi t)\big] dt = &-\frac{\mathfrak{Q}(x + \xi)}{\xi} \cos \big[\Psi_N(x + \xi t)\big] \\
&+ \int_0^1 \mathfrak{Q}(x + \xi t) \cos \big[\Psi_N(x + \xi t)\big] dt,
\end{aligned}
$$

and obviously both terms are bounded by the constant C_ϵ. $\qquad \square$

3.2 Proof of theorem 1.1

In this section, we focus on a varying weight of the type $\omega(x) = e^{-NV(x)}$ where the potential V is real-analytic on $\mathbb{R}$ and satisfies the condition (1.2). Then, the equilibrium density ϱ_V is smooth on the set $J_V = J_{\varrho_V}$ which is composed of finitely many bounded intervals; see [13, 12] for further references. As we already mentioned, by theorem 1.2, it suffices to know the global asymptotics (K) of the correlation kernel which follow from the results of the steepest descent for the OP Riemann-Hilbert problem developed in [13]. The formulae referenced by $\{\#\}$ belwo are taken from this paper. The authors of [13] were interested in universality of the local correlations and convergence of the gap probability for the eigenvalue process, so that the asymptotics (K) are not stated explicitly in their paper and we may only refer to [13, Lemma 6.1] for an analogous result valid in the local regime. So, for completeness, we review below the main steps of the proof of (K).

We fix a component of J_V, denoted (b, a), and according to formula $\{6.7\}$, we let

$$
\phi(x) = \int_x^a \varrho_V(s) ds. \tag{3.21}
$$

Note that in [13], the equilibrium density is denoted by Ψ, $\{1.6\}$, instead of ϱ_V. By $\{2.2\}$, we can write the correlation kernel (3.7) as

$$
K_N^\omega(x, y) = -e^{-N(V(x) + V(y))/2} \frac{Y_{11}(x) Y_{21}(y) - Y_{11}(y) Y_{21}(x)}{2\pi i (x - y)}, \tag{3.22}
$$

where the 2×2 matrix Y is the solution of an appropriate Riemann-Hilbert problem. Transforming the problem, cf. $\{6.8 - 6.9\}$, the authors proved that for any $x \in (a, b)$,

$$
\begin{cases}
Y_{11}(x) = M_{11}(x) \exp \big[N(V(x) + \ell + 2\pi i \phi(x))/2\big] + M_{12}(x) \exp \big[N(V(x) + \ell - 2\pi i \phi(x))/2\big] \\
Y_{21}(x) = M_{21}(x) \exp \big[N(V(x) - \ell + 2\pi i \phi(x))/2\big] + M_{22}(x) \exp \big[N(V(x) - \ell - 2\pi i \phi(x))/2\big]
\end{cases}
\tag{3.23}
$$

where the 2×2 matrices $M(z)$ and $\frac{d}{dz}M(z)$, which depends on the dimension N, are uniformly bounded for all z in a complex neighborhood of any point $x_0 \in J_V$ and for all $N > N_0$, cf. $\{5.161\}$. Using formulae (3.23), a little of algebra shows that for all $x, y \in I$,

$$e^{-N(V(x)+V(y))/2}\big(Y_{11}(x)Y_{21}(y) - Y_{11}(y)Y_{21}(x)\big)$$

$$= e^{i\pi N(\phi(x)-\phi(y))}\left\{\det M(x) - M_{11}(x)\big(M_{22}(x) - M_{22}(y)\big) + M_{21}(x)\big(M_{12}(x) - M_{12}(y)\big)\right\}$$

$$+ e^{-i\pi N(\phi(x)-\phi(y))}\left\{-\det M(x) + M_{22}(x)\big(M_{11}(x) - M_{11}(y)\big) - M_{12}(x)\big(M_{21}(x) - M_{21}(y)\big)\right\}$$

$$+ e^{i\pi N(\phi(x)+\phi(y))}\left\{M_{21}(y)\big(M_{11}(x) - M_{11}(y)\big) - M_{11}(y)\big(M_{21}(x) - M_{21}(y)\big)\right\}$$

$$+ e^{-i\pi N(\phi(x)+\phi(y))}\left\{M_{22}(y)\big(M_{12}(x) - M_{12}(y)\big) - M_{12}(y)\big(M_{22}(x) - M_{22}(y)\big)\right\}$$

$$= 2i\det M(x)\sin\big(\pi N(\phi(x) - \phi(y))\big) + O(x - y).$$

Hence, since $\det M(z) = 1$ for all $z \in \mathbb{C}$, by formula (3.22), we obtain

$$K_N^\omega(x, y) = \frac{\sin \pi N\big(\phi(y) - \phi(x)\big)}{\pi(x - y)} + O(1). \tag{3.24}$$

Moreover, formula (3.24) holds uniformly for any points $x, y \in I$ where $I \subset (a, b)$ is any closed interval. To conclude it remains to observe that by (3.21), one has $\phi(y) - \phi(x) = F_\omega(x) - F_\omega(y)$.

3.3 Extension of the mesoscopic CLT in the one-cut case

The goal of this section is to provide estimates on the variance of linear statistics valid for determinantal process whose correlation kernel are of the type (3.7) and the OPs satisfy certain *semiclassical asymptotics*. Our main motivation is to upgrade the CLT of theorem 1.1 to a larger class of test functions. In particular, we obtain the following result:

Proposition 3.5. *Let $V : \mathbb{R} \to \mathbb{R}$ be a real-analytic function which satisfies the conditions* (1.2) *and* (1.6) *with $c_0 = 0$ and $\ell = 1$. Under $\mathbb{P}_N^V$, the CLT* (1.3) *holds for any $x_0 \in J_V$, any $0 < \alpha < 1$, and for all test function $f \in H^{1/2}(\mathbb{R})$ such that there exists $\mathrm{L} > 0$ and*

$$\limsup_{|x|\to\infty} \sup\left\{|x|\left|\frac{f(x) - f(y)}{x - y}\right| : |y| \le |x|\right\} < \mathrm{L}. \tag{H.1}$$

Remark 3.6. It is straightforward to check that $(\mathrm{H.1})$ holds in both cases:
  i) $f \in C^1(\mathbb{R})$ and $|f'(x)| \le \mathrm{L}/|x|$.
  ii) f is bounded and has compact support.

We will focus on the one-cut regime because the asymptotics of the OPs are simpler, but we expect that similar variance estimates holds in the multi-cut as well even though the analysis would be more involved. According to formula (3.5), we let $\Phi_k(x) = \gamma_k \pi_k(x)\sqrt{\omega(x)}$. Thus $(\Phi_k)_{k=0}^\infty$ is an orthonormal family in $L^2(\mathbb{R})$ and the Christoffel-Darboux kernel for the weight $\omega(x)$ on $\mathbb{R}$ is given by

$$K_N^\omega(x, y) = \frac{\gamma_{N-1}}{\gamma_N}\frac{\Phi_N(x)\Phi_{N-1}(y) - \Phi_{N-1}(x)\Phi_N(y)}{x - y}. \tag{3.25}$$

For now on, we suppose that $\mathrm{supp}(\varrho_\omega) = (-1, 1)$ and that the OPs have the following asymptotics for all $|x| < 1$,

$$\Phi_{N-\kappa}(x) = \sqrt{\frac{2}{\pi}}\frac{\cos\big[N\pi F_\omega(x) + \psi_\omega(x) + \kappa\theta(x)\big]}{(1 - x^2)^{1/4}} + \underset{N\to\infty}{o_\epsilon(1)}, \tag{3.26}$$

where $\kappa = 0, 1$, $\theta(x) = \arccos(x)$, the function $\psi_\omega \in C(-1, 1)$. Moreover, we also suppose that

$$\lim_{N\to\infty}\frac{\gamma_{N-1}}{\gamma_N} = \frac{1}{2}. \tag{3.27}$$

Remark 3.7. Formulae (3.26–3.27) are sometimes referred as *semiclassical* or Plancherel-Rotach asymptotics, [38]. This terminology comes by analogy with the GUE case, see proposition 3.3. Generally, if $\omega(x) = e^{-NV(x)}$, $V(x)$ is real-analytic, and supp $\varrho_V = [-1, 1]$, these asymptotics follow from the steepest descent for the OP Riemann-Hilbert problem, see [13, Thm 1.1 and formula (1.64)] and $\psi_\omega(x) = \arcsin(x)/2$. Finally, the case of non-varying weights is discussed by theorem 3.11 in section 3.4 below.

We consider the determinantal process Ξ_N with correlation kernel (3.25) and let $\Xi_N f := \sum f(\lambda_k)$ where the sum is over the configuration $\{\lambda_k\}_{k=1}^N$. To prove theorem 3.5, our goal is to first show that there exists a constant $C > 0$ (independent of N and f) so that

$$\text{Var}\left[\Xi_N f\right] \leq C\widetilde{\Sigma}(f)^2 \quad \text{where} \quad \widetilde{\Sigma}(f)^2 = \frac{1}{\pi^2} \iint_{[-1,1]^2} \left|\frac{f(x) - f(y)}{x - y}\right|^2 \frac{dxdy}{\sqrt{1-x^2}\sqrt{1-y^2}}, \quad (3.28)$$

for any function $f \in H^{1/2}(\mathbb{R})$ such that there exists $\delta > 0$ and $L > 0$ so that

$$\sup\left\{\left|\frac{f(x) - f(y)}{x - y}\right| : |x| \vee |y| > 1 - \delta\right\} < L. \quad \text{(H.2)}$$

To obtain the estimate (3.28), we use that the kernel K_N^ω defines a projection on $L^2(\mathbb{R})$, cf. e.g. [25, Lemma 3.1], so that

$$\text{Var}\left[\Xi_N f\right] = \frac{1}{2} \iint |f(x) - f(y)|^2 \left|K_N^\omega(x,y)\right|^2 dxdy, \quad (3.29)$$

and the following lemma:

Lemma 3.8. *For any $0 < \epsilon < 1$, if $J_\epsilon = [-1 + \epsilon, 1 - \epsilon]$, we have for $\kappa = 0, 1$,*

$$\int_{J_\epsilon} \left|\Phi_{N-\kappa}(x)\right|^2 dx = \frac{2\arcsin(1 - \epsilon)}{\pi} + \underset{N\to\infty}{o_\epsilon(1)}.$$

Proof. By (3.26), for any $|x| < 1$,

$$\left|\Phi_N(x)\right|^2 = \frac{2}{\pi\sqrt{1-x^2}}\left\{\frac{1 + \cos\left[2N\pi F_\omega(x) + 2\psi_\omega(x)\right]}{2} + o_\epsilon(1)\right\},$$

and this implies that

$$\int_{J_\epsilon} \left|\Phi_N(x)\right|^2 dx = \int_{J_\epsilon} \frac{1 + o_\epsilon(1)}{\pi\sqrt{1-x^2}} dx + \int_{J_\epsilon} \frac{\cos\left[2N\pi F_\omega(x) + 2\psi_\omega(x)\right]}{\pi\sqrt{1-x^2}} dx.$$

The first integral converges to $2\arcsin(1-\epsilon)/\pi$ as $N \to \infty$ and it remains to show that the second integral which is oscillatory converges to 0. By assumption, $F_\omega'(x) = \varrho_\omega(x) > 0$ for all $x \in (-1, 1)$ and we can make the change of variable $x = G(y)$ where $G = F_\omega^{-1}$, (2.10). So that, if $g \in L^1(\mathbb{R})$ with compact support in $(-1, 1)$, we have

$$\int g(x)e^{i2N\pi F_\omega(x)}dx = \int g(G(y))G'(y)e^{i2N\pi y}dy,$$

and, since $y \mapsto g(G(y))G'(y)$ is integrable, $\lim_{N\to\infty} \int g(x)e^{i2N\pi F_\omega(x)}dx = 0$ by the Riemann-lebesgue lemma. Applying this argument to the function $g(x) = \frac{e^{i2\psi_\omega(x)}}{\pi\sqrt{1-x^2}}\mathbb{1}_{J_\epsilon}(x)$, we conclude that for any $0 < \epsilon < 1$,

$$\lim_{N\to\infty} \int_{J_\epsilon} \frac{\cos\left[2N\pi F_\omega(x) + 2\psi_\omega(x)\right]}{\pi\sqrt{1-x^2}} dx = 0$$

and this complete the proof (of course, we may use the same argument when $\kappa = 1$). $\square$

Proposition 3.9. *Suppose that the OPs with respect to the weight $\omega(x)$ on $\mathbb{R}$ satisfy (3.26–3.27), then for any function $f \in H^{1/2}(\mathbb{R})$ which satisfies the condition (H.2), we have for any $0 < \epsilon < \delta$,*

$$\operatorname{Var}\left[\Xi_N f\right] \leq \Delta_N(\epsilon)\widetilde{\Sigma}(f)^2 + \underset{N\to\infty}{O}\left(\mathrm{L}^2\Theta(\epsilon)\right), \tag{3.30}$$

where $\Theta(\epsilon) = 1 - \frac{2\arcsin(1-\epsilon)}{\pi}$, $\Delta_N(\epsilon) > 0$ and $\lim_{N\to\infty} \Delta_N(\epsilon) = 8$ for any $0 < \epsilon < 1$.

Proof. For any $0 < \epsilon < 1$, let $J_\epsilon = [-1+\epsilon, 1-\epsilon]$. By (3.25), we have

$$\left|K_N^\omega(x,y)\right|^2 \leq \frac{2\gamma_{N-1}}{\gamma_N}\frac{\left|\Phi_N(x)\Phi_{N-1}(y)\right|^2 + \left|\Phi_{N-1}(x)\Phi_N(y)\right|^2}{|x-y|^2}. \tag{3.31}$$

Using the asymptotics (3.26–3.27), we get for all $|x|, |y| < 1-\epsilon$,

$$\left|K_N^\omega(x,y)\right|^2 \leq \frac{1}{\pi^2}\frac{\Delta_N(\epsilon)}{\sqrt{1-x^2}\sqrt{1-y^2}}\frac{1}{|x-y|^2}, \tag{3.32}$$

where $\Delta_N(\epsilon) > 0$ and $\lim_{N\to\infty}\Delta_N(\epsilon) = 8$ for any $0 < \epsilon < 1$. By (3.28), this implies that

$$\iint_{J_\epsilon^2} |f(x) - f(y)|^2 \left|K_N^\omega(x,y)\right|^2 dxdy \leq \Delta_N(\epsilon)\widetilde{\Sigma}(f)^2. \tag{3.33}$$

On the other hand, if f satisfies the hypothesis (H.2) and $0 < \epsilon < \delta$, by formula (3.31),

$$\iint_{\mathbb{R}^2\setminus J_\epsilon^2} |f(x) - f(y)|^2 \left|K_N^\omega(x,y)\right|^2 dxdy \leq \frac{4\gamma_{N-1}}{\gamma_N}\mathrm{L}^2\iint_{\mathbb{R}^2\setminus J_\epsilon^2}\left|\Phi_N(x)\Phi_{N-1}(y)\right|^2 dxdy. \tag{3.34}$$

By symmetry

$$\iint_{\mathbb{R}^2\setminus J_\epsilon^2}\left|\Phi_N(x)\Phi_{N-1}(y)\right|^2 dxdy$$

$$\leq \int_{\mathbb{R}}\left|\Phi_{N-1}(y)\right|^2 dy\int_{\mathbb{R}\setminus J_\epsilon}\left|\Phi_N(x)\right|^2 dx + \int_{\mathbb{R}}\left|\Phi_N(x)\right|^2 dx\int_{\mathbb{R}\setminus J_\epsilon}\left|\Phi_{N-1}(y)\right|^2 dy,$$

and since $\|\Phi_N\|_{L^2} = \|\Phi_{N-1}\|_{L^2} = 1$, by lemma 3.8, we obtain

$$\limsup_{N\to\infty}\iint_{\mathbb{R}^2\setminus J_\epsilon^2}\left|\Phi_N(x)\Phi_{N-1}(y)\right|^2 dxdy \leq 2\left(1 - \frac{2\arcsin(1-\epsilon)}{\pi}\right).$$

Since $\frac{\gamma_{N-1}}{\gamma_N} \to \frac{1}{2}$, this upper-bound and (3.34) imply that for any $0 < \epsilon < \delta$,

$$\iint_{\mathbb{R}^2\setminus J_\epsilon^2} |f(x) - f(y)|^2 \left|K_N^\omega(x,y)\right|^2 dxdy = \underset{N\to\infty}{O}\left(\mathrm{L}^2\Theta(\epsilon)\right). \tag{3.35}$$

The claim follows from formula (3.29) by combining the estimates (3.33) and (3.35). $\square$

We are now ready to complete the proof of proposition 3.5. The method consists in using a density argument, combined with the variance estimate of proposition 3.9 and lemma A.5 which is borrowed from [48].

Proof of proposition 3.5. The assumption (H.1) implies that there exists $C > 0$ so that, if $|x| \geq C$, for all $|y| \leq |x|$,

$$\left| \frac{f(x) - f(y)}{x - y} \right| \leq \frac{L}{|x|}.$$

Fixing $x_0 \in J$, if we let $g_N(x) = f\big(N^\alpha(x - x_0)\big)$, this implies that for all $|x - x_0| > CN^{-\alpha}$ and for all $|y - x_0| < |x - x_0|$,

$$\left| \frac{g_N(x) - g_N(y)}{x - y} \right| \leq \frac{L}{|x - x_0|}.$$

This inequality shows that, if $|x_0 - 1| \wedge |x_0 + 1| = 2\delta$, then for all $N > (C/\delta)^{1/\alpha}$ and for all $|x| > 1 - \delta$,

$$\left| \frac{g_N(x) - g_N(y)}{x - y} \right| \leq \frac{L}{\delta}. \tag{3.36}$$

Hence, since the LHS of (3.36) is symmetric, this proves that for sufficiently large N, the functions g_N satisfy the condition (H.2) and by proposition 3.9, we have for any $0 < \epsilon < 1$,

$$\text{Var}\left[\Xi_N^{x_0,\alpha} f\right] = \text{Var}\left[\Xi_N g_N\right] \leq \Delta_N(\epsilon) \widetilde{\Sigma}(g_N)^2 + \underset{N \to \infty}{O}\big(L^2 \delta^{-2} \Theta(\epsilon)\big). \tag{3.37}$$

By (3.28), if we set $J_\epsilon = [-1 + \epsilon, 1 - \epsilon]$, then

$$\widetilde{\Sigma}(g_N)^2 = I_1(f; N, \epsilon) + I_2(f; N, \epsilon) \tag{3.38}$$

$$= \frac{1}{\pi^2} \iint_{J_\epsilon} \left| \frac{g_N(x) - g_N(y)}{x - y} \right|^2 \frac{dx\,dy}{\sqrt{1 - x^2}\sqrt{1 - y^2}}$$

$$+ \frac{1}{\pi^2} \iint_{[-1,1]^2 \setminus J_\epsilon} \left| \frac{g_N(x) - g_N(y)}{x - y} \right|^2 \frac{dx\,dy}{\sqrt{1 - x^2}\sqrt{1 - y^2}}.$$

By a change of variables,

$$I_1(f; N, \epsilon) = \frac{1}{\pi^2} \iint_{\mathcal{B}_N} \left| \frac{f(u) - f(v)}{u - v} \right|^2 \frac{du\,dv}{\sqrt{1 - (x_0 + N^{-\alpha}u)^2}\sqrt{1 - (x_0 + N^{-\alpha}v)^2}},$$

where $\mathcal{B}_N = \big[N^\alpha(-1 + \epsilon - x_0), N^\alpha(1 - \epsilon - x_0)\big]^2$. Since $f \in H^{1/2}(\mathbb{R})$, by the dominated convergence theorem, we obtain for any $0 < \epsilon < 1$,

$$\lim_{N \to \infty} I_1(f; N, \epsilon) = \frac{1}{\pi^2} \iint_{\mathbb{R}^2} \left| \frac{f(u) - f(v)}{u - v} \right|^2 du\,dv = 4\|f\|_{H^{1/2}}^2.$$

On the other hand, using the estimate (3.36), we have for all $0 < \epsilon < \delta$,

$$I_2(f; N, \epsilon) \leq \frac{L^2}{\pi^2 \delta^2} \iint_{[-1,1]^2 \setminus J_\epsilon} \frac{dx\,dy}{\sqrt{1 - x^2}\sqrt{1 - y^2}} = \frac{L^2}{4\delta^2} \Theta(\epsilon)^2.$$

Thus, according to formula (3.38), we obtain

$$\widetilde{\Sigma}(g_N)^2 = 4\|f\|_{H^{1/2}}^2 + \underset{N \to \infty}{O}\big(\Theta(\epsilon)^2\big).$$

and combined with (3.37), it shows that there exists a constant $C > 0$ so that for any $0 < \epsilon < \delta$,

$$\limsup_{N \to \infty} \text{Var}\left[\Xi_N g_N\right] \leq 32\|f\|_{H^{1/2}}^2 + C\Theta(\epsilon).$$

Since this holds for any $0 < \epsilon < \delta$ and $\lim_{\epsilon \searrow 0} \Theta(\epsilon) = 0$, this implies that

$$\limsup_{N \to \infty} \text{Var}\left[\Xi_N^{x_0,\alpha} f\right] \leq 32\|f\|_{H^{1/2}}^2. \tag{3.39}$$

Using this estimate, we can apply lemma A.5 with $\mathfrak{F} = \big\{ f \in H^{1/2}(\mathbb{R}) : f \text{ satisfies (H.1)} \big\}$ and $\mathfrak{X} = C_0^\infty(\mathbb{R})$ in order to complete the proof. $\qquad\square$

3.4 Modified Jacobi ensembles

In this section, we apply theorem 1.2 to orthogonal polynomial ensembles with respect to a non-varying measure with support on $[-1, 1]$, hence providing an alternative proof of the main result of [10]. In particular, we obtain an elementary proof of the mesoscopic CLT for the classical Jacobi ensembles. We consider the following family of weights:

$$\omega(x) = \begin{cases} h(x)(1-x)^{\gamma_-}(1+x)^{\gamma_+} & \text{if } |x| \leq 1 \\ 0 & \text{else} \end{cases}, \tag{3.40}$$

where $\gamma_+, \gamma_- > -1$ and $h(x)$ is a function which is real-analytic and strictly positive on the interval $(-1-\epsilon, 1+\epsilon)$ for some $\epsilon > 0$ and normalized so that $\omega(x)dx$ is a probability measure. Note that these correspond to unitary invariant measures, which are called **the modified Jacobi unitary ensembles**, of the form:

$$\mathbb{P}_N^\omega = Z_{\omega,N}^{-1} \det\left[\omega(\mathrm{H})\right] \mathbb{1}_{\|\mathrm{H}\| \leq 1} d\mathrm{H}, \tag{3.41}$$

where $\|\mathrm{H}\|$ is the operator norm of H. Therefore, $\mathbb{P}_N^\omega$ induces a determinantal process on $J = (-1, 1)$ with correlation kernel (3.7). Moreover, if h is constant, the OPs with respect to ω are the classical Jacobi polynomials and their asymptotics are well-known, cf. [49, Thm 8.21.8 and Thm 12.1.4]. In general, the asymptotics of the OPs with respect to the weight (3.40) have been analyzed in [29, 30] using the Deift–Zhou steepest descent method. We note that the formulation is analogous to [13] but the set-up is more elementary since the weight (3.40) does not vary with the dimension N. Thus, going through the proof of section 3.2, we obtain the following asymptotics for the correlation kernels of the modified Jacobi ensembles:

Proposition 3.10. *Let $\varrho(x) = \frac{1}{\pi\sqrt{1-x^2}}\mathbb{1}_{|x|\leq 1}$ be the **arcsine measure** on J and*

$$F_\varrho(x) = \frac{\arcsin(x)}{\pi} \qquad \text{if } |x| \leq 1. \tag{3.42}$$

For any weight ω of the type (3.40) and for any $\epsilon > 0$, we have for all $|x|, |y| \leq 1 - \epsilon$,

$$K_N^\omega(x,y) = \frac{\sin\left[\pi N\big(F_\varrho(x) - F_\varrho(y)\big)\right]}{\pi(x-y)} + \underset{N\to\infty}{O_\epsilon}(1). \tag{3.43}$$

Proposition 3.10 implies that the arcsine law ϱ is the equilibrium density for the eigenvalue process of the modified Jacobi ensembles. In contrast to the varying weights $e^{-NV(x)}$ analyzed in section 3.2, the global eigenvalues distribution is independent of the parameters of the model and is always one-cut. Moreover, it was proved in [29] that the OPs satisfy formulae (3.26–3.27). This follows readily from the next theorem and the fact that with our notation $\theta(x) = \arccos(x)$ and $-\theta(x) = \pi\big(F_\varrho(x) - 1/2\big)$.

Theorem 3.11 (Thm. 1.6, Thm. 1.12, [29]). *For any weight ω of type (3.40), there exists $D_\infty > 0$ and $\vartheta \in C^1(-1, 1)$ such that the OPs with respect to $\omega(x)dx$ satisfy*

$$\pi_N(x) = \frac{D_\infty}{2^N\sqrt{\pi\omega(x)\sqrt{1-x^2}}} \cos\left[(N+1/2)\theta(x) + \vartheta(x) - \pi/4\right] + \underset{N\to\infty}{O}(N^{-1}),$$

uniformly for all $|x| \leq 1 - \epsilon$ for any $\epsilon > 0$ and $\gamma_N = 2^N \frac{\sqrt{2}}{D_\infty}\left\{1 + \underset{n\to\infty}{O}(N^{-1})\right\}$.

Hence, by theorem 1.2, combining the asymptotics of proposition 3.10 with the variance estimates of section (3.3), we obtain the following result:

Theorem 3.12. *If $\{\lambda_1, \ldots, \lambda_N\}$ denote the eigenvalues of a random matrix distributed according to $\mathbb{P}_N^\omega$, (3.41), then for any $x_0 \in (-1, 1)$, any $0 < \alpha < 1$, and for all $f \in H^{1/2} \cap L^\infty(\mathbb{R})$ with compact support, one has*

$$\sum_{k=1}^N f\big(N^\alpha(\lambda_k - x_0)\big) - \mathbb{E}_N^\omega \left[\sum_{k=1}^N f\big(N^\alpha(\lambda_k - x_0)\big)\right] \underset{N \to \infty}{\Longrightarrow} \mathcal{N}\big(0, \|f\|_{H^{1/2}}^2\big).$$

As mentioned in the introduction, this is an extension of [10, Thm 1.1] for bounded test functions which are in the Sobolev class $H_0^{1/2}(\mathbb{R})$. In the remainder of this section, we will provide an elementary proof of proposition 3.10 for **the Chebyshev ensemble** and explain how to general the argument to all weights of the type (3.40). In particular, for the classical Jacobi ensembles, this provides a proof of theorem 3.12 which does not relie on the Riemann–Hilbert machinery. The main observation is that when $\gamma_+ = \gamma_- = -1/2$ and $h = 1/\pi$, i.e. $\omega = \varrho$, the OPs which appear in the correlation kernel (3.7) are the Chebyshev polynomials of the first kind, $\pi_k = 2^{-k}T_k$, and they satisfy for all $k \geq 0$ and $x \in [-1, 1]$,

$$\pi_k(x) = 2^{-k} \cos\big[(k+1)\theta(x)\big] \qquad \text{and} \qquad \gamma_k = 2^k\sqrt{2}.$$

Thus, the correlation kernel of the Chebychev process is given explicitly by

$$K_N^\varrho(x, y) = \frac{\cos[(N+1)\theta(x)]\cos[N\theta(y)] - \cos[(N+1)\theta(y)]\cos[N\theta(x)]}{\pi(1-x^2)^{1/4}(1-y^2)^{1/4}(x-y)}. \tag{3.44}$$

In this case, the proof of proposition 3.10 just relies on elementary trigonometric identities which are summarized by the following lemma.

Lemma 3.13. *For any function $\Psi_N \in C(J)$, we define the kernel*

$$K_{\Psi_N}(x, y) = \frac{\cos\big[\Psi_N(x)\big]\cos\big[\Psi_N(y) - \theta(y)\big] - \cos\big[\Psi_N(y)\big]\cos\big[\Psi_N(x) - \theta(x)\big]}{\pi(1-x^2)^{1/4}(1-y^2)^{1/4}(x-y)}. \tag{3.45}$$

For any $\epsilon > 0$, one has for all $|x|, |y| < 1 - \epsilon$,

$$K_{\Psi_N}(x, y) = \frac{\sin\big[\Psi_N(y) - \Psi_N(x)\big]}{\pi(x-y)} + \underset{|x-y| \to 0}{O_\epsilon(1)}. \tag{3.46}$$

Proof. Using that $\cos\big[\Psi_N(x) - \theta(x)\big] = x\cos\big[\Psi_N(x)\big] + \sqrt{1-x^2}\sin\big[\Psi_N(x)\big]$, we deduce that for all $|x|, |y| < 1$,

$$\begin{aligned}
K_{\Psi_N}(x, y) = &- \cos\big[\Psi_N(x)\big]\cos\big[\Psi_N(y)\big] \\
&+ \frac{\sqrt{1-y^2}\cos\big[\Psi_N(x)\big]\sin\big[\Psi_N(y)\big] - \sqrt{1-x^2}\cos\big[\Psi_N(y)\big]\sin\big[\Psi_N(x)\big]}{\pi(1-x^2)^{1/4}(1-y^2)^{1/4}(x-y)}.
\end{aligned}$$

Then, the estimate $\left|\left(\frac{1-x^2}{1-y^2}\right)^{1/4} - 1\right| \leq \frac{|x-y|}{1-y^2}$ implies that for all $|x|, |y| < 1 - \epsilon$,

$$K_{\Psi_N}(x, y) = \frac{\cos\big[\Psi_N(x)\big]\sin\big[\Psi_N(y)\big] - \cos\big[\Psi_N(y)\big]\sin\big[\Psi_N(x)\big]}{\pi(x-y)} + \underset{|x-y| \to 0}{O_\epsilon(1)},$$

and formula (3.46) follows directly from another trigonometric identity. $\square$

In particular, for the Chebyshev process, by formula (3.44), we have $K_N^\varrho = K_{\Psi_N}$ with the phase $\Psi_N(x) = (N+1)\theta(x)$ and the asymptotics (3.43) follow directly from lemma 3.13 and the fact that for all $x, y \in [-1, 1]$,

$$\theta(y) - \theta(x) = \pi\big(F_\varrho(x) - F_\varrho(y)\big). \tag{3.47}$$

Remark 3.14. By theorem 1.2, this argument provides an elementary proof of the mesoscopic CLT for the Chebyshev eigenvalue process. Then, we note that universality for the modified Jacobi ensembles can be deduced directly from [10, Thm 1.2] and the asymptotics of the recurrence coefficients which are given by [29, Thm 1.10] for any weight of the form (3.40).

Proof of proposition 3.10. For a general modified Jacobi ensemble, according to theorem 3.11, we may sill apply lemma 3.13 with $\Psi_N(x) = (N + 1/2)\theta(x) + \vartheta(x) - \pi/4$. Then, by formula (3.47), one has

$$\Psi_N(y) - \Psi_N(x) = N\pi\big(F_\varrho(x) - F_\varrho(y)\big) + \underset{|x-y|\to 0}{O_\epsilon}\big(|x-y|\big),$$

and we obtain the asymptotics which are valid for all $|x|, |y| \le 1 - \epsilon$,

$$K_N^\omega(x,y) = \frac{\sin\big[N\pi\big(F_\varrho(x) - F_\varrho(y)\big)\big]}{\pi(x-y)} + \underset{N\to\infty}{O_\epsilon}\left(1 + \frac{N^{-1}}{|x-y|}\right). \tag{3.48}$$

The error term in formula (3.48) is not uniform. However, we may observe that in the regime $|x - y| \le 1/N$, the asymptotics of the correlation kernel already follow from *local universality considerations*. Local asymptotics for Christoffel-Darboux kernels have been extensively studied and are known in great generality, see e.g. the work of Lubinsky [34] or the surveys [43, 35]. Moreover, they can be derived without using the Riemann-Hilbert formulation for the OPs. Local universality is usually formulated stating that for any $L, \epsilon > 0$,

$$\frac{1}{N}K_N^\omega\left(x_0 + \frac{\xi}{N}, x_0 + \frac{\zeta}{N}\right) = \frac{\sin\big[\pi\varrho(x_0)(\xi - \zeta)\big]}{\pi(\xi - \zeta)} + \underset{N\to\infty}{O_{\epsilon,L}}\big(N^{-1}\big), \tag{3.49}$$

for all $|x_0| < 1 - \epsilon$ and all $\xi, \zeta \in [-L, L]$. Now, by *reverse engineering* the argument of remark 2.4 (cf. the Taylor expansion (2.19)), we deduce from formula (3.49) that if $|x - y| \le 1/N$,

$$K_N^\omega(x,y) = \frac{\sin\big[N\pi\big(F_\varrho(x) - F_\varrho(y)\big)\big]}{\pi(x-y)} + \underset{N\to\infty}{O_\epsilon}(1).$$

Combining these asymptotics and formula (3.48), it completes the proof. $\qquad\square$

4 Regularized characteristic polynomial and log-correlated Gaussian processes

In [21], it was first established that the logarithm of the modulus of the characteristic polynomial of a CUE random matrix converges weakly to a random generalized function on the unit circle whose correlation kernel has a logarithmic singularity. In [17], Fyodorov et al. recently explained to how to perform an analogous construction at mesoscopic scales for a certain regularization of the characteristic polynomial of a GUE matrix. Namely, let $0 < \alpha < 1$, $x_0 \in \mathbb{R}$, $\eta > 0$, $z_t = t + i\eta$, and define

$$W_N(t) = \log\big|\det\big[\mathrm{H} - x_0 - z_t N^{-\alpha}\big]\big| - \log\big|\det\big[\mathrm{H} - x_0 - z_0 N^{-\alpha}\big]\big|.$$

Then, if $\mathrm{H} \sim \mathbb{P}_N^{\mathrm{GUE}}$ is a GUE matrix, (3.1), the main result of [17] states that the random process $t \mapsto W_N(t) - \mathbb{E}[W_N(t)]$ converges weakly in $L^2[a,b]$ (for any $a, b \in \mathbb{R}$) to a Gaussian Gaussian process B_η defined as follows:

Definition 4.1. *The η-regularized fractional Brownian motion with Hurst exponent $H = 0$ is a real-valued Gaussian process B_η characterized by the following properties:*

    i) B_η is a continuous process with mean 0 and $B_\eta(0) = 0$ almost surely.

    ii) B_η has stationary increments.

    iii) $\mathrm{Var}\left[B_\eta(t)\right] = \dfrac{1}{2}\log\left(1 + \dfrac{t^2}{4\eta^2}\right)$ *for any $t \in \mathbb{R}$.*

In particular, the process B_η converges as $\eta \to 0$ to a Gaussian generalized function whose correlation kernel has a logarithmic singularity along the diagonal; we refer to [17] for some background and references on fractional Brownian motion. Let us just point out that the process B_η has the following representation, for any $t \in \mathbb{R}$,

$$B_\eta(t) = \Re\left\{\int_0^\infty e^{-\eta s}(e^{-its} - 1)\frac{dZ_s}{\sqrt{2s}}\right\},$$

where Z is a complex Brownian motion with unit variance. Inspired by certain Fisher–Hartwig asymptotics obtained by Krasovsky in [28], the authors of [17] computed the limits of the Laplace transform of the random variable $W_N(t)$ for any $t \in \mathbb{R}$ and show that the finite-dimentionnal distributions of $W_N - \mathbb{E}\left[W_N\right]$ converges to that of B_0. In the following, using the central limit theorem 3.5, we are able to generalize their result to other unitary invariant ensembles. In fact, theorem 4.2 should be true in the multi-cut situation as well, but we have not pushed the estimates of section 3.3 in this case.

Theorem 4.2. *Let ω be any weight satisfying* (3.3) *and the one-cut condition* $\mathrm{supp}(\varrho_\omega) = [-1, 1]$ *and let* $\mathrm{H} \sim \mathbb{P}_N^\omega$, (3.2)*. Then, for any $|x_0| < 1$, any $0 < \alpha < 1$, the stochastic process $t \mapsto W_N(t) - \mathbb{E}_N^\omega\left[W_N(t)\right]$ converges weakly as $N \to \infty$ in $L^2[a, b]$ (for any $a, b \in \mathbb{R}$) to the Gaussian process B_η characterized by definition 4.1.*

First, let us observe that the random variable $W_N(t)$ is a linear statistic:

$$\begin{aligned}
W_N(t) &= \Re\left\{\log\det\left[\frac{M - x_0 - z_t N^{-\alpha}}{M - x_0 - z_0 N^{-\alpha}}\right]\right\}\\
&= \Re\left\{\mathrm{Tr}\left[\log\left(\frac{M - x_0 - z_t N^{-\alpha}}{M - x_0 - z_0 N^{-\alpha}}\right)\right]\right\}\\
&= \Xi_N^{x_0,\alpha} g_t
\end{aligned} \tag{4.1}$$

where the function $g_t(x) = \Re\left\{\log\left(\dfrac{x - z_t}{x - z_0}\right)\right\}$ is defined using the principal branch of the logarithm and $z_t = t + i\eta$. It is easily seen that, even though $g_t \notin L^1(\mathbb{R})$, its Fourier transform is well defined in $L^2(\mathbb{R})$ and, by lemma 4.3 below, it is given by

$$\widehat{g_t}(u) = (1 - e^{-2\pi i u t})\frac{e^{-2\pi|u|\eta}}{2|u|}. \tag{4.2}$$

Lemma 4.3. *For any $\eta > 0$ and $x, t \in \mathbb{R}$, one has*

$$\int_{\mathbb{R}} e^{2\pi i u x}(1 - e^{-2\pi i u t})\frac{e^{-2\pi\eta|u|}}{2|u|}\,du = g_t(x) = \frac{1}{2}\log\left(\frac{(x - t)^2 + \eta^2}{x^2 + \eta^2}\right).$$

Proof. Proof. This identity is classical and it can be proved by observing that, for any $t > 0$,

$$\frac{1 - e^{-2\pi i t u}}{2|u|} = i\,\mathrm{sgn}(u)\pi\int_0^t e^{-2\pi i s u}\,ds,$$

and, by Fubini's theorem,

$$
\int_{\mathbb{R}} e^{2\pi iux}(1 - e^{-2\pi iut})\frac{e^{-2\pi\eta|u|}}{2|u|}du = i\pi \int_0^t \int_{\mathbb{R}} e^{2\pi iu(x-s)}\operatorname{sgn}(u)e^{-2\pi\eta|u|}du\,ds
$$

$$
= -2\pi \int_0^t \Im\left\{\int_0^\infty e^{-2\pi u(\eta - i(x-s))}du\right\}ds
$$

$$
= -\Im\left\{\int_0^t \frac{ds}{\eta - i(x-s)}\right\}.
$$

We conclude by observing that, by definition, for any $t > 0$, $g_t(x) = \Re\left\{\int_{x-t}^x \frac{dv}{v + i\eta}\right\}$. The proof in the case $t < 0$ is almost identical. $\qquad\square$

Formula (4.2) shows that the functions $g_t \in H^{1/2}(\mathbb{R})$. Indeed, one has

$$
\widehat{g_t}(u)\widehat{g_s}(-u)|u| = (1 - e^{-i2\pi ut} - e^{i2\pi us} + e^{i2\pi u(s-t)})\frac{e^{-4\pi|u|\eta}}{4|u|}.
$$

According to formula (1.4) and lemma 4.3 with $x = 0$, we obtain for any $t, s \in \mathbb{R}$

$$
\langle g_t, g_s \rangle_{H^{1/2}} = \int_{\mathbb{R}} \widehat{g_t}(u)\widehat{g_s}(-u)|u|
$$

$$
= \frac{1}{4}\left\{\log\left(1 + \frac{t^2}{4\eta^2}\right) + \log\left(1 + \frac{s^2}{4\eta^2}\right) - \log\left(1 + \frac{(t-s)^2}{4\eta^2}\right)\right\}. \tag{4.3}
$$

Hence, by (4.1) and theorem 3.5, it suffices to check that the functions g_t satisfies the condition (H.1) to deduce that for any $t \in \mathbb{R}$,

$$
W_N(t) - \mathbb{E}_N^\omega\left[W_N(t)\right] \underset{N\to\infty}{\Longrightarrow} \mathcal{N}\left(0, \frac{1}{2}\log\left(1 + \frac{t^2}{4\eta^2}\right)\right).
$$

By remark 3.6, this follows immediately from the facts that $g_t \in C^1(\mathbb{R})$ and for any $t \in \mathbb{R}$,

$$
g_t'(x) = \Re\left\{\frac{-t}{(x - z_0)(x - z_t)}\right\} \underset{|x|\to\infty}{=} O\left(\frac{1}{x^2}\right).
$$

In fact, given $t_1 < \cdots < t_k$ and $\xi_1, \ldots, \xi_k \in \mathbb{R}$, the test function $f = \sum_{j=1}^k \xi_j g_{t_j}$ also satisfies the assumption of theorem 3.5 and according to definition 4.1, a similar argument shows that for any $k \in \mathbb{N}$,

$$
\left(W_N(t_1) - \mathbb{E}_N^\omega\left[W_N(t_1)\right], \ldots, W_N(t_k) - \mathbb{E}_N^\omega\left[W_N(t_k)\right]\right) \underset{N\to\infty}{\Longrightarrow} \left(B_\eta(t_1), \ldots, B_\eta(t_k)\right). \tag{4.4}
$$

Note that the fact that the Gaussian process B_η has independent increments follows immediately from the covariance structure (4.3) and the continuity of its sample paths follows from Kolmogorov's regularity theorem. Following [17, Thm. 2.3], the convergence (4.4) of the finite-dimensional distributions and the estimate (3.39):

$$
\limsup_{N\to\infty} \operatorname{Var}\left[\Xi_N^{x_0,\alpha}f\right] \leq 32\|f\|_{H^{1/2}}^2
$$

allows us to conclude that the random process W_N converges in distribution to B_η in $L^2[a,b]$ for any $a, b \in \mathbb{R}$.

A Variance estimate in the global regime

In this section, we consider the unitary invariant ensemble $\mathbb{P}_N^V$ introduced in theorem 1.1 and we assume that there exists $B > 2$ and $\eta > 0$ so that for all $|x| > B$.

$$V(x) \geq 2(1 + \eta) \log |x|. \tag{A.1}$$

We also suppose that the potential V satisfies the one-cut condition and $J_V = (-1, 1)$ so that we may use the results of section 3.3. The aim of this appendix is to derive an estimate for the variance of global linear statistics valid for rather general continuously differentiable test functions.

Proposition A.1. *Suppose that the potential V satisfies the assumptions above. We denote $\Xi_N h = \sum h(\lambda_k)$ where the sum is over the eigenvalues of a random matrix distributed according to $\mathbb{P}_N^V$. If $h \in C^1(\mathbb{R})$ and there exists $Q, n > 0$ so that $|h'(x)| \leq Q|x|^n$ for all $|x| \geq 1$, then*

$$\limsup_{N \to \infty} \operatorname{Var}\left[\Xi_N h\right] \leq 16\widetilde{\Sigma}(h)^2, \tag{A.2}$$

where

$$\tilde{\Sigma}(h)^2 = \frac{1}{\pi^2} \iint\limits_{[-1,1]^2} \left|\frac{h(x) - h(y)}{x - y}\right|^2 \frac{dx\,dy}{\sqrt{1 - x^2}\sqrt{1 - y^2}}. \tag{A.3}$$

Note that since the point process Ξ_N is supported in a small neighborhood of J_V with very high probability, we expect that, apart form some mild growth assumption, the behavior of the test function h outside of J_V should be irrelevant to estimate the variance of the linear statistics $\Xi_N h$. Moreover, as Johansson put forward in [23], the only important regularity condition should be that the quantity $\Sigma(h) < \infty$, see (1.8). However, mostly because of the effects of the spectral edges, it is a very difficult task to obtain an estimate of the form $\operatorname{Var}\left[\Xi_N h\right] \leq C\Sigma(h)^2$ under optimal conditions. It seems that this question remains open even for the GUE; see [48, Theorem 2] for a very general result. Note also that even though the asymptotic variance $\Sigma(f) \leq \widetilde{\Sigma}(f)$ in the estimate (A.2), if it exists, it is difficult to exhibit a function $h \in H^{1/2}(\mathbb{R})$ such that $\Sigma(h) < \infty$ and $\widetilde{\Sigma}(h) = \infty$.

The remainder of this section is devoted to the proof of proposition A.1. The main motivation for these variance estimates is to extend a CLT from [31] from polynomials to general test functions; see theorem A.4 below. We say that a real-valued function f belong to the space $\mathcal{H}_0^{1/2}$ and we denote $f \in \mathcal{H}_0^{1/2}$ if $f \in L^\infty(\mathbb{R})$ with compact support and

$$\iint\limits_{[-1,1]^2} \left|\frac{f(x) - f(y)}{x - y}\right|^2 dx\,dy < \infty.$$

Lemma A.2. *If $f \in \mathcal{H}_0^{1/2}$ and satisfies the condition* (H. 2), *then $f \in H^{1/2}(\mathbb{R})$.*

Proof. Suppose that $\operatorname{supp}(f) \subseteq [-A, A]$ and let $\mathcal{K} = \left\{|x| \leq A, 1 \leq |y| \leq A + 1\right\}$ and $\mathcal{B} = [-A, A] \times [A + 1, \infty)$. By symmetry, we have

$$\|f\|_{H^{1/2}}^2 \leq \iint\limits_{[-1,1]^2} \left|\frac{f(x) - f(y)}{x - y}\right|^2 dx\,dy + 2\iint\limits_{\mathcal{K}} \left|\frac{f(x) - f(y)}{x - y}\right|^2 dx\,dy + 4\iint\limits_{\mathcal{B}} \left|\frac{f(x) - f(y)}{x - y}\right|^2 dx\,dy.$$

By definition, the first term is finite. Since f satisfies the condition (H.2), the second term is bounded by $4A^2L$. By construction of the set $\mathcal{B}$, the third term satisfies

$$\iint\limits_{\mathcal{B}} \left|\frac{f(x) - f(y)}{x - y}\right|^2 dx\,dy \leq \iint\limits_{\mathcal{B}} \left|\frac{f(x)}{y - A}\right|^2 dx\,dy$$

$$\leq 2A\|f\|_{L^\infty}^2,$$

and we conclude that $\|f\|_{H^{1/2}}^2 < \infty$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

The proof of proposition A.1 is based on the result of proposition 3.9 and the exponential decay of the Christoffel-Darboux kernel outside of the bulk; see lemma A.3 below. We suppose that $h \in C^1(\mathbb{R})$ in order to simplify the proof, however this condition is not necessary. In fact, by a simple modification of our method, it suffices to suppose that $h \in \mathcal{H}^{1/2}$ and there exists $Q > 0$ and $n > 0$ so that for all $|x| > 1 - \delta$,

$$\sup\left\{ \left|\frac{h(x) - h(y)}{x - y}\right| : |y| \le |x| \right\} \le Q|x|^n.$$

Lemma A.3. *Under the assumption of proposition A.1, if we also suppose that $h(x) = 0$ for all $|x| \le \mathrm{B}$. Then, there exists $C > 0$ so that*

$$\mathrm{Var}\left[\Xi_N h\right] \le C\mathrm{B}^{-\eta N}.$$

Proof. By [13, formula 1.58], for any $\epsilon > 0$, we have

$$\left|\Phi_N(x)\right| \le \left(\frac{1}{2\sqrt{\pi}}\left|\frac{x+1}{x-1}\right|^{1/4} + \underset{N\to\infty}{O_\epsilon}\left(N^{-1}\right)\right)e^{-N\mathfrak{H}_V(x)}, \qquad \forall |x| > 1 + \epsilon, \qquad (\mathrm{A.4})$$

where for all $x \in \mathbb{R}$,

$$\mathfrak{H}_V(x) = \frac{V(x) + \ell}{2} - \int \log|x - s|\varrho_V(s)ds \qquad \text{and} \qquad \ell \in \mathbb{R}.$$

This function appears in the determination of the equilibrium density ϱ_V. In fact, $\varrho_V(x)dx$ is the unique minimizer of a weighted energy functional and it is uniquely determined by the following Euler-Lagrange conditions:

$$\begin{cases} \mathfrak{H}_V(x) = 0 & \forall x \in \overline{J_V} \\ \mathfrak{H}_V(x) \ge 0 & \forall x \in \mathbb{R}\backslash\overline{J_V}. \end{cases}$$

Moreover, since $\mathrm{supp}(\varrho_V) = [-1, 1]$, we have for all $|x| > 1$

$$\int \log|x - s|\varrho_V(s) \le \log(2|x|).$$

Hence, if the potential $V(x)$ satisfies the condition (A.1), then $\mathfrak{H}_V(x) \ge \eta \log|x| + \frac{\ell}{2} - \log 2$ for all $|x| > \mathrm{B}$. In fact, choosing a larger constant B if necessary, we can suppose that $\mathfrak{H}_V(x) \ge \frac{\eta \log|x|}{2}$. By formula (A.4), this implies that there exists $C > 0$ so that for all $|x| > \mathrm{B}$,

$$\left|\Phi_N(x)\right| \le \sqrt{C/2}\, e^{-N\eta \log|x|/2}. \qquad\qquad\qquad\qquad\qquad (\mathrm{A.5})$$

Using [13, formula 1.59] instead, we can show that the estimate (A.5) holds for the function Φ_{N-1} as well. By formula (3.31), this implies that for all $|x| \ge \mathrm{B}$,

$$\left|K_N^\omega(x,y)\right|^2 \le C\frac{\gamma_{N-1}}{\gamma_N}\frac{\left|\Phi_{N-1}(y)\right|^2 + \left|\Phi_N(y)\right|^2}{|x - y|^2}e^{-N\eta \log|x|}.$$

Hence, since $\|\Phi_N\|_{L^2} = \|\Phi_{N-1}\|_{L^2} = 1$, we obtain for all $|x| \ge \mathrm{B}$,

$$\int_{\mathbb{R}} \left|(x - y)K_N^\omega(x,y)\right|^2 dy \le C\frac{\gamma_{N-1}}{\gamma_N}e^{-N\eta \log|x|}. \qquad\qquad\qquad (\mathrm{A.6})$$

On the other hand, by assumptions, we have for all $|y| \leq |x|$,

$$\left| \frac{h(x) - h(y)}{x - y} \right| \leq \mathbb{1}_{|x| > \mathrm{B}} \sup \left\{ h'(t) : |t| \leq |x| \right\}$$

$$\leq Q|x|^n \mathbb{1}_{|x| > \mathrm{B}}.$$

According to formula (3.29) and (A.6), we obtain

$$\mathrm{Var}\left[\Xi_N h\right] = \frac{1}{2} \iint |h(x) - h(y)|^2 \left| K_N^\omega(x, y) \right|^2 dx dy$$

$$\leq \frac{Q^2}{2} \int_{\mathbb{R} \setminus [-\mathrm{B}, \mathrm{B}]} |x|^{2n} \left(\int_{\mathbb{R}} \left| (x - y) K_N^\omega(x, y) \right|^2 dy \right) dx$$

$$\leq CQ^2 \frac{\gamma_{N-1}}{\gamma_N} \int_{\mathrm{B}}^{\infty} x^{2n} e^{-N\eta \log(x)} dx. \tag{A.7}$$

Because of the asymptotics (3.27), $C := CQ^2 \sup\limits_{N \in \mathbb{N}} \{ \frac{\gamma_{N-1}}{\gamma_N} \} \mathrm{B}^{2n} < \infty$ and this completes the proof. $\qquad\square$

Proof of proposition A.1. Let $\mathrm{A} > \mathrm{B}$ and $\chi \in C^1\left(\mathbb{R}_+ \to [0, 1]\right)$ such that $-\chi' \in [0, 1]$ and

$$\chi(x) = \begin{cases} 1 & \text{if } x \leq \mathrm{B} \\ 0 & \text{if } x \geq \mathrm{A} \end{cases}.$$

We decompose $h = f + g$ where $f = \chi h$ has compact support and $g = (1 - \chi)h$, then one has

$$\mathrm{Var}\left[\Xi_N h\right] \leq 2 \left(\mathrm{Var}\left[\Xi_N f\right] + \mathrm{Var}\left[\Xi_N g\right] \right). \tag{A.8}$$

First, since $g(x) = 0$ for all $|x| \leq \mathrm{B}$ and $|g'(x)| \leq |h(x)| + |h'(x)|$, by assumptions there exists a constant $\tilde{Q}$ so that $|g'(x)| \leq \tilde{Q}|x|^{n+1}$ for all $|x| \geq 1$. Then, by lemma A.3,

$$\limsup_{N \to \infty} \mathrm{Var}\left[\Xi_N g\right] = 0. \tag{A.9}$$

Next, we will show that the function f satisfies the condition (H.2). By definition, one has

$$\left| \frac{f(x) - f(y)}{x - y} \right| \leq \left| \frac{h(x) - h(y)}{x - y} \right| \chi(x) + |h(y)| \left| \frac{\chi(x) - \chi(y)}{x - y} \right|.$$

Hence, using the properties of the cutoff function χ, one has

$$\sup_{\substack{|y| \leq |x| \\ |x| < \mathrm{A}+1}} \left| \frac{f(x) - f(y)}{x - y} \right| \leq \sup_{|t| \leq \mathrm{A}+1} \left\{ |h'(t)| + |h(t)| \right\}.$$

On the other hand, if $|x| \geq \mathrm{A} + 1$, for all $|y| \leq |x|$,

$$\left| \frac{f(x) - f(y)}{x - y} \right| \leq \frac{|h(y)| \chi(y)}{|x - y|} \leq \begin{cases} 0 & \text{if } |y| \geq \mathrm{A} \\ \sup \left\{ |h(t)| : |t| \leq \mathrm{A} \right\} & \text{else} \end{cases}$$

Hence, there exists $\mathrm{L} > 0$ so that

$$\sup_{x \in \mathbb{R}} \left\{ \left| \frac{f(x) - f(y)}{x - y} \right| : |y| \leq |x| \right\} = \mathrm{L},$$

and, by symmetry, the function f satisfies the condition (H.2). Moreover, by lemma A.2, $f \in H^{1/2}(\mathbb{R})$ and by proposition 3.9, this implies that

$$\limsup_{N \to \infty} \mathrm{Var}\left[\Xi_N f\right] \leq 8\tilde{\Sigma}(f)^2. \tag{A.10}$$

Combining the estimates (A.8–A.10), we conclude that

$$\limsup_{N\to\infty} \mathrm{Var}\left[\Xi_N h\right] \leq 16\tilde{\Sigma}(f)^2.$$

It completes the proof since $\tilde{\Sigma}(f) = \tilde{\Sigma}(h)$ because $h(x) = f(x)$ for all $|x| \leq 1$. □

Proposition A.1 is used in [31] to give a new proof of the CLT (1.7). Moreover, for general one-cut potential, theorem A.4 is to the author's knowledge up to now valid for the most general class of test functions.

Theorem A.4. *Let $V : \mathbb{R} \to \mathbb{R}$ be a real-analytic function which satisfies the condition $(A.1)$ and such that $J_V = (-1,1)$. If $(\lambda_1, \ldots, \lambda_N)$ denote the eigenvalues of a random matrix distributed according to $\mathbb{P}_N^V$, then for any $f \in C^1(\mathbb{R})$ such that there exists $Q, n > 0$ so that $|f'(x)| \leq Q|x|^n$ for all $|x| \geq 1$, one has*

$$\sum_{k=1}^{N} f(\lambda_k) - \mathbb{E}\left[\sum_{k=1}^{N} f(\lambda_k)\right] \underset{N\to\infty}{\Longrightarrow} \mathcal{N}\big(0, \Sigma(f)^2\big). \tag{A.11}$$

The proof relies on the following approximation result:

Lemma A.5 ([48], Lemma 2.1)**.** *Let $\mathfrak{F}(\mathbb{R})$ be a vector space of functions equipped with a seminorm $\tilde{\Sigma}$ and let Ξ_N be a sequence of point processes on $\mathbb{R}$ such that for all $f \in \mathfrak{F}(\mathbb{R})$,*

$$\limsup_{N\to\infty} \mathrm{Var}\left[\Xi_N f\right] \leq C\tilde{\Sigma}(f)^2.$$

If there exists a subspace $\mathfrak{X}$ which is dense in $\mathfrak{F}(\mathbb{R})$ such that for all $g \in \mathfrak{X}$,

$$\Xi_N g - \mathbb{E}\left[\Xi_N g\right] \underset{N\to\infty}{\Longrightarrow} \mathcal{N}\big(0, \Sigma(g)^2\big), \tag{A.12}$$

where $\Sigma(g) \leq \tilde{\Sigma}(g)$, then the CLT (A.12) holds as well for all test function $f \in \mathfrak{F}(\mathbb{R})$.

The formulation is slightly more general than in [48] but the proof which is based on the characteristic function method and Lévy's theorem is exactly the same.

Proof of theorem A.4. Let

$$\mathfrak{F}(\mathbb{R}) = \big\{ f \in C^1(\mathbb{R}) : \exists\, n, Q, C > 0 \text{ so that } |f'(x)| \leq C + Q|x|^n \big\}.$$

It is not difficult to check that the functional $\tilde{\Sigma}$ given by (A.3) defines a seminorm on $\mathfrak{F}(\mathbb{R})$ and, by using Weierstrass approximation theorem, to check that real-valued polynomials are dense in $\mathfrak{F}(\mathbb{R})$. Now, if (1.7) holds for any polynomials (this follows e.g. by [31, Theorem 2.6]), the estimate (A.2) combined with lemma A.5 implies that the CLT must hold for all test functions in the class $\mathfrak{F}(\mathbb{R})$. □

Remark A.6. For the GUE kernel, using Cramér's inequality, $\|\phi_k\|_\infty \leq 1$ for all $k \geq 0$, so that

$$\left|K_N^{\mathrm{GUE}}(x,y)\right| \leq N, \qquad \forall x, y \in \mathbb{R}.$$

Moreover, by Theorem 5.2.3 in [40], for any $\epsilon > 0$ there exists $\beta, C > 0$ so that $K_N^{\mathrm{GUE}}(x,x) \leq CNe^{-\beta Nx^2}$ for all $|x| \geq 1 + \epsilon$. Hence, by the Cauchy-Schwartz inequality, we obtain for all $|x| \geq 1 + \epsilon$ and $y \in \mathbb{R}$,

$$\left|K_N^{\mathrm{GUE}}(x,y)\right|^2 \leq K_N^{\mathrm{GUE}}(y,y) K_N^{\mathrm{GUE}}(x,x)$$
$$\leq CN^2 e^{-\beta Nx^2}.$$

This implies that proposition A.1 and the CLT (A.11) hold for any test function $h(x) = \underset{x\to\infty}{o}\big(e^{|x|^\alpha}\big)$ with $0 < \alpha < 2$ and such that there exists $0 < \delta < 1$ and $\mathrm{L} > 0$ so that

$$\sup\left\{ \left|\frac{h(x) - h(y)}{x - y}\right| : |y| \leq |x|,\ 1 - \delta < |x| < 1 + \delta \right\} \leq \mathrm{L}.$$

References

[1] Anderson, G. W.; Guionnet, A.; Zeitouni, O.: An introduction to random matrices. Cambridge Studies in Advanced Mathematics, 118. Cambridge University Press, Cambridge, 2010. MR-2760897

[2] Borodin, A.: Determinantal point processes. The Oxford handbook of random matrix theory, 231–249, Oxford Univ. Press, Oxford, 2011. MR-2932631

[3] Borot, G.; Guionnet, A: Asymptotic expansion of β matrix models in the one-cut regime. Comm. Math. Phys. 317 (2013), no. 2, 447–483. MR-3010191

[4] Borot, G.; Guionnet, A.: Asymptotic expansion of β matrix models in the multi-cut regime, arXiv:1303.1045

[5] Bourgade, P.; Erdős, L.; Yau, H.-T.; Yin, J.: Fixed energy universality for generalized Wigner matrices. Comm. Pure Appl. Math. 69 (2016), no. 10, 1815–1881. MR-3541852

[6] Bourgade, P.; Kuan, J.: Strong Szegő asymptotics and zeros of the zeta-function. Comm. Pure Appl. Math. 67 (2014), no. 6, 1028–1044 MR-3193966

[7] Boutet de Monvel, A.; Khorunzhy, A.: Asymptotic distribution of smoothed eigenvalue density I. Gaussian random matrices. Random Oper. Stochastic Equations 7 (1999), no. 1, 1–22. MR-1678012

[8] Boutet de Monvel, A.; Khorunzhy, A.: Asymptotic distribution of smoothed eigenvalue density II. Wigner random matrices. Random Oper. Stochastic Equations 7 (1999), no. 2, 149–168. MR-1689027

[9] Breuer, J.; Duits, M.: The Nevai condition and a local law of large numbers for orthogonal polynomial ensemble. Adv. Math. 265 (2014), 441–484. MR-3255467

[10] Breuer, J.; Duits, M.: Universality of mesoscopic fluctuations for orthogonal polynomial ensembles. Comm. Math. Phys. 342 (2016), no. 2, 491–531. MR-3459158

[11] Breuer, J.; Duits, M.: Central limit theorems for biorthogonal ensembles and asymptotics of recurrence coefficients. J. Amer. Math. Soc. 30 (2017), no. 1, 27–66. MR-3556288

[12] Deift, P.: Orthogonal polynomials and random matrices: a Riemann-Hilbert approach. Courant Lecture Notes in Mathematics, 3. New York University, Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, RI, 1999. MR-1677884

[13] Deift, P.; Kriecherbauer, T.; McLaughlin, K. T.-R.; Venakides, S.; Zhou, X.: *Uniform asymptotics for polynomials orthogonal with respect to varying exponential weights and applications to universality questions in random matrix theory*. Comm. Pure Appl. Math. 52 (1999), no. 11, 1335–1425. MR-1702716

[14] Deift, P.; Zhou, X.: A steepest descent method for oscillatory Riemann-Hilbert problems. Bull. Amer. Math. Soc. (N.S.) 26 (1992), no. 1, 119–123. MR-1108902

[15] Duplantier, B.; Rhodes, R.; Sheffield, S.; Vargas, V.: Log-correlated Gaussian Fields: an overview in Geometry, Analysis and Probability. Vol. 310 of the series Progress in Mathematics, 191–216, 2017.

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

[17] Fyodorov, Y. V.; Khoruzhenko, B. A.; Simm, N. J.: Fractional Brownian motion with Hurst index $H = 0$ and the Gaussian unitary ensemble. Ann. Probab. 44 (2016), no. 4, 2980–3031. MR-3531684

[18] Fyodorov, Y. V.; Simm, N. J.: On the distribution of the maximum value of the characteristic polynomial of GUE random matrices. Nonlinearity 29 (2016), no. 9, 2837–2855. MR-3544809

[19] Hardy, A.: Average characteristic polynomials of determinantal point processes. Ann. Inst. Henri Poincaré Probab. Stat. 51 (2015), no. 1, 283–303. MR-3300971

[20] Hough, J. B.; Krishnapur, M.; Peres, Y.; Virág, B.: Determinantal processes and independence. Probab. Surv. 3 (2006), 206–229. MR-2216966

[21] Hughes, C. P.; Keating, J. P.; O'Connell, N.: On the characteristic polynomial of a random unitary matrix. Comm. Math. Phys. 220 (2001), no. 2, 429–451. MR-1844632

[22] Janson, S.: Gaussian Hilbert spaces. Cambridge Tracts in Mathematics, 129. Cambridge University Press, Cambridge, 1997. MR-1474726

[23] Johansson, K.: On fluctuations of eigenvalues of random Hermitian matrices. Duke Math. J. 91 (1998), no. 1, 151–204. MR-1487983

[24] Johansson, K.: Random matrices and determinantal processes. Mathematical statistical physics, 1–55, Elsevier B. V., Amsterdam, 2006. MR-2581882

[25] Johansson, K., Lambert G.: Gaussian and non-Gaussian fluctuations for mesoscopic linear statistics in determinantal processes, arXiv:1504.06455

[26] He, Y.; Knowles, A.: Mesoscopic eigenvalue statistics of Wigner matrices. Ann. Appl. Probab. 27 (2017), no. 3, 1510–1550. MR-3678478

[27] König, W.: *Orthogonal polynomial ensembles in probability theory*. Probab. Surv. 2 (2005), 385–447. MR-2203677

[28] Krasovsky, I. V.: Correlations of the characteristic polynomials in the Gaussian unitary ensemble or a singular Hankel determinant. Duke Math. J. 139 (2007), no. 3, 581–619. MR-2350854

[29] Kuijlaars, A. B. J.; McLaughlin, K. T.-R.; Van Assche, W.; Vanlessen, M.: The Riemann-Hilbert approach to strong asymptotics for orthogonal polynomials on [?1,1]. Adv. Math. 188 (2004), no. 2, 337–398. MR-2087231

[30] Kuijlaars, A. B. J.; Vanlessen, M.: Universality for eigenvalue correlations from the modified Jacobi unitary ensemble. Int. Math. Res. Not. 2002, no. 30, 1575–1600. MR-1912278

[31] Lambert, G.: CLT for biorthogonal ensembles and related combinatorial identities, arXiv:1511.06121

[32] Levin, E., Lubinsky, D. S.: Universality limits in the bulk for varying measures. Adv. Math. 219 (2008), 743–779. MR-2442052

[33] Levin, E., Lubinsky, D. S.: Universality limits for exponential weights. Constr. Approx. 29 (2009), 247–275. MR-2481591

[34] Lubinsky, D. S.: A new approach to universality limits involving orthogonal polynomials. Ann. of Math. (2) 170 (2009), no. 2, 915–939. MR-2552113

[35] Lubinsky, D. S.: Some recent methods for establishing universality limits. Nonlinear Anal. 71 (2009), no. 12, 2750–2765. MR-2672047

[36] Lyons, R.: Determinantal probability: Basic properties and conjectures in Proceedings of the Intl. Congress Math., vol. IV, Seoul, 2014.

[37] Mehta, M. L., Gaudin, M.: On the density of eigenvalues of a random matrix. Nuclear Phys. 18 (1960), 420–427.

[38] Pastur, L.: Limiting laws of linear eigenvalue statistics for Hermitian matrix models. J. Math. Phys. 47 (2006), no. 10, 103303. MR-2268864

[39] Pastur, L.; Shcherbina, M.: Universality of the local eigenvalue statistics for a class of unitary invariant random matrix ensembles. J. Statist. Phys. 86 (1997), no. 1-2, 109–147. MR-1435193

[40] Pastur, L.; Shcherbina, M.: Eigenvalue distribution of large random matrices. Mathematical Surveys and Monographs, 171. American Mathematical Society, Providence, RI, 2011. MR-2808038

[41] Plancherel; M.; Rotach, W.: Sur les valeurs asymptotiques des polynomes d'Hermite $H_n(x) = (-1)^n e^{\frac{x^2}{2}} \frac{d^n}{dx^n}(e^{-\frac{x^2}{2}})$. Comment. Math. Helv. 1 (1929), 227–254. MR-1509395

[42] Rodgers, B.: A central limit theorem for the zeroes of the zeta function. Int. J. Number Theory 10 (2014), no. 2, 483–511. MR-3189991

[43] Simon, B.: The Christoffel-Darboux kernel. Perspectives in partial differential equations, harmonic analysis and applications, 295–335, Proc. Sympos. Pure Math. 79. Amer. Math. Soc., Providence, RI, 2008. MR-2500498

[44] Schubert, K.; Venker, M.: Empirical spacings of unfolded eigenvalues. Electron. J. Probab. 20 (2015), no. 120, 37 pp. MR-3425540

[45] Soshnikov, A.: The Central Limit Theorem for local linear statistics in classical compact groups and related combinatorial identities. Ann. Probab. 28 (2000), 1353–1370. MR-1797877

[46] Soshnikov, A.: Determinantal random point fields. (Russian) Uspekhi Mat. Nauk 55 (2000), no. 5(335), 107–160; translation in Russian Math. Surveys 55 (2000), no. 5, 923–975. MR-1799012

[47] Soshnikov, A.: Gaussian limit for determinantal random point fields. Ann. Probab. 30 (2001), 1–17. MR-1894104

[48] Sosoe, P., Wong, P.: Regularity conditions in the CLT for linear eigenvalue statistics of Wigner matrices. Adv. Math. 249 (2013), 37–87. MR-3116567

[49] Szegő, G.: Orthogonal polynomials. Fourth edition. American Mathematical Society, Colloquium Publications, Vol. XXIII. American Mathematical Society, Providence, R.I., 1975. MR-0372517.

[50] Webb, C.: The characteristic polynomial of a random unitary matrix and Gaussian multiplicative chaos — the L2-phase. Electron. J. Probab. 20 (2015), no. 104, 21 pp. MR-3407221

Electronic Journal of Probability
Electronic Communications in Probability

Advantages of publishing in EJP-ECP

- Very high standards

- Free for authors, free for readers

- Quick publication (no backlog)

- Secure publication (LOCKSS[1])

- Easy interface (EJMS[2])

Economical model of EJP-ECP

- Non profit, sponsored by IMS[3], BS[4] , ProjectEuclid[5]

- Purely electronic

Help keep the journal free and vigorous

- Donate to the IMS open access fund[6] (click here to donate!)

- Submit your best articles to EJP-ECP

- Choose EJP-ECP over for-profit journals

---

[1] LOCKSS: Lots of Copies Keep Stuff Safe http://www.lockss.org/
[2] EJMS: Electronic Journal Management System http://www.vtex.lt/en/ejms.html
[3] IMS: Institute of Mathematical Statistics http://www.imstat.org/
[4] BS: Bernoulli Society http://www.bernoulli-society.org/
[5] Project Euclid: https://projecteuclid.org/
[6] IMS Open Access Fund: http://www.imstat.org/publications/open.htm