Incomplete Determinantal Processes: From Random Matrix to Poisson Statistics

Gaultier Lambert[1] (ORCID)

Abstract

We study linear statistics of a class of determinantal processes which interpolate between Poisson and GUE/Ginibre statistics in dimension 1 or 2. These processes are obtained by performing an independent Bernoulli percolation on the particle configuration of a log-gas confined in a general potential. We show that, depending on the expected number of deleted particles, there is a universal transition for mesoscopic linear statistics. Namely, at small scales, the point process behave according to random matrix theory, while, at large scales, it needs to be renormalized because the variance of any linear statistic diverges. The crossover is explicitly characterized as the superposition of a H^1- or $H^{1/2}$-correlated Gaussian noise depending on the dimension and an independent Poisson process. The proof consists in computing the limits of the cumulants of linear statistics using the asymptotics of the correlation kernel of the process.

Keywords Random matrix theory · Determinantal point processes · Cumulants method

1 Introduction and Results

1.1 Introduction

In these notes, we consider a log-gas, also known as β-ensemble or one component plasma, in dimension 1 or 2 at inverse temperature $\beta = 2$. Let $\mathfrak{X} = \mathbb{R}$ equipped with the Lebesgue measure $d\mu = dx$ or $\mathbb{C}$ equipped with the area measure $d\mu = dA = \frac{r dr d\theta}{\pi}$. Let also $V \in C^2(\mathfrak{X})$ be a real-valued function such that for a $\nu > 0$,

$$V(z) \geq (1 + \nu) \log |z| \text{ as } |z| \to \infty. \tag{1.1}$$

✉ Gaultier Lambert
 gaultier.lambert@math.uzh.ch

[1] Institut für Mathematik, Universität Zürich, Zürich, Switzerland

We consider the probability measure on $\mathfrak{X}^N$ with density $G_N(x) = e^{-\beta \mathscr{H}_V^N(x)}/Z_V^N$ where the Hamiltonian is

$$\mathscr{H}_V^N(x) = \sum_{1 \le i < j \le N} \log |x_i - x_j|^{-1} + N \sum_{j=1}^N V(x_j). \qquad (1.2)$$

Regardless of the dimension d, the condition $\beta = 2$ implies that if a configuration $(\lambda_1, \ldots, \lambda_N)$ is sampled from G_N, then the point process $\Xi := \sum_{k=1}^N \delta_{\lambda_k}$ is determinantal with a correlation kernel

$$K_V^N(z, w) = \sum_{k=0}^{N-1} \varphi_k(z) \overline{\varphi_k(w)}, \qquad (1.3)$$

with respect to μ. Moreover, for all $k \ge 0$,

$$\varphi_k(x) = P_k(x) e^{-NV(x)}, \qquad (1.4)$$

where $\{P_k\}_{k=0}^\infty$ is the sequence of orthonormal polynomials[1] with respect to the weight $e^{-2NV(x)}$ on $L^2(\mu)$. It turns out that for $\beta = 2$, the density G_N also corresponds to the joint law of the eigenvalues of the ensemble of Hermitian (or normal) matrices with weight $e^{-2N \operatorname{Tr} V(M)}$ on $\mathfrak{X} = \mathbb{R}$ (or $\mathfrak{X} = \mathbb{C}$). In particular, when $V(z) = |z|^2$, these correspond to the well-know Gaussian Unitary (GUE) and Ginibre ensembles respectively. It is well known that if the condition (1.1) holds, the thermodynamical limit of the log-gas is described by an equilibrium measure which has compact support. Moreover, if the potential $V \in C^2(\mathfrak{X})$, then the equilibrium measure is absolutely continuous and we let ϱ_V be its density. This implies that for any bounded test function $f \in C(\mathfrak{X})$, as $N \to +\infty$,

$$\frac{1}{N} \mathbb{E}[\Xi(f)] = \int f(x) u_V^N(x) d\mu(x) \to \int f(x) \varrho_V(x) d\mu(x), \qquad (1.5)$$

where the expected density of states is given by $u_V^N(x) = N^{-1} K_V^N(x, x)$. The asymptotics (1.5) follows either from potential theory for general $\beta > 0$ or from the asymptotics of the correlation kernel (1.3) when $\beta = 2$. We refer to [4, Section 2.6] for a proof of the large deviation principle in dimension 1 and to [18] for analogous results for Coulomb gases in higher dimension and further references.

In the following, we consider the problem of describing the fluctuations of the so-called *thinned* log-gases in dimension d = 1, 2. In general, a *thinned* or *incomplete point process* is defined by performing a Bernoulli percolation on the configuration of a the original process. That is the incomplete log-gas, denoted by $\widehat{\Xi}$, is obtained by deleting independently each particle with probability $q_N \in (0, 1)$ or by keeping it with probability $p_N = 1 - q_N$. It turns out that the incomplete process $\widehat{\Xi}$ is also determinantal with correlation kernel $\widehat{K}_V^N(z, w) = p_N K_V^N(z, w)$; see the appendix A for a short proof. In the context of random matrix theory, this procedure was first considered by Bohigas and Pato [8,9] who showed that it gives rise to a crossover to Poisson statistics and the problem of rigorously analyzing this transition in the context of Coulomb gases was popularized by Deift in [23, Problem 2]. Indeed, these types of transitions are supposed to arise in many different contexts in statistical physics, such as the localization/delocalization phenomena, the crossover from chaotic to regular dynamics in the theory of quantum billiards, or in the spectrum of certain band random matrices, see [27,28,50] and reference therein. Although such transitions are believed to be non-universal,

---

[1] For any $k \ge 0$, P_k is a polynomial of degree k and its leading coefficient is positive.

the model of Bohigas and Pato is arguably one of the most tractable to study this phenomenon because it is determinantal. In a different context, the effect of thinning determinantal process on statistical inferences has been recently discussed in [40] and it should be emphasized that the general strategy explained in Sect. 2 applies to more general determinantal processes, see Theorem 2.2. For instance, our method applies to the Sine and the ∞-Ginibre processes which describes the local limits of the log-gases in dimensions 1 and 2 respectively. In fact, this paper is motivated by an analogous result obtained recently by Berggren and Duits for smooth linear statistics of the incomplete Sine and CUE processes [6]. Based on the fact that these processes come from integrable operators, they fully characterized the transition for a large class of mesoscopic linear statistics and suggested that it should be universal for thinned point processes coming from random matrix theory. There are also results for the gap probabilities of the critical thinned ensembles. In [14,15], for the Sine process, Deift et al. computed very detailed asymptotics for the crossover from the *Wigner surmise* to the exponential distribution making rigorous a prediction of Dyson [26], and Charlier–Claeys obtained an analogous result for the CUE [19]. The contribution of this paper is to elaborate on universality for smooth linear statistics of β-ensemble in dimension 1 or 2 when $\beta = 2$. Although our proof relies on the determinantal structure of these models, instead of the connection with Riemann-Hilbert problems used in the previous works, we apply the cumulants method which appears to be very robust to study the asymptotic fluctuations of smooth linear statistics.

Let us point out that based on the theory of [30], an alternative correlation kernel for the incomplete process is

$$\widehat{K}_V^N(z, w) = \sum_{k=0}^{\infty} J_k^N \varphi_k(z) \varphi_k(w), \qquad (1.6)$$

where $(J_k^N)_{k=1}^{\infty}$ is a sequence of i.i.d. Bernoulli random variables with expected values $\mathbb{E}[J_k^N] = p_N \mathbf{1}_{k<N}$. This shows that removing particles builds up randomness in the system and when the disorder becomes sufficiently strong, it will behave like a Poisson process rather than according to random matrix theory.

To keep the analysis as simple as possible, we will restrict ourself to real-analytic V, although the results should be valid for more general potential as well (especially in dimension 1 where the asymptotics of the correlation kernels have been studied in great generality). We keep track of the transition by looking at linear statistics $\widehat{\Xi}(f) = \sum f(\lambda)$ for smooth test functions, where the sum is over the configuration of the incomplete log-gas. The random matrix regime is characterized by the property that the fluctuations of $\widehat{\Xi}(f)$ are of order 1 and described by a *universal Gaussian noise* as the number of particles tends to infinity. On the other hand, in the Poisson regime, the variance of any non-trivial statistic diverges and, once properly renormalized, the point process converges in distribution to a *white noise*. In the remainder of this introduction, we first formulate our assumptions and main results for the fluctuations of the incomplete 2-dimensional log-gases. Then, we will present analogous results in dimension 1.

In what follows, we let $\mathcal{C}_c^k(\mathscr{S})$ be the set of functions with k continuous derivatives and compact support in $\mathscr{S} \subset \mathfrak{X}$ and we use the notation:

$$\partial = (\partial_x - i\partial_y)/2 \,, \ \overline{\partial} = (\partial_x + i\partial_y)/2 \text{ and } \Delta = \partial\overline{\partial}.$$

If ϱ_V is the equilibrium density function, we also denote

$$\mathscr{S}_V := \left\{ x \in \mathfrak{X} : \varrho_V(x) > 0 \right\}. \qquad (1.7)$$

1.2 Main Results for 2-Dimensional Coulomb Gases

If the potential V is real-analytic and satisfies the condition (1.1), then the log-gas lives on the compact set $\overline{\mathscr{S}_V} \subset \mathbb{C}$ which is called the *droplet* and the equilibrium density is given by $\varrho_V = 2\Delta V \mathbf{1}_{\mathscr{S}_V}$. It is also well known that the bulk fluctuations of a two dimensional log-gas around its equilibrium configuration are described by a centered Gaussian process X with correlation structure:

$$\mathbb{E}\big[\mathrm{X}(f)\mathrm{X}(g)\big] = \frac{1}{4}\int_{\mathbb{C}} \nabla f(z) \cdot \nabla g(z)\, d\mathrm{A}(z) = \int_{\mathbb{C}} \partial f(z)\overline{\partial}g(z)d\mathrm{A}(z), \qquad (1.8)$$

for any (real-valued) smooth functions f and g. Modulo constants, the RHS of formula (1.8) defines a Hilbert space, denoted $H^1(\mathbb{C})$, with norm:

$$\|f\|_{H^1(\mathbb{C})}^2 = \int_{\mathbb{C}} |\partial f(z)|^2\, d\mathrm{A}(z). \qquad (1.9)$$

Therefore the stochastic process X is called a H^1-Gaussian noise. The central limit theorem (CLT) was first established for the Ginibre process by Rider and Viràg [45] for $\mathcal{C}^1$ test functions with at most exponential growth at infinity. For general real-analytic potentials, it was proved in [3] that for any smooth function f with compact support, one has as $N \to \infty$,

$$\Xi(f) - \mathbb{E}\big[\Xi(f)\big] \Rightarrow \mathrm{X}(f^{\dagger}) \qquad (1.10)$$

where $f^{\dagger}$ is the (unique) continuous and bounded function on $\mathbb{C}$ such that $f^{\dagger} = f$ on the droplet $\overline{\mathscr{S}_V}$ and $\Delta f^{\dagger} = 0$ on $\mathbb{C}\backslash\overline{\mathscr{S}_V}$. Actually, when $\mathrm{supp}(f) \subset \mathscr{S}_V$, we have $f^{\dagger} = f$ on $\mathbb{C}$ and the CLT was obtained previously in the paper [2] from which part of our method is inspired. We also refer to [10,41] for more recent proofs which hold for general $\beta > 0$. By convention, in (1.8) and below, $\Rightarrow$ means that the convergence holds in distribution and that all moments of the random variable converge.

In order to describe the crossover from the H^1-Gaussian noise to white noise, let Λ_η be a mean-zero Poisson process with intensity $\eta \in L^\infty(\mathfrak{X})$. This process is characterized by the fact that for any function $f \in \mathcal{C}_c(\mathfrak{X})$, the Laplace transform of the random variable $\Lambda_\eta(f)$ is well-defined and given by

$$\log \mathbb{E}\big[\exp \Lambda_\eta(f)\big] = \int_{\mathfrak{X}} \big(e^{f(z)} - 1 - f(z)\big)\eta(z)d\mu(z). \qquad (1.11)$$

Theorem 1.1 *Let* X *be a* H^1-*Gaussian noise and* $\Lambda_{\tau\varrho_V}$ *be an independent Poisson process with intensity* $\tau\varrho_V$ *where* $\tau > 0$ *defined on the same probability space. Let* $f \in \mathcal{C}_c^3(\mathscr{S}_V)$, $p_N = 1 - q_N$, *and let* $T_N = Nq_N$. *As* $N \to \infty$ *and* $q_N \to 0$, *we have*

$$\widehat{\Xi}(f) - \mathbb{E}\big[\widehat{\Xi}(f)\big] \Rightarrow \mathrm{X}(f) \qquad\qquad\qquad \text{if } T_N \to 0, \qquad (1.12)$$

$$\frac{\widehat{\Xi}(f) - \mathbb{E}\big[\widehat{\Xi}(f)\big]}{\sqrt{T_N}} \Rightarrow \mathcal{N}\left(0, \int f(z)^2 \varrho_V(z) d\mathrm{A}(z)\right) \qquad \text{if } T_N \to \infty, \qquad (1.13)$$

$$\widehat{\Xi}(f) - \mathbb{E}\big[\widehat{\Xi}(f)\big] \Rightarrow \mathrm{X}(f) + \Lambda_{\tau\varrho_V}(-f) \qquad\quad \text{if } T_N \to \tau. \qquad (1.14)$$

The proof of Theorem 1.1 is based on the cumulants' method and it is explained in details in Sect. 2. In particular, we formulate a result—Theorem 2.2—valid for general determinantal point process and might be of independent interest. The details of the proof of Theorem 1.1 are given in Sect. 3.2. Our method relies on the approximations of the correlation kernel K_V^N from [2]—see Lemma 3.1 below—and it restricts us to work with test functions which

are supported inside the bulk. However, the result should be true for general functions if we replace f by $f^{\dagger}$ on the RHS of (1.12) and (1.14).

Theorem 1.1 can be interpreted as follows. In the regime $Nq_N \to 0$, virtually no particles are deleted and linear statistics behave according to random matrix theory. On the other hand, in the regime $Nq_N \to \infty$, the variance of a linear statistic diverge. So, if we renormalize the random variable $\widehat{\Xi}(f)$, we obtain a classical CLT and (1.13) shows that the limit is described by a white noise supported on $\mathscr{S}_V$ whose intensity is the equilibrium measure ϱ_V. In the critical regime, when the expected number of deleted particles equals $\tau > 0$, the limiting process is the superposition of a H^1-correlated Gaussian noise and an independent mean-zero Poisson process applied to $-f$. Finally, by using formula (1.11), it is not difficult to check that as $\tau \to \infty$, the random variable

$$\frac{\Lambda_{\tau \varrho_V}(-f)}{\sqrt{\tau}} \Rightarrow \mathcal{N}\left(0, \int f(z)^2 \varrho_V(z) d\mathrm{A}(z)\right),$$

so that the critical regime clearly interpolates between (1.12) and (1.13).

In fact, the crossover is more interesting at mesoscopic scales. Namely, the density of a log-gas is of order N and one can also investigate fluctuations at small scales by zooming inside the bulk of the process. If $L_N \nearrow \infty$, $x_0 \in \mathscr{S}_V$, and $f \in \mathcal{C}_c(\mathbb{C})$, we consider the test function

$$f_N(z) = f\left(L_N(z - x_0)\right). \tag{1.15}$$

The regime $L_N = N^{1/d}$ is called microscopic and it was shown in [2, Proposition 7.5.1] that when $d = 2$,

$$\Xi(f_N) \Rightarrow \Xi^{\infty}_{\varrho_V(x_0)}(f),$$

where the process Ξ^{∞}_{ρ} is called the ∞-Ginibre **process** with density $\rho > 0$. It is a determinantal process on $\mathbb{C}$ with correlation kernel

$$K^{\infty}_{\rho}(z, w) = \rho e^{\rho(2z\overline{w} - |z|^2 - |w|^2)/2}. \tag{1.16}$$

Based on the argument from [2], it is straightforward to verify that the incomplete process has a local limit as well:

$$\widehat{\Xi}(f_N) \Rightarrow \widehat{\Xi}^{\infty}_{\varrho_V(x_0);p}(f) \text{ as } p_N \to p \text{ and } N \to \infty. \tag{1.17}$$

For any $0 < p \leq 1$, $\widehat{\Xi}^{\infty}_{\varrho;p}$ is a (translation invariant) determinantal process on $\mathbb{C}$ with correlation kernel $pK^{\infty}_{\varrho}(z, w)$. This process is constructed by running an independent Bernoulli percolation with parameter p on the point configuration of the ∞-Ginibre process with density $\rho > 0$. In particular, (1.17) shows that one needs to delete a non-vanishing fraction of the N particles of the gas in order to get a local limit which is different from random matrix theory. It was proved in [44] that, as the density $\rho \to \infty$, the fluctuations of the ∞-Ginibre process are of order 1 and described by the H^1-Gaussian noise:

$$\Xi^{\infty}_{\rho}(f) - \rho \int_{\mathbb{C}} f(z) d\mathrm{A}(z) \Rightarrow \mathrm{X}(f)$$

for any $f \in H^1 \cap L^1(\mathbb{C})$. Therefore it is expected that, in the mesoscopic regime, i.e. $L_N = o(\sqrt{N})$, the asymptotic fluctuations of the linear statistic $\Xi(f_N)$ are universal and described by $\mathrm{X}(f)$. However, to our best knowledge, a proof was missing from the literature and, in Sect. 3, we show that this fact follows quite simply by combining the ideas from [44] and [2].

Theorem 1.2 *Let $x_0 \in \mathscr{S}_V$, $f \in \mathcal{C}_c^3(\mathbb{C})$, $\alpha \in (0, 1/2)$, and let f_N be given by formula (1.15) with $L_N = N^\alpha$. Then, we have as $N \to \infty$,*

$$\Xi(f_N) - \mathbb{E}\big[\Xi(f_N)\big] \Rightarrow X(f).$$

Using the same method, we can also describe the fluctuations of smooth mesoscopic linear statistics of a incomplete Coulomb gas.

Theorem 1.3 *Let X be a H^1-Gaussian noise and Λ_τ be an independent Poisson process with constant intensity $\tau > 0$ on $\mathbb{C}$. Let $x_0 \in \mathscr{S}_V$, $f \in \mathcal{C}_c^3(\mathscr{S}_V)$, $\alpha \in (0, 1/2)$, and let f_N be the mesoscopic test function given by formula (1.15) with $L_N = N^\alpha$. We also let $p_N = 1 - q_N$ and $T_N = Nq_N L_N^{-2} \varrho_V(x_0)$. We have as $N \to \infty$ and $q_N \to 0$,*

$$\widehat{\Xi}(f_N) - \mathbb{E}\big[\widehat{\Xi}(f_N)\big] \Rightarrow X(f) \qquad\qquad\qquad \text{if } T_N \to 0,$$

$$\frac{\widehat{\Xi}(f_N) - \mathbb{E}\big[\widehat{\Xi}(f_N)\big]}{\sqrt{T_N}} \Rightarrow \mathcal{N}\left(0, \int f(z)^2 dA(z)\right) \qquad \text{if } T_N \to \infty,$$

$$\widehat{\Xi}(f_N) - \mathbb{E}\big[\widehat{\Xi}(f_N)\big] \Rightarrow X(f) + \Lambda_\tau(-f) \qquad\quad \text{if } T_N \to \tau.$$

The proof of Theorem 1.3 follows the same strategy are that of Theorem 1.1 and the technical differences are explained in Sect. 3.2. This result shows that, at mesoscopic scales, the transition occurs when the mesoscopic density of deleted particles which is given by the parameter $T_N > 0$ converges to a positive constant τ. In contrast to previous results, this transition appears to be non-Gaussian and it is somewhat surprising that it can also be described in an elementary way. In dimension 1, one can obtain a crossover from GUE eigenvalues to a Poisson process by letting independent points evolved according to Dyson's Brownian motion. This leads to a determinantal process sometimes called the *deformed GUE* whose kernel depends on the diffusion time, see [32]. This model also exhibit a transition which has been analyzed for mesoscopic linear statistics in [25] and it was proved that the critical fluctuations are Gaussian. One can also consider non-intersecting Brownian motions on a cylinder. It turns out that this point process describes the positions of free fermions confined in a harmonic trap at a temperature $\tau > 0$. It was established in [34], see also [21], that the corresponding grand canonical ensemble is determinantal with a correlation kernel of the form (1.6) with for $k \geq 0$,

$$\mathbb{E}[J_k^N] = \frac{1}{1 + \exp\left(\frac{k-N}{\tau}\right)}.$$

For sufficiently small temperature, this system behaves like its ground-state, the GUE, while it behaves like a Poisson process at larger temperature. It was proved in [35] that this leads to yet another crossover where non-Gaussian fluctuations are observed at the critical temperature. However, to the author's knowledge, in contrast to the incomplete ensembles considered here, the critical processes discovered in [35] cannot be described in simple terms like Theorem 1.5 below.

1.3 Main Results for Eigenvalues of Unitary Invariant Hermitian Random Matrices

For 1-dimensional log-gases, for general $\beta > 0$ and for a large class of potentials, Johansson established in [31] the existence of the equilibrium measure and also managed to describe the fluctuations around the equilibrium configuration. To state the result in a universal way, note that one can make an affine rescaling of the potential and assume that $\mathscr{I}_V \subset [-1, 1]$. If

V is a polynomial and $\mathscr{I}_V = (-1, 1)$, Johansson proved that linear statistics of the process Ξ satisfy a central limit theorem:

$$\Xi(f) - N \int_{\mathbb{R}} f(x) \varrho_V(x) dx \Rightarrow \mathrm{Y}(f) \text{ as } N \to \infty, \tag{1.18}$$

for any $f \in \mathcal{C}^2(\mathbb{R})$ such that $f'(x)$ grows at most polynomially as $|x| \to \infty$. The process Y is a centered Gaussian noise defined on $[-1, 1]$ with covariance structure:

$$\mathbb{E}\big[\mathrm{Y}(f)\mathrm{Y}(g)\big] = \frac{1}{4} \sum_{k=1}^{\infty} k \mathrm{c}_k(f) \mathrm{c}_k(g). \tag{1.19}$$

In (1.19), $\mathrm{c}_k(f)$ denote the Fourier–Chebyshev coefficients of the function f:

$$\mathrm{c}_k(f) = \frac{2}{\pi} \int_{-1}^{1} f(x) T_k(x) \frac{dx}{\sqrt{1 - x^2}}, \tag{1.20}$$

where $(T_k)_{k=0}^{\infty}$ are the Chebyshev polynomials of the first kind.[2] The CLT (1.18) holds for more general potentials and for other orthogonal polynomial ensembles as well, see [43, Section 11.3] or [12,17,39] and it is known that *the one-cut condition*, i.e. the assumption that the support of the equilibrium measure is connected, is necessary. Otherwise, the asymptotic fluctuations of a generic linear statistic $\Xi(f)$ are still of order 1 but are not Gaussian, see [13,42,46]. In fact, the one-cut condition is closely related to the fact the recurrence coefficients [see formula (4.1)] which defines the orthogonal polynomials $(P_k)_{k \geq 0}$ appearing in the correlation kernel (1.3) satisfy for any $j \in \mathbb{Z}$,

$$\lim_{N \to \infty} a_{N+j}^N = 1/2 \text{ and } \lim_{N \to \infty} b_{N+j}^N = 0; \tag{1.21}$$

see Remark 4.2 below. Like for 2-dimensional Coulomb gas, we obtain analogous transitions for the eigenvalues of random unitary invariant Hermitian matrices.

Theorem 1.4 *Let $p_N = 1 - q_N$, $T_N = N q_N$, and suppose that the recurrence coefficients of the orthogonal polynomials $\{P_k\}_{k=0}^{\infty}$ satisfy the conditions (1.21). Then, for any polynomial Q, we obtain as $N \to \infty$ and $q_N \to 0$,*

$$\widehat{\Xi}(Q) - \mathbb{E}\big[\widehat{\Xi}(Q)\big] \Rightarrow \mathrm{Y}(Q) \qquad\qquad \text{if } T_N \to 0,$$

$$\frac{\widehat{\Xi}(Q) - \mathbb{E}\big[\widehat{\Xi}(Q)\big]}{\sqrt{T_N}} \Rightarrow \mathcal{N}\left(0, \int_{\mathbb{R}} Q(x)^2 \varrho_V(x) dx\right) \qquad \text{if } T_N \to \infty,$$

$$\widehat{\Xi}(Q) - \mathbb{E}\big[\widehat{\Xi}(Q)\big] \Rightarrow \mathrm{Y}(Q) + \Lambda_{\tau \varrho_V}(-Q) \qquad \text{if } T_N \to \tau,$$

where the Poisson process $\Lambda_{\tau \varrho_V}$ is independent from the Gaussian process Y and both are defined on $\mathscr{I}_V = (-1, 1)$.

The proof of Theorem 1.4 is also based on the cumulants' method and on Theorem 2.2. However, the technical details, which are explained in Sects. 4.1 and 4.2, rely on the formulation from [39] and are very different from that of the proof of Theorem 1.1.

We also obtain the counterpart of Theorem 1.4 for mesoscopic linear statistics. For any function $f \in L^1(\mathbb{R})$, we define its Fourier transform:

$$\hat{f}(u) = \int_{\mathbb{R}} f(x) e^{-2\pi i x u} dx.$$

---

[2] $T_k(\cos\theta) = \cos(k\theta)$ for any $k \geq 0$ and $\theta \in \mathbb{R}$.

We let Z be a mean-zero Gaussian process on $\mathbb{R}$ with correlation structure:

$$\mathbb{E}\big[Z(h)Z(g)\big] = \int_0^\infty u\hat{h}(u)\overline{\hat{g}(u)}du. \tag{1.22}$$

Since $\mathrm{Var}\big[Z(f)\big] = \|f\|_{H^{1/2}(\mathbb{R})}$, the process Z is usually called the $H^{1/2}$—Gaussian noise. It describes the mesoscopic fluctuations of the eigenvalues of Hermitian random matrices, see [16,29,38], as well as the mesoscopic fluctuations of the log-gases for general $\beta > 0$, [5], and of certain random band matrices in the appropriate regime [27,28].

Theorem 1.5 *We let $x_0 \in \mathscr{I}_V$, $f \in \mathcal{C}_c^2(\mathbb{R})$, $\alpha \in (0,1)$, and $f_N(x) = f\big(N^\alpha(x - x_0)\big)$. We also let $p_N = 1 - q_N$ and $T_N = q_N N^{1-\alpha}\varrho_V(x_0)$. We obtain as $N \to \infty$ and $q_N \to 0$,*

$$\widehat{\Xi}(f_N) - \mathbb{E}\big[\widehat{\Xi}(f_N)\big] \Rightarrow Z(f) \qquad\qquad \text{if } T_N \to 0, \tag{1.23}$$

$$\frac{\widehat{\Xi}(f_N) - \mathbb{E}\big[\widehat{\Xi}(f_N)\big]}{\sqrt{T_N}} \Rightarrow \mathcal{N}\left(0, \int_\mathbb{R} f(x)^2 dx\right) \qquad \text{if } T_N \to \infty, \tag{1.24}$$

$$\widehat{\Xi}(f_N) - \mathbb{E}\big[\widehat{\Xi}(f_N)\big] \Rightarrow Z(f) + \Lambda_\tau(-f) \qquad \text{if } T_N \to \tau, \tag{1.25}$$

where the Poisson process Λ_τ has constant intensity $\tau > 0$ on $\mathbb{R}$ and is independent from the $H^{1/2}$—Gaussian noise Z.

The proof of Theorem 1.5 is quite similar to that of Theorem 1.3. It follows the strategy explained in Sect. 2 and is based on the asymptotics for the correlation kernel K_V^N in terms of the sine-kernel, see [37]. The details relies on the method from [38] and are given in Sect. 4.3.

1.4 Overview of the Rest of the Paper

In Sect. 2, we present the strategy of the proofs of the results from Sects. 1.2 and 1.3. We begin by reviewing Soshnikov's cumulants' method. Then, we explain how to apply it to the incomplete ensemble $\widehat{\Xi}$ and we obtain a general result—Theorem 2.2—which characterizes the transition from Gaussian to Poisson statistics for general determinantal point processes. The rest of the paper consists in verifying the assumptions of Theorem 2.2 for determinantal log-gases in dimensions 1 and 2. In Sect. 3, we prove Theorems 1.1–1.3 for the 2-dimensional log-gases. The proof relies on estimates for the correlation kernel (1.3) which come from the paper [2] and are collected in the Appendix B. In Sect. 4, we provide the details of the proofs of Theorems 1.4 and 1.5 by relying on the method from [39] and [38] respectively.

In the following, $C > 0$ denotes a numerical constant which changes from line to line. For any $n \in \mathbb{N}$, we let for all $\mathrm{x} \in \mathfrak{X}^n$,

$$d\mu^n(\mathrm{x}) = d\mu(x_1)\cdots d\mu(x_n).$$

If (u_N) and (v_N) are two sequences, we use the notation:

$$u_N \simeq v_N \text{ if } \lim_{N\to\infty}(u_N - v_N) = 0.$$

2 Outline of the Proof

In this section, we consider a general state space $\mathfrak{X}$ which is a complete separable metric space equipped with a Radon measure μ as in [30,47] and let Ξ be a sequence of determinantal point processes on $\mathfrak{X}$ with correlation kernels K^N which are reproducing:

$$\int_{\mathfrak{X}} K^N(z, x) K^N(x, w) d\mu(x) = K^N(z, w). \tag{2.1}$$

One may think of the parameter $N \in \mathbb{N}$ as the *density* of particles. Since it is generally the case in the context of random matrix theory, we shall also assume that the kernels K^N are continuous on $\mathfrak{X} \times \mathfrak{X}$, Hermitian symmetric, and that they define locally trace-class integral operators acting on $L^2(\mathfrak{X}, \mu)$.

The cumulants' method to analyze the asymptotic distribution of linear statistic of determinantal processes goes back to the work of Costin and Lebowitz [20] for count statistics of the Sine process. The general theory was developed by Soshnikov in [47–49] and subsequently applied to many different ensembles coming from random matrix theory, see for instance [2,16,17,35,38,39,44,45].

In this section, we show how to implement the cumulants' method to describe the asymptotics law of linear statistics of the incomplete ensemble $\widehat{\Xi}$ with correlation kernel $p_N K^N(z, w)$ when the density of particles $0 < p_N < 1$ converges to 1 in the large N limit.

Let

$$\mho = \bigcup_{l=1}^{\infty} \left\{ \mathbf{k} = (k_1, \ldots, k_l) \in \mathbb{N}^l \right\}$$

and let $\ell(\mathbf{k}) = l$ denote the length of the tuple $\mathbf{k}$. Let us denote the set of compositions of the integer $n > 0$ by

$$\{\mathbf{k} \vdash n\} = \left\{ \mathbf{k} \in \mho : k_1 + \cdots + k_l = n \right\}.$$

We also denote by $n \in \mho$ the trivial composition. For any map $\Upsilon : \mho \mapsto \mathbb{R}$, for any function $f : \mathfrak{X} \to \mathbb{R}$, and for any $n \in \mathbb{N}$, we define for all $x \in \mathfrak{X}^n$,

$$\Upsilon^n[f](x) = \sum_{\mathbf{k} \vdash n} \Upsilon(\mathbf{k}) \prod_{1 \le j \le \ell(\mathbf{k})} f(x_j)^{k_j}. \tag{2.2}$$

If $\mathbf{k} \vdash n$, we let $M(\mathbf{k}) = \dfrac{n!}{k_1! \cdots k_l!}$ be the corresponding multinomial coefficient and for all integers $n \ge 1$ and $m \in \{0, \ldots, n\}$, we define the coefficients

$$\gamma_m^n = \sum_{\mathbf{k} \vdash n} \frac{(-1)^{\ell(\mathbf{k})}}{\ell(\mathbf{k})} \binom{\ell(\mathbf{k})}{m} M(\mathbf{k}). \tag{2.3}$$

We will also use the notation: $\delta_k(n) = \begin{cases} 1 & \text{if } n = k \\ 0 & \text{else} \end{cases}$ for any $k \in \mathbb{Z}$.

Lemma 2.1 *For all $n \in \mathbb{N}$, we have $\gamma_0^n = \delta_1(n)$ and $\gamma_1^n = (-1)^n$.*

Proof The coefficients (2.3) have the generating function:

$$\sum_{n=1}^{\infty} \sum_{m=0}^{n} \gamma_m^n \frac{x^n q^m}{n!} = -\log\left(1 + (1 + q)(e^x - 1)\right).$$

In particular, setting $q = 0$, we see that $\gamma_0^n = \delta_1(n)$. Moreover, since

$$\sum_{n=1}^{\infty} \gamma_1^n \frac{x^n}{n!} = \frac{d}{dq} \log\left(1 + (1 + q)(e^x - 1)\right)\bigg|_{q=0} = 1 - e^{-x},$$

we also see that $\gamma_1^n = (-1)^n$. $\qquad \square$

Given a test function $f : \mathcal{X} \to \mathbb{R}$, say locally integrable with compact support, the cumulant generating function of the random variable $\Xi(f)$ is

$$\log \mathbb{E}\big[\exp(\lambda \Xi(f)) \big] = \sum_{n=1}^{\infty} \frac{\lambda^n}{n!} \mathrm{C}_{K^N}^n[f].$$

It was proved by Soshnikov that under our general assumptions, the cumulants $\mathrm{C}_K^n[f]$ characterize the law of the linear statistics $\Xi(f)$ and that for any $n \in \mathbb{N}$,

$$\mathrm{C}_{K^N}^n[f] = - \sum_{l=1}^{n} \frac{(-1)^l}{l} \sum_{\substack{\mathbf{k} \vdash n \\ \ell(\mathbf{k})=l}} \mathrm{M}(\mathbf{k}) \int_{\substack{\mathcal{X}^l \\ x_0=x_l}} f(x_1)^{k_1} \cdots f(x_l)^{k_l} \prod_{1 \le j \le l} K^N(x_j, x_{j-1}) d\mu^l(x).$$

(2.4)

Under stronger assumptions, for instance if the kernel K^N has finite rank, this formula makes sense also for test functions which are not necessarily compactly supported. We use the convention that the variables x_0 and x_l are identified in the previous integral. Since we assume that the correlation kernel K^N is reproducing, we can rewrite this formula:

$$\mathrm{C}_{K^N}^n[f] = - \sum_{\mathbf{k} \vdash n} \frac{(-1)^{\ell(\mathbf{k})}}{\ell(\mathbf{k})} \mathrm{M}(\mathbf{k}) \int_{\substack{\mathcal{X}^n \\ x_0=x_n}} f(x_1)^{k_1} \cdots f(x_l)^{k_l} \prod_{1 \le j \le n} K^N(x_j, x_{j-1}) d\mu^n(x)$$

$$= - \int_{\substack{\mathcal{X}^n \\ x_0=x_n}} \Upsilon_0^n[f](x) \prod_{1 \le j \le n} K^N(x_j, x_{j-1}) d\mu^n(x),$$

(2.5)

where for any $\mathbf{k} \in \mho$,

$$\Upsilon_0(\mathbf{k}) = \frac{(-1)^{\ell(\mathbf{k})}}{\ell(\mathbf{k})} \mathrm{M}(\mathbf{k}).$$

(2.6)

A simple observation which turns out to be very important when it comes to asymptotics is that, by Lemma 2.1, for all $n \ge 2$,

$$\sum_{\mathbf{k} \vdash n} \Upsilon_0(\mathbf{k}) = \gamma_0^n = 0.$$

(2.7)

For any $m \in \mathbb{N}$, we define[3]

$$\begin{cases} \Upsilon_m(\mathbf{k}) = \dfrac{(-1)^{\ell(\mathbf{k})}}{\ell(\mathbf{k})} \dbinom{\ell(\mathbf{k})}{m} \mathrm{M}(\mathbf{k}) & \text{if } \ell(\mathbf{k}) \ge 2 \\ \Upsilon_m(\mathbf{k}) = -\delta_1(m) - \gamma_m^n & \text{if } \mathbf{k} = n \end{cases}.$$

(2.8)

These functions are constructed so that we have for all $n, m \in \mathbb{N}$,

$$\sum_{\mathbf{k} \vdash n} \Upsilon_m(\mathbf{k}) = 0.$$

(2.9)

According to formula (2.4), the cumulants of the process $\widehat{\Xi}$ with correlation kernel $\widehat{K}_N = p_N K^N$ are given by

$$\mathrm{C}_{\widehat{K}^N}^n[f] = - \sum_{l=1}^{n} \frac{(-1)^l}{l} p_N^l \sum_{\substack{\mathbf{k} \vdash n \\ \ell(\mathbf{k})=l}} \mathrm{M}(\mathbf{k}) \int_{\substack{\mathcal{X}^l \\ x_0=x_l}} f(x_1)^{k_1} \cdots f(x_l)^{k_l} \prod_{1 \le j \le l} K^N(x_j, x_{j-1}) d\mu^l(x).$$

---

[3] By convention $\binom{l}{m} = 0$ if $m > l$.

Since the kernel K^N is reproducing, if we set $q_N = 1 - p_N$, using the binomial formula, we obtain

$$
\begin{aligned}
C^n_{\widehat{K}^N}[f] = C^n_{K^N}[f] &- \sum_{m=1}^n (-q_N)^m \gamma^n_m \int_{\mathfrak{X}} f(x)^n K^N(x, x) d\mu(x) \\
&- \sum_{m=1}^n (-q_N)^m \int_{\substack{\mathfrak{X}^n \\ x_0 = x_n}} \Upsilon^n_m[f](x) \prod_{1 \le j \le n} K^N(x_j, x_{j-1}) d\mu^n(x).
\end{aligned}
\tag{2.10}
$$

We are now ready to state our general result from which Theorems 1.1, 1.3, 1.4 and 1.5 in the introduction follow.

Theorem 2.2 *Let $0 < q_N < 1$ be a sequence which converges to 0 as $N \to +\infty$. Under our general assumptions above, let f_N be a sequence of functions for which the cumulants $C^n_{K^N}[f_N]$ are well-defined for all n, $N \in \mathbb{N}$ and the following conditions hold:*

1. *There exists a (Radon) measure η on $\mathfrak{X}$, a function $f \in L^p(\eta)$ for any $p \ge 2$, and a sequence $M_N \nearrow +\infty$ as $N \to +\infty$ such that for all $n \ge 1$,*

$$
\frac{1}{M_N} \int_{\mathfrak{X}} f_N(x)^n K^N(x, x) d\mu(x) \simeq \int_{\mathfrak{X}} f(x)^n d\eta(x).
\tag{2.11}
$$

2. *For all $n \ge 2$ and all $m \ge 1$, as $N \to +\infty$*

$$
\left| \int_{\substack{\mathfrak{X}^n \\ x_0 = x_n}} \Upsilon^n_m[f_N](x) \prod_{1 \le j \le n} K^N(x_j, x_{j-1}) d\mu^n(x) \right| = o\left(q_N^{-1} \vee M_N\right).
\tag{2.12}
$$

3. *There exists $\sigma > 0$ such that for all $n \in \mathbb{N}$,*

$$
\lim_{N \to \infty} C^n_{K^N}[f_N] = \begin{cases} \sigma^2 & \text{if } n = 2 \\ 0 & \text{if } n > 2 \end{cases}.
\tag{2.13}
$$

Then, depending on the parameter $T_N = q_N M_N > 0$, we distinguish three different asymptotic regimes for the linear statistic $\widehat{\Xi}(f_N)$ of the thinned point process with density $p_N = 1 - q_N$:

(i) *If $T_N \to 0$ as $N \to +\infty$,*

$$
\widehat{\Xi}(f_N) - \mathbb{E}\big[\widehat{\Xi}(f_N)\big] \Rightarrow \mathcal{N}(0, \sigma^2).
$$

(ii) *If $T_N \to \tau$ with $\tau > 0$ as $N \to +\infty$,*

$$
\widehat{\Xi}(f_N) - \mathbb{E}\big[\widehat{\Xi}(f_N)\big] \Rightarrow X + \Lambda_{\tau\eta}(f),
$$

where $X \sim \mathcal{N}(0, \sigma^2)$ and $\Lambda_{\tau\eta}$ is Poisson process on $\mathfrak{X}$ with intensity $\tau\eta$ independent from X.

(iii) *If $T_N \to +\infty$ as $N \to +\infty$,*

$$
\frac{\widehat{\Xi}(f_N) - \mathbb{E}\big[\widehat{\Xi}(f_N)\big]}{\sqrt{T_N}} \Rightarrow \mathcal{N}\left(0, \int f(x)^2 d\eta(x)\right).
$$

Before giving our proof of Theorem 2.2, let us make two remarks about the Assumptions (2.11)–(2.13).

Remark 2.1 For any function $g : \mathfrak{X} \to \mathbb{R}_+$, we have $\mathbb{E}[\Xi(g)] = \int_{\mathfrak{X}} g(x) K^N(x, x) d\mu(x)$, so that one can interpret (2.11) as a condition about the mean of the point process Ξ. In contrast, (2.12) can be seen as a condition about the fluctuations of the incomplete point process. We also implicitly assume that for $n = 2$, the RHS of (2.11) is positive, so that the measure η is non-trivial and puts mass on the support of f. Then the random variables $\mathcal{N}(0, \int f(x)^2 d\eta(x))$ and $\Lambda_{\tau\eta}(f)$ are non zero for any $\tau > 0$. In order to handle mesoscopic linear statistics, we have allowed our test functions to depend on the parameter N. However, for simplicity, one can think of the case where $f_N = f$ is a smooth and compactly supported test function.

Remark 2.2 Instead of (2.13), we could assume that the cumulants of the linear statistics $\Xi(f_N)$ converge to that of a random variable X which is not necessarily Gaussian. Then, the conclusion (ii) of Theorem 2.2 remains true and we obtain a crossover from a non-Gaussian process to a Poisson process. For instance, this more general situation arises when considering linear statistics of 1-dimensional log-gases in the multi-cut regime.

Proof Observe that it follows from formula (2.10) and the condition (2.12) that the cumulants of the linear statistic $\widehat{\Xi}(f_N)$ satisfy for all $n \geq 2$ as $N \to +\infty$,

$$C_{\widehat{K}^N}^n[f_N] = C_{K^N}^n[f_N] - \sum_{m=1}^{n} (-q_N)^m \gamma_m^n \int_{\mathfrak{X}} f_N(x)^n K^N(x, x) d\mu(x) + o(1 \vee T_N),$$

where $T_N = q_N L_N$. Then, using the condition (2.11), we obtain for any $n \geq 2$, as $N \to +\infty$,

$$C_{\widehat{K}^N}^n[f_N] = C_{K^N}^n[f_N] + T_N \sum_{m=0}^{n-1} (-q_N)^m \gamma_{m+1}^n \left(\int_{\mathfrak{X}} f_N(x)^n d\eta(x) + o(1) \right) + o(1 \vee T_N).$$
(2.14)

Let us observe that in the previous sum, regardless of the regime we consider, since $q_N \to 0$ as $N \to +\infty$, only the term $m = 0$ is asymptotically relevant. For instance, if we assume that $T_N = \tau + o(1)$ with $\tau \geq 0$, this implies that

$$C_{\widehat{K}^N}^n[f_N] = C_{K^N}^n[f_N] + \tau \gamma_1^n \int_{\mathfrak{X}} f_N(x)^n d\eta(x) + o(1).$$

On the one hand by Lemma 2.1, since $\gamma_1^n = (-1)^n$ and using the condition (2.13), we obtain for any $n \geq 2$

$$\lim_{N \to +\infty} C_{\widehat{K}^N}^n[f_N] = \sigma^2 \mathbf{1}_{n=2} + (-1)^n \tau \int_{\mathfrak{X}} f_N(x)^n d\eta(x).$$
(2.15)

In the regime (i)—which corresponds to $\tau = 0$—this shows that the linear statistic $\widehat{\Xi}(f_N)$, once centered, converges in distribution (as well as in the sense of moments) to a Gaussian random variable with variance σ^2. Let us observe that by formula (1.11), if $\tau > 0$, the second term on the RHS of (2.15) corresponds to the nth cumulant of the random variable $\Lambda_{\tau\eta}(f)$. This proves the claim in the regime (ii).

On the other hand, in the regime (iii) where $T_N \to +\infty$, we see from (2.14) that the variance $C_{\widehat{K}^N}^n[f_N]$ diverges as $N \to +\infty$. Thus, in order to have a non-trivial limit, we need to renormalize the linear statistic $\widehat{\Xi}(f_N)$. Namely, we consider instead the test function $g_N = f_N/\sqrt{T_N}$ and it follows from (2.14) that for all $n \geq 2$, as $N \to +\infty$

$$C_{\widehat{K}^N}^n[g_N] = \mathbf{1}_{n=2} \left(\int_{\mathfrak{X}} f(x)^2 d\eta(x) + o(q_N) \right) + o(1).$$

These asymptotics show that in the the regime (iii), $\frac{\widehat{\Xi}(f_N) - \mathbb{E}\left[\widehat{\Xi}(f_N)\right]}{\sqrt{T_N}}$ converges in distribution (as well as in the sense of moments) to a centered Gaussian random variable with variance $\int_{\mathfrak{X}} f(x)^2 d\eta(x)$. □

3 Transition for Coulomb Gases in Two Dimensions

3.1 Asymptotics of the Correlation Kernel

In this section, we begin by reviewing the basics of the theory of eigenvalues of random normal matrices developed by Ameur, Hedenmalm and Makarov. In particular, we are interested in the properties of the correlation kernel (1.3) in the bulk of the gas. It has been established in [2] that, if the potential V is real-analytic, the equilibrium measure is $\varrho_V = 2\Delta V \mathbf{1}_{\mathscr{S}_V}$ and the droplet $\overline{\mathscr{S}_V}$ is a compact set with a nice boundary. Moreover, in order to compute the asymptotics of the cumulants of a smooth linear statistic, instead of working with the correlation kernel K_V^N, one can use the so-called approximate Bergman kernel:

$$B_V^N(z, w) = \left(N b_0(z, \overline{w}) + b_1(z, \overline{w})\right) e^{N\{2\Phi(z,\overline{w}) - V(z) - V(w)\}}. \tag{3.1}$$

The functions $b_0(z, w)$, $b_1(z, w)$ and $\Phi(z, w)$ are the (unique) bi-holomorphic functions defined in a neighborhood in $\mathbb{C}^2$ of the set $\left\{(z, \overline{z}) : z \in \mathscr{S}_V\right\}$ such that $b_0(z, \overline{z}) = 2\Delta V(z)$, $b_1(z, \overline{z}) = \frac{1}{2}\Delta \log(\Delta V)(z)$, and $\Phi(z, \overline{z}) = V(z)$.

Lemma 3.1 (Lemma 1.2 in [2], proved in [1,7]) *For any $x_0 \in \mathscr{S}_V$, there exists $\epsilon_0 > 0$ and $C_0 > 0$ so that when the dimension N is sufficiently large, we have for all $z, w \in \mathbb{D}(x_0, \epsilon_0)$,*

$$\left|K_V^N(z, w) - B_V^N(z, w)\right| \le C_0 N^{-1}. \tag{3.2}$$

Moreover, at sufficiently small mesoscopic scale, up to a gauge transform, the asymptotics of the approximate Bergman kernel B_V^N is universal.

Lemma 3.2 *Let $\kappa > 0$ and $\epsilon_N = \kappa N^{-1/2} \log N$ for all $N \in \mathbb{N}$. For any $x_0 \in \mathscr{S}_V$, there exists $\varepsilon_0 > 0$ and a function $\mathfrak{h} : \mathbb{D}(0, \varepsilon_0) \to \mathbb{R}$ such that if the parameter N is sufficiently large, the function*

$$\widetilde{B}_{V,x_0}^N(u, v) = \frac{B_V^N(x_0 + u, x_0 + v) e^{iN\mathfrak{h}(u)}}{e^{iN\mathfrak{h}(v)}} \tag{3.3}$$

satisfies

$$\widetilde{B}_{V,x_0}^N(u, v) = K_{N\varrho_V(x_0)}^\infty(u, v) \left\{1 + \underset{N\to\infty}{O}\left((\log N)^2 \epsilon_N\right)\right\} \tag{3.4}$$

uniformly for all $u, v \in \mathbb{D}(0, \epsilon_N)$, where $K_{N\varrho_V(x_0)}^\infty$ is the ∞-Ginibre kernel with density $N\varrho_V(x_0) = 2N\Delta V(x_0)$.

A key ingredient in the paper [2], as well as [44], is to reduce the domain of integration in formula (2.5), using the exponential off-diagonal decay of the correlation kernels K_V^N, to a set where we can use the asymptotics (3.4)—see the next Lemma. For completeness, the proofs of Lemmas 3.2 and 3.3 are given in the Appendix B.

Lemma 3.3 *Let $n \in \mathbb{N}$ and $\epsilon_N = \kappa N^{-1/2} \log N$ for some constant $\kappa > 0$ which is sufficiently large compared to n. We let*

$$\mathscr{A}(z_0; \epsilon) = \left\{ z \in \mathbb{C}^n : |z_j - z_{j-1}| \leq \epsilon \text{ for all } j = 1, \ldots n \right\}. \tag{3.5}$$

Let $\mathscr{S}$ be a compact subset of $\mathscr{S}_V$, $N_0 \in \mathbb{N}$, and $F_N : \mathbb{C}^{n+1} \to \mathbb{R}$ be a sequence of continuous functions such that

$$\sup \left\{ |F_N(z_0, z)| : z \in \mathbb{C}^n, N \geq N_0 \right\} \leq C \mathbf{1}_{z_0 \in \mathscr{S}}. \tag{3.6}$$

We have

$$\int_{\substack{\mathbb{C}^{n+1} \\ z_0 = z_{n+1}}} F_N(z_0, z) \prod_{j=0}^n K_V^N(z_j, z_{j+1}) dA(z_0) dA^n(z)$$

$$= \int_{\mathscr{S}} dA(z_0) \int_{\substack{\mathscr{A}(z_0; \epsilon_N) \\ z_{n+1} = z_0}} F_N(z_0, z) \prod_{j=0}^n K_V^N(z_j, z_{j+1}) dA^n(z) + \underset{N \to \infty}{O}(N^{-1}). \tag{3.7}$$

Remark 3.1 Recall that K_ρ^∞, (1.16), denotes the correlation kernel of the ∞-Ginibre process with density $\rho > 0$. Using the fact that for all $z, w \in \mathbb{C}$,

$$\left| K_\rho^\infty(z, w) \right| = \rho e^{-\rho |z - w|^2 / 2}, \tag{3.8}$$

it is easy to obtain the counterpart of Lemma 3.3 for the ∞-Ginibre kernel, see Lemma 4.3.

3.2 Proof of Theorem 1.1

In this section, we show how to apply Theorem 2.2 for Coulomb gases in the global regime by relying on the asymptotics from Sect. 3.1 for the correlation kernel K_V^N. First, observe that with $f_N = f$ and $M_N = N$, the Assumptions (2.11) follow immediately from the law of large numbers (1.5). Then the measure $d\eta = \varrho_V d\mu$ is absolutely with respect to μ with compact support. Moreover, if $f \in \mathcal{C}_c^3(\mathscr{S}_V)$, then the conditions (2.13) are well-known from [2, Theorem 4.4] with $\sigma^2 = \mathbb{E}[X(f)]$ according to formula (1.8). So, our main technical challenge is to obtain the estimates (2.12) for a large class of test functions. The first step of the proof is the following approximation.

Proposition 3.4 *Let $\mathcal{K} \subset \mathbb{C}$ be a compact set, $f \in \mathcal{C}_c^3(\mathcal{K})$ and let $\Upsilon : \mathfrak{V} \to \mathbb{R}$ be any map such that $\sum_{\mathbf{k} \vdash n} \Upsilon(\mathbf{k}) = 0$ for all $n \geq 2$. Let L_N be an increasing sequence such that $L_N^{-1} (\log N)^4 = o(1)$ and $L_N = o(\sqrt{N}/(\log N)^3)$ as $N \to +\infty$. We also denote by $H^n(\lambda; w)$ the second order Taylor polynomial at 0 of the function $w \in \mathbb{C}^n \mapsto \Upsilon^{n+1}[f](\lambda, \lambda + w)$. Fix $x_0 \in \mathscr{S}_V$ and let $f_N(z) = f(L_N(z - x_0))$ as in (1.15). Then, we have for any $n \geq 1$, as $N \to +\infty$,*

$$\int_{\substack{\mathbb{C}^{n+1} \\ z_0 = z_{n+1}}} \Upsilon^{n+1}[f_N](z_0, z) \prod_{j=0}^n K_V^N(z_j, z_{j+1}) dA(z_0) dA^n(z)$$

$$\simeq \int_{\mathcal{K}} dA(\lambda) \int_{\substack{\mathbb{C}^n \\ w_0 = w_{n+1} = 0}} H^n(\lambda; w) \prod_{j=0}^n K_{\eta_N(\lambda)}^\infty(w_j, w_{j+1}) dA^n(w), \tag{3.9}$$

where the density is given by $\eta_N(\lambda) = N L_N^{-2} \varrho_V(x_0 + \lambda/L_N)$.

Remark 3.2 Observe that under the assumptions of Proposition 3.4, we have $\eta_N(\lambda) \to +\infty$ as $N \to +\infty$ uniformly for all $\lambda \in \mathcal{K}$. Moreover, it follows from the proof below that in the global regime where $L_N = 1$ and $x_0 = 0$, provided that $\mathcal{K} \subset \mathscr{S}_V$, the estimates (3.9) remain valid with an extra error. Namely, we obtain for all $n \geq 1$,

$$
\int_{\substack{\mathbb{C}^{n+1} \\ z_0 = z_{n+1}}} \Upsilon^{n+1}[f_N](z_0, \mathbf{z}) \prod_{j=0}^{n} K_V^N(z_j, z_{j+1}) d\mathrm{A}(z_0) d\mathrm{A}^n(\mathbf{z})
$$

$$
= \int_{\mathcal{K}} d\mathrm{A}(\lambda) \int_{\substack{\mathbb{C}^n \\ w_0 = w_{n+1} = 0}} H^n(\lambda; \mathbf{w}) \prod_{j=0}^{n} K_{N\varrho V(\lambda)}^{\infty}(w_j, w_{j+1}) \, d\mathrm{A}^n(\mathbf{w}) \tag{3.10}
$$

$$
+ \underset{N \to \infty}{O}\left((\log N)^4\right).
$$

Proof We let $F_N = \Upsilon^{n+1}[f_N]$ and

$$
\mathscr{I}_N^n := \int_{\substack{\mathbb{C}^{n+1} \\ z_{n+1} = z_0}} F_N(z_0, \mathbf{z}) \prod_{j=0}^{n} K_V^N(z_j, z_{j+1}) d\mathrm{A}(z_0) d\mathrm{A}^n(\mathbf{z}).
$$

Since $x_0 \in \mathscr{S}_V$, there exists a compact set $\mathscr{S} \subset \mathscr{S}_V$ so that $\mathrm{supp}(f_N) \subseteq \mathscr{S}$ when the parameter N is sufficiently large. Then, according to formula (2.2), the function F_N satisfies the Assumption (3.6). Thus, by Lemma 3.3, we obtain as $N \to +\infty$

$$
\mathscr{I}_N^n \simeq \int_{\mathscr{S}} d\mathrm{A}(z_0) \int_{\substack{\mathscr{A}(z_0; \epsilon_N) \\ z_{n+1} = z_0}} F_N(z_0, \mathbf{z}) \prod_{j=0}^{n} K_V^N(z_j, z_{j+1}) d\mathrm{A}^n(\mathbf{z}). \tag{3.11}
$$

By (3.5), the set

$$
\mathscr{A}(z_0; \epsilon) \subset \left\{ \mathbf{z} \in \mathbb{C}^n : z_1, \ldots, z_n \in \mathbb{D}(z_0, n\epsilon_N) \right\}
$$

and we can apply Lemma 3.1 to replace the kernels $K_V^N(z_j, z_{j+1})$ in formula (3.11). Namely, if $z_{n+1} = z_0$ and $\mathbf{z} \in \mathscr{A}(z_0; \epsilon)$, then

$$
\left| \prod_{j=0}^{n} K_V^N(z_j, z_{j+1}) - \prod_{j=0}^{n} B_V^N(z_j, z_{j+1}) \right| \leq C \sum_{k=1}^{n+1} N^{-k} S_N^{n+1-k},
$$

where $S_N = \sup\left\{ |B_V^N(z, w)| : z, w \in \mathbb{D}(z_0, n\epsilon_N), z_0 \in \mathscr{S} \right\}$. By Lemma 3.2, we have for all $u, v \in \mathbb{D}(0, n\epsilon_N)$,

$$
\left| B_V^N(z_0 + u, z_0 + v) \right| \leq C \left| K_{N\varrho V(z_0)}^{\infty}(u, v) \right|.
$$

and, by formula (3.8), this implies that $S_N \leq CN$. If we combine these estimates with formula (3.11), since the functions F_N are uniformly bounded, we obtain

$$
\mathscr{I}_N^n = \int_{\mathscr{S}} d\mathrm{A}(z_0) \int_{\substack{\mathscr{A}(z_0; \epsilon_N) \\ z_{n+1} = z_0}} F_N(z_0, \mathbf{z}) \prod_{j=0}^{n} B_V^N(z_j, z_{j+1}) d\mathrm{A}^n(\mathbf{z})
$$

$$
+ \underset{N \to \infty}{O}\left(N^{n-1} \int_{\mathscr{S}} d\mathrm{A}(z_0) \left| \mathscr{A}(z_0; \epsilon_N) \right| \right),
$$

where $|\mathscr{A}|$ denotes the Lebesgue measure of the set $\mathscr{A}$. By definition, $\epsilon_N = \kappa N^{-1/2} \log N$ so that $\left| \mathscr{A}(z_0; \epsilon_N) \right| \leq CN^{-n}(\log N)^{2n}$ for all $z_0 \in \mathbb{C}$. Thus, the previous error term converges

to 0 like $(\log N)^{2n}/N$. Hence, if we make the change of variables $z = z_0 + u$ and the appropriate *gauge transform* in the previous integral, according to formula (3.3), we obtain

$$\mathscr{I}_N^n \simeq \int_{\mathscr{S}} d\mathrm{A}(z_0) \int_{\substack{\mathscr{A}(0;\epsilon_N) \\ u_{n+1}=u_0=0}} F_N(z_0 + u) \prod_{j=0}^{n} \widetilde{B}_{V,z_0}^N(u_j, u_{j+1}) d\mathrm{A}^n(u). \qquad (3.12)$$

Note that in formula (3.12), the integral is over a small subset of the surface $\{u \in \mathbb{C}^{n+2} : u_0 = u_{n+1} = 0\}$ and we denote $F_N(z_0 + u) = F_N(z_0, z_0 + u_1, \ldots, z_0 + u_n)$. Then, we can apply Lemma 3.2 to replace the kernel $\widetilde{B}_{V,z_0}^N$ by $K_{N\varrho_V(z_0)}^{\infty}$ in formula (3.12), we obtain

$$\mathscr{I}_N^n \simeq \int_{\mathscr{S}} d\mathrm{A}(z_0) \int_{\substack{\mathscr{A}(0;\epsilon_N) \\ u_{n+1}=u_0=0}} F_N(z_0 + u)\chi_N(z_0, u) \prod_{j=0}^{n} K_{N\varrho_V(z_0)}^{\infty}(u_j, u_{j+1}) d\mathrm{A}^n(u),$$

where $\chi_N(z_0, u) = 1 + \underset{N\to\infty}{O}\big((\log N)^2 \epsilon_N\big)$ uniformly for all $z_0 \in \mathscr{S}$ and all $u \in \mathscr{A}(0; \epsilon_N)$.

Let $F = \Upsilon^{n+1}[f]$, $\delta_N = \epsilon_N L_N$ and $\eta_N(\lambda) = N L_N^{-2} \varrho_V(x_0 + \lambda/L_N)$. By definition, $F_N(z_0 + u) = F\big(L_N(z_0 - x_0 + u)\big)$ and we can make the change of variables $\lambda = L_N(z_0 - x_0)$ and $w = L_N u$ to get rid of the scale L_N and x_0 in the previous integral. Using the obvious scaling property of the ∞-Ginibre kernel, (1.16), we obtain

$$\mathscr{I}_N^n \simeq \int_{\mathcal{K}} d\mathrm{A}(\lambda) \int_{\substack{\mathscr{A}(0;\delta_N) \\ w_{n+1}=w_0=0}} F(\lambda + w)\widetilde{\chi}_N(\lambda, w) \prod_{j=0}^{n} K_{\eta_N(\lambda)}^{\infty}(w_j, w_{j+1}) d\mathrm{A}^n(w), \quad (3.13)$$

where $\widetilde{\chi}_N(\lambda, w) = 1 + \underset{N\to\infty}{O}\big((\log N)^2 \epsilon_N\big)$ uniformly for all $\lambda \in \mathcal{K}$ and for all $u \in \mathscr{A}(0; \delta_N)$. Here we used that the test function f is supported in the set $\mathcal{K}$. The condition $\sum_{k \vdash n+1} \Upsilon(k) = 0$ implies that $F(\lambda + 0) = 0$ for all $\lambda \in \mathbb{C}$ so that for all $w \in \mathscr{A}(0; \delta_N)$,

$$\big|F(\lambda + w)\big| \le C\delta_N.$$

Moreover, by formula (3.8), we have for any $n \in \mathbb{N}$,

$$\prod_{j=0}^{n} \big|K_{\rho}^{\infty}(w_j, w_{j+1})\big| \le \rho^{n+1} \prod_{j=1}^{n} e^{-\rho|v_j|^2/2} \qquad (3.14)$$

where $v_j = w_j - w_{j-1}$ for all $j = 1, \ldots, n$. Hence, we see that

$$\left| \int_{\substack{\mathscr{A}(0;\delta_N) \\ w_{n+1}=w_0=0}} F(\lambda + w) \prod_{j=0}^{n} K_{\eta_N(\lambda)}^{\infty}(w_j, w_{j+1}) d\mathrm{A}^n(w) \right| \le C\delta_N \eta_N(\lambda)$$

and, since $\eta_N(\lambda) \le CNL_N^{-2}$ for all $\lambda \in \mathcal{K}$, we deduce from formula (3.13) that

$$\mathscr{I}_N^n = \int_{\mathcal{K}} d\mathrm{A}(\lambda) \int_{\substack{\mathscr{A}(0;\delta_N) \\ w_{n+1}=w_0=0}} F(\lambda + w) \prod_{j=0}^{n} K_{\eta_N(\lambda)}^{\infty}(w_j, w_{j+1}) d\mathrm{A}^n(w)$$

$$+ \underset{N\to\infty}{O}\big(NL_N^{-2}\delta_N\epsilon_N(\log N)^2\big). \qquad (3.15)$$

Recall that $\delta_N = L_N\epsilon_N$ and $\epsilon_N = \kappa N^{-1/2} \log N$, so that the error term in (3.15) is of order $(\log N)^4 L_N^{-1}$. Moreover, if $L_N = o\big(\sqrt{N}/\log N\big)$, a Taylor approximation shows that for any $w \in \mathscr{A}(0; \delta_N)$,

$$F(\lambda, \lambda + w_1, \ldots, \lambda + w_n) = H^n(\lambda; \mathrm{w}) + \underset{N \to \infty}{O}(\delta_N^3).$$

Using the estimate (3.14) once more, by formula (3.15), this implies that

$$\mathscr{I}_N^n = \int_{\mathcal{K}} d\mathrm{A}(\lambda) \int_{\substack{\mathscr{A}(0; \delta_N) \\ w_{n+1}=w_0=0}} H^n(\lambda; \mathrm{w}) \prod_{j=0}^{n} K^\infty_{\eta_N(\lambda)}(w_j, w_{j+1}) d\mathrm{A}^n(\mathrm{w})$$

$$+ \underset{N \to \infty}{O}\left(N L_N^{-2} \delta_N^3 \vee (\log N)^4 L_N^{-1}\right). \tag{3.16}$$

By Lemma 4.3, the leading term in formula (3.16) has the same limit (up to an arbitrary small error term) as

$$\int_{\mathcal{K}} d\mathrm{A}(\lambda) \int_{\substack{\mathbb{C}^n \\ w_{n+1}=w_0=0}} H^n(\lambda; \mathrm{w}) \prod_{j=0}^{n} K^\infty_{\eta_N(\lambda)}(w_j, w_{j+1}) d\mathrm{A}^n(\mathrm{w})$$

and, since $N L_N^{-2} \delta_N^3 \to 0$ when $L_N = o\left(\sqrt{N}/(\log N)^3\right)$, this completes the proof. $\square$

Since the function $\mathrm{w} \mapsto H^n(\lambda; \mathrm{w})$ is a multivariate polynomial of degree 2, the leading term in the Asymptotics (3.9) can be computed explicitly using the reproducing property of the ∞-Ginibre kernel; see for instance [44]. For any $\rho > 0$, the function $(z, w) \mapsto e^{\rho z \overline{w}}$ is the reproducing kernel for the Bergman space with weight $\rho e^{-\rho |z|^2/2}$ on $\mathbb{C}$. This implies that for any $w_1, w_2 \in \mathbb{C}$ and for all integer $k \geq 0$,

$$\begin{cases} \int_{\mathbb{C}} K^\infty_\rho(w_1, w_2) w_2^k K^\infty_\rho(w_2, w_3) d\mathrm{A}(w_2) = w_1^k K^\infty_\rho(w_1, w_3) \\ \int_{\mathbb{C}} K^\infty_\rho(w_1, w_2) \overline{w_2}^k K^\infty_\rho(w_2, w_3) d\mathrm{A}(w_2) = \overline{w_3}^k K^\infty_\rho(w_1, w_3) \end{cases}.$$

As a basic application of these identities, we obtain the following Lemma.

Lemma 3.5 *Let $n \geq 1$ and $\rho > 0$. For any polynomial $H(\mathrm{w})$ of degree at most 2 in the variables $w_1, \ldots, w_n, \overline{w_1}, \ldots, \overline{w_n}$ such that $H(0) = 0$, we have*

$$\int_{\substack{\mathbb{C}^n \\ w_{n+1}=w_0=0}} H^n(\mathrm{w}) \prod_{j=0}^{n} K^\infty_\rho(w_j, w_{j+1}) \, d\mathrm{A}^n(\mathrm{w}) = \sum_{1 \leq r \leq s \leq n} \partial_s \overline{\partial}_r H|_{\mathrm{w}=0}. \tag{3.17}$$

Under the assumptions of Proposition 3.4, since $\eta_N(\lambda) \to +\infty$ as $N \to +\infty$ uniformly for all $\lambda \in \mathcal{K}$, we deduce from Lemma 3.5 that for any test function $f \in \mathcal{C}_c^3(\mathcal{K})$, we have for all integers $n \geq 1$ and $m \geq 0$, as $N \to +\infty$,

$$\int_{\substack{\mathbb{C}^{n+1} \\ z_0=z_{n+1}}} \Upsilon_m^{n+1}[f_N](z_0, \mathrm{z}) \prod_{j=0}^{n} K_V^N(z_j, z_{j+1}) d\mathrm{A}(z_0) d\mathrm{A}^n(\mathrm{z})$$

$$\simeq \sum_{2 \leq r \leq s \leq n+1} \int_{\mathcal{K}} \partial_s \overline{\partial}_r \Upsilon_m^{n+1}[f](\lambda, \ldots, \lambda) d\mathrm{A}(\lambda). \tag{3.18}$$

Here we used that according to formulae (2.7) and (2.9), we have for any $m \geq 0$ and $n \geq 1$,

$$\Upsilon_m^{n+1}[f](\lambda, \ldots, \lambda) = f(\lambda)^{n+1} \sum_{\mathbf{k} \vdash n+1} \Upsilon_m(\mathbf{k}) = 0.$$

In the macroscopic regime ($L_N = 1$, $x_0 = 0$ and $\mathcal{K} = \text{supp}(f) \subset \mathscr{S}_V$), by Remark 3.2, this also shows that for any $n, m \geq 1$,

$$\int_{\substack{\mathbb{C}^{n+1} \\ z_0 = z_{n+1}}} \Upsilon_m^{n+1}[f](z) \prod_{j=0}^{n} K_V^N(z_j, z_{j+1}) dA^{n+1}(z) = \underset{N \to \infty}{O} \left((\log N)^4 \right).$$

This shows that the estimate (2.12) with $M_N = N$ holds for any sequence $q_N \searrow 0$ as $N \to +\infty$. By Theorem 2.2, this completes the proof of Theorem 1.1.

3.3 Mesoscopic Fluctuations for 2-Dimensional Coulomb Gases and the Proofs of Theorems 1.2 and 1.3

In the mesoscopic regime, we claim that the asymptotics (3.18) with $m = 0$ implies the Central Limit Theorem 1.2. Indeed, the fact that the cumulants of order $n \geq 3$ vanish in the large N limit comes from the following combinatorial Lemma.

Lemma 3.6 ([44], Lemma 9) *For any $n \geq 1$, let*

$$\mathscr{Y}_n = -\sum_{k \vdash n} \Upsilon_0(\mathbf{k}) \left\{ \sum_{2 \leq r < s \leq \ell(\mathbf{k})} k_r k_s + \sum_{r=2}^{\ell(\mathbf{k})} k_r (k_r - n) \right\}.$$

We have $\mathscr{Y}_n = \begin{cases} 1 & \text{if } n = 2 \\ 0 & \text{else} \end{cases}$.

Proof of Theorem 1.2 Let $\lambda \in \mathbb{C}$ and $\boldsymbol{\lambda} = (\lambda, \ldots, \lambda) \in \mathbb{C}^{n+1}$. According to formula (2.2), an elementary computation shows that for any $2 \leq r < s \leq n + 1$,

$$\partial_s \overline{\partial}_r \Upsilon_0^{n+1}[f](\boldsymbol{\lambda}) = \partial f(\lambda) \overline{\partial} f(\lambda) f(\lambda)^{n-1} \sum_{\mathbf{k} \vdash n+1} \Upsilon_0(\mathbf{k}) k_r k_s \mathbf{1}_{s \leq \ell(k)} \tag{3.19}$$

and

$$\partial_r \overline{\partial}_r \Upsilon_0^{n+1}[f]((\boldsymbol{\lambda}) = \partial f(\lambda) \overline{\partial} f(\lambda) f(\lambda)^{n-1} \sum_{\mathbf{k} \vdash n+1} \Upsilon_0(\mathbf{k}) k_r (k_r - 1) \mathbf{1}_{r \leq \ell(k)}$$
$$+ \Delta f(\lambda) f(\lambda)^n \sum_{\mathbf{k} \vdash n+1} \Upsilon_0(\mathbf{k}) k_r \mathbf{1}_{r \leq \ell(k)}. \tag{3.20}$$

Since, by integration by parts,

$$\int_{\mathbb{C}} \Delta f(\lambda) f(\lambda)^n dA(\lambda) = -n \int_{\mathbb{C}} \partial f(\lambda) \overline{\partial} f(\lambda) f(\lambda)^{n-1} dA(\lambda),$$

we deduce from formulae (3.19) and (3.20) that

$$\sum_{2 \leq r \leq s \leq n+1} \int_{\mathbb{C}} \partial_s \overline{\partial}_r \Upsilon_m^{n+1}[f](\lambda) dA(\lambda) = \mathscr{Y}_{n+1} \int_{\mathbb{C}} \partial f(\lambda) \overline{\partial} f(\lambda) f(\lambda)^{n-1} dA(\lambda).$$

When $L_N = N^\alpha$ and $0 < \alpha < 1/2$, formulae (2.5) and (3.18) with $m = 0$ imply that for any $n \geq 1$,

$$\lim_{N \to \infty} C_{K_V^N}^{n+1}[f_N] = \mathscr{Y}_{n+1} \int_{\mathbb{C}} \partial f(\lambda) \overline{\partial} f(\lambda) f(\lambda)^{n-1} dA(\lambda). \tag{3.21}$$

By Lemma 3.6, this proves that for any test function $f \in \mathcal{C}_0^3(\mathbb{C})$ and any $n \geq 2$,

$$\lim_{N \to \infty} C_{K_V^N}^n[f_N] = \begin{cases} \|f\|_{H^1(\mathbb{C})}^2 & \text{if } n = 2 \\ 0 & \text{else} \end{cases}. \tag{3.22}$$

This shows that the centered mesoscopic linear statistics $\Xi(f_N) - \mathbb{E}[\Xi(f_N)]$ converges in distribution as $N \to \infty$ to the mean-zero Gaussian random variable $X(f)$. □

We are now ready to finish the proof of Theorem 1.3. By Lemma 3.1, for any bounded function f with compact support, we have for any $n \geq 1$,

$$\int_{\mathbb{C}} f_N(z)^n K_V^N(z, z) dA(z) = N \int_{\mathbb{C}} f_N(z)^n 2\Delta V(z) dA(z) + \underset{N \to \infty}{O}(1)$$

$$= N L_N^{-2} \varrho_V(x_0) \int_{\mathbb{C}} f(z)^n dA(z) + \underset{N \to \infty}{O}(N L_N^{-3}).$$

Here we used that the potential V is smooth and $\rho_V = 2\Delta V > 0$ on a small neighborhood of the point $x_0 \in \mathscr{S}_V$. This implies the Assumption (2.11) with $M_N = N L_N^{1-2\alpha} \varrho_V(x_0)$—since the parameter $\alpha < 1/2$, $M_N \nearrow +\infty$ as $N \to +\infty$. As we already pointed out, the asymptotics (3.18) yield the Assumption (2.12) with an error which is $O(1)$. Finally, the Assumption (2.13) was proved just above—see (3.22). So, by Theorem 2.2, this completes the proof of Theorem 1.3.

4 Transition for 1-Dimensional Log-Gases

4.1 Asymptotics of Orthogonal Polynomials

In this section, we begin by reviewing basic facts about the asymptotics of orthogonal polynomials which are required for the proofs of Theorems 1.4 and 1.5. A comprehensive reference for the results discussed in this section is the book of Deift [22]. We assume that the potential $V \in C^2(\mathbb{R})$ is a function which satisfies the condition (1.1) and we let Ξ and $\widehat{\Xi}$ be the determinantal processes with correlation kernels K_V^N and $\widehat{K}_V^N = p_N K_V^N$ respectively.

The proof of Theorem 1.4 relies on a combinatorial method introduced in [17] which consists in using the three-terms recurrence relation of the orthogonal polynomials $\{P_k\}_{k=0}^{\infty}$ with respect to the measure $d\mu_N = e^{-2NV(x)}dx$ to compute the cumulants of polynomial linear statistics. For any $N \in \mathbb{N}$, there exists two sequences $a_k^N > 0$ and $b_k^N \in \mathbb{R}$ such that the orthogonal polynomials P_k in (1.4) satisfy

$$x P_k(x) = a_k^N P_{k+1}(x) + b_k^N P_k(x) + a_{k-1}^N P_{k-1}(x). \tag{4.1}$$

In particular, the completion $\mathscr{L}_N$ of the space of polynomials with respect to $L^2(\mathbb{R}, \mu_N)$ is isomorphic to $L^2(\mathbb{N}_0)$ and formula (4.1) implies that the multiplication by x on $\mathscr{L}_N$ is unitary equivalent to applying the **Jacobi matrix**

$$\mathbf{J} := \begin{bmatrix} b_0^N & a_0^N & 0 & 0 & 0 & \\ a_0^N & b_1^N & a_1^N & 0 & 0 & \mathbf{0} \\ 0 & a_1^N & b_2^N & a_2^N & 0 & \\ 0 & 0 & a_2^N & b_3^N & a_3^N & \\ & \mathbf{0} & & \ddots & \ddots & \ddots \end{bmatrix}. \tag{4.2}$$

We also let $\mathbf{\Pi}_N$ be the orthogonal projection on span$\{e_0, \ldots, e_{N-1}\}$ acting on $L^2(\mathbb{N}_0)$. The connection with eigenvalues statistics comes from the fact that for any polynomial Q and for any composition $\mathbf{k} \vdash n$, one has

$$\int_{\substack{\mathfrak{X}^l \\ x_0 = x_l}} Q(x_1)^{k_1} \cdots Q(x_l)^{k_l} \prod_{0 \le j \le l} K_V^N(x_j, x_{j-1}) dx_1 \cdots dx_l$$

$$= \mathrm{Tr}\left[Q(\mathbf{J})^{k_1} \mathbf{\Pi}_N \cdots Q(\mathbf{J})^{k_l} \mathbf{\Pi}_N\right]$$

$$= \sum_{m=0}^{N-1} \sum_{\pi \in \Gamma_m^n} \prod_{j=1}^{l} \mathbf{1}_{\pi(k_1 + \cdots + k_j) < N} \prod_{i=0}^{n-1} Q(\mathbf{J})_{\pi(i)\pi(i+1)},$$

where $\mathcal{G}$ denotes the adjacency graph of the matrix $Q(\mathbf{J})$ and

$$\Gamma_m^n = \left\{\text{paths } \pi \text{ on the graph } \mathcal{G} \text{ of length } n \text{ such that } \pi(0) = \pi(n) = m\right\}.$$

Given a path π of length n and a composition $\mathbf{k} \vdash n$, we let

$$\Phi_\pi^N(\mathbf{k}) := \mathbf{1} \max_{1 \le j < \ell(\mathbf{k})} \pi(k_1 + \cdots + k_j) \ge N. \tag{4.3}$$

Observe that

$$\prod_{1 \le j < \ell(\mathbf{k})} \mathbf{1}_{\pi(k_1 + \cdots + k_j) < N} = 1 - \Phi_\pi^N(\mathbf{k}),$$

so that by formula (2.5), the cumulants of a polynomial linear statistics are given by

$$C_{K_V^N}^n[Q] = -\sum_{m=0}^{N-1} \sum_{\pi \in \Gamma_m^n} \prod_{i=0}^{n-1} Q(\mathbf{J})_{\pi(i)\pi(i+1)} \sum_{\mathbf{k} \vdash n} \Upsilon_0(\mathbf{k})\left(1 - \Phi_\pi^N(\mathbf{k})\right). \tag{4.4}$$

By definitions, there exists a constant $M > 0$ which only depends on the degree of Q and n so that $\Phi_\pi^N = 0$ for any path $\pi \in \Gamma_m^n$ as long as $m < N - M$. Since $\sum_{\mathbf{k} \vdash n} \Upsilon_0(\mathbf{k}) = 0$ for all $n \ge 2$, formula (4.4) implies that

$$C_{K_V^N}^n[Q] = \sum_{m=N-M}^{N-1} \sum_{\pi \in \Gamma_m^n} \prod_{i=0}^{n-1} Q(\mathbf{J})_{\pi(i)\pi(i+1)} \sum_{\mathbf{k} \vdash n} \Upsilon_0(\mathbf{k}) \Phi_\pi^N(\mathbf{k}).$$

In particular, if the Jacobi matrix has a right-limit, i.e. there exists an (infinite) matrix $\mathbf{L}$ such that for all $i, j \in \mathbb{Z}$,

$$\lim_{N \to \infty} \mathbf{J}_{N+i, N+j} = \mathbf{L}_{i,j}$$

then

$$\lim_{N \to \infty} C_{K_V^N}^n[Q] = \sum_{m=1}^{M} \sum_{\pi \in \widetilde{\Gamma}_m^n} \prod_{i=0}^{n-1} Q(\mathbf{L})_{\pi(i)\pi(i+1)} \sum_{\mathbf{k} \vdash n} \Upsilon_0(\mathbf{k}) \Phi_\pi^0(\mathbf{k}), \tag{4.5}$$

where $\widetilde{\mathcal{G}}$ denotes the adjacency graph of the matrix $Q(\mathbf{L})$ and

$$\widetilde{\Gamma}_m^n = \left\{\text{paths } \pi \text{ on the graph } \widetilde{\mathcal{G}} \text{ of length } n \text{ such that } \pi(0) = \pi(n) = -m\right\}.$$

The condition (1.21) implies that the right-limit of the Jacobi matrix is a tridiagonal matrix $\mathbf{L}$ such that $\mathbf{L}_{jj} = 0$ and $\mathbf{L}_{j,j\pm1} = 1/2$ for all $j \in \mathbb{Z}$ and it was proved in [39], see also [17], that in this case:

$$\lim_{N \to \infty} C^n_{K^N_V}[Q] = \begin{cases} \sum_{k=1}^{\deg Q} k \left(\int_{-1}^{1} Q(x) T_k(x) \frac{dx}{\pi \sqrt{1 - x^2}} \right)^2 & \text{if } n = 2 \\ 0 & \text{else} \end{cases}. \quad (4.6)$$

The combinatorial method used in the previous section is well-suited to investigate the global fluctuations of 1-dimensional log-gas, but it is difficult to implement in the mesoscopic regime since we cannot use polynomials as test functions. So, to describe the transition for mesoscopic linear statistics and to prove Theorem 1.5, we rely on the asymptotics of the correlation kernel K^N_V from [24] and the method from [38] that we review below. Recall that ϱ_V is the equilibrium density of the gas and define *the integrated density of states*:

$$F_V(x) = \int_0^x \varrho_V(s) ds. \quad (4.7)$$

Let us fix $x_0 \in \mathscr{I}_V$, $0 < \alpha < 1$, and set

$$\widetilde{K}^N_{V,x_0}(x, y) = \frac{1}{N^\alpha} K^N_V \left(x_0 + \frac{x}{N^\alpha}, x_0 + \frac{y}{N^\alpha} \right). \quad (4.8)$$

Based on the results of [24], we have for any $\alpha \in (0, 1]$,

$$\widetilde{K}^N_{V,x_0}(x, y) = \frac{\sin \left[\pi N \left((F_V(x_0 + x N^{-\alpha}) - F_V(x_0 + y N^{-\alpha}) \right) \right]}{\pi (x - y)} + \underset{N \to \infty}{O} \left(N^{-\alpha} \right), \quad (4.9)$$

uniformly for all x, y in compact subsets of $\mathbb{R}$; c.f. [38, Proposition 3.5]. The main idea of the method of [38] is to compare the kernel (4.9) to the sine-kernel and use the results from Soshnikov [48] for the cumulants of linear statistics of the Sine process. We define the sine-kernel with density $\rho > 0$ on $\mathbb{R}$ by

$$K^{\sin}_\rho(x, y) = \frac{\sin[\pi \rho (x - y)]}{\pi (x - y)}. \quad (4.10)$$

We see by taking $\alpha = 1$ in formula (4.9) that the Sine process with correlation kernel (4.10) describes the local limit in the bulk of the 1-dimensional log-gases. In the mesoscopic regime, it was proved in [38] that, up to a change of variable, it is possible to replace the kernel $\widetilde{K}^N_{V,x_0}$ by an appropriate sine-kernel using the asymptotics (4.9) in the cumulant formulae. Namely, for any $n \geq 2$,

$$C^n_{K^N_V}[f_N] \simeq C^n_{K^{\sin}_{\eta_N(x_0)}} [f \circ \zeta_N] \quad (4.11)$$

where

$$\zeta_N(x) = N^\alpha \left\{ G_V \left(F_V(x_0) + \varrho_V(x_0) \frac{x}{N^\alpha} \right) - x_0 \right\}. \quad (4.12)$$

Here, the function G_V denotes the inverse of the integrated density of sates F_V, (4.7). By the inverse function theorem, it exists in a neighborhood of any point $F_V(x_0)$ when $x_0 \in \mathscr{I}_V$ and the map ζ_N is well-defined on any compact subset of $\mathbb{R}$ as long as the parameter N is sufficiently large. Then, using Soshnikov's main combinatorial Lemma, it was proved in [38]

that

$$
\lim_{N \to \infty} C^n_{K^N_V}[f_N] = \begin{cases} \int_0^\infty u|\hat{f}(u)|^2 du & \text{if } n = 2 \\ 0 & \text{if } n \geq 3 \end{cases}. \tag{4.13}
$$

4.2 Proof of Theorem 1.4: The Global Regime

In this section, we modify the strategy described above in order to deduce Theorem 1.4 from our general Theorem 2.2. The first step is to verify the Assumption (2.11). By [43, Theorem 11.1.2], the expected density of states satisfies for all $x \in \mathbb{R}$,

$$
u^N_V(x) \leq e^{-2N\{V(x) - \log(1+|x|) - C\}},
$$

where C is a constant which depends only on the potential V. This implies that the law of large numbers (1.5) can be extended to all continuous functions with polynomial growth. Moreover, by (4.6), we obtain the asymptotics (2.13) for the cumulants $C^n_{K^N_V}[Q]$ of the linear statistic $\Xi(Q)$. Then, it only remains to verify that the estimates (2.12) hold for any polynomial test function. By (2.9), since $\sum_{\mathbf{k} \vdash n} \Upsilon_m(\mathbf{k}) = 0$, the very same computation leading to (4.5) shows that for any integers $n, m \geq 1$

$$
\lim_{N \to \infty} \int_{\substack{\mathfrak{X}^n \\ x_0 = x_n}} \Upsilon^n_m[Q](x) \prod_{1 \leq j \leq n} K^N_V(x_j, x_{j-1}) d^n x
$$

$$
= \sum_{m=1}^M \sum_{\pi \in \widetilde{\Gamma}^n_m} \prod_{i=0}^{n-1} Q(\mathbf{L})_{\pi(i)\pi(i+1)} \sum_{\mathbf{k} \vdash n} \Upsilon_m(\mathbf{k}) \Phi^0_\pi(\mathbf{k}).
$$

Since the (infinite) matrix $\mathbf{L}$ is bounded with $\|\mathbf{L}\| \leq 1/2$, we obtain the estimates (2.12) with an error which is $O(1)$. This completes the proof of Theorem 1.4.

Remark 4.1 (Generalizations of Theorem 1.4) Note that we have formulated Theorem 1.4 for a log-gas at inverse temperature $\beta = 2$, but the previous proof can be generalized to other one-dimensional biorthogonal ensemble with a correlation kernel of the form:

$$
K^N(z, w) = \sum_{k=0}^{N-1} \varphi^N_k(z) \varpi^N_k(w).
$$

The appropriate assumptions are that there exists an equilibrium density and a law of large numbers holds for all polynomials, the family $\{\varphi^N_k\}_{k=0}^\infty$ satisfies a q-term recurrence relation for all $N \in \mathbb{N}$ and the corresponding recurrence matrix $\mathbf{J}$ has a right-limit $\mathbf{L}$ as $N \to \infty$. This applies to other orthogonal polynomial ensembles, such as the discrete point processes coming from domino tilings of hexagons, as well as some non-symmetric biorthogonal ensembles such as the Muttalib-Borodin ensembles, square singular values of product of complex Ginibre matrices or certain two-matrix models, see [17,39] for more details. Moreover, we only require that the right-limit $\mathbf{L}$ exists but it need not be a Toeplitz matrix. Then, we obtain a crossover from a non-Gaussian process (described by $\mathbf{L}$ in the regime where $T_N \to 0$) to a Poisson process (when $T_N \to \infty$). For instance, such a transition arises when considering linear statistics of the log-gases in the multi-cut regime [39].

Remark 4.2 It was proved in [31, Section 5] that when V is a convex polynomial, then $\mathscr{S}_V = (-1, 1)$ and the conditions (1.21) are satisfied. In fact, Johansson's argument shows that these

conditions are also necessary to have a CLT for polynomial test functions. Therefore, it is an interesting question to know whether (1.21) and the one-cut condition $\mathscr{S}_V = (-1, 1)$ are equivalent.

4.3 Proof of Theorem 1.5: The mesoscopic regime

Let us fix $x_0 \in \mathscr{I}_V$, $0 < \alpha < 1$ and let $f_N = f(N^\alpha(\cdot - x_0))$ where $f \in C_c^2(\mathbb{R})$. First, observe that by a change of variable, we have for all $n \geq 1$,

$$\int_\mathbb{R} f_N(x)^n K_V^N(x, x) dx = \int_\mathbb{R} f(x)^n \widetilde{K}_{V,x_0}^N(x, x) dx,$$

where $\widetilde{K}_{V,x_0}^N$ is given by (4.8). Using the asymptotics (4.9), we have that

$$\widetilde{K}_{V,x_0}^N(x, x) = N^{1-\alpha} \varrho_V(x_0) + \underset{N \to \infty}{O}\left(N^{-\alpha}\right)$$

uniformly for all $x \in \mathrm{supp}(f)$. If $M_N = N^{1-\alpha} \varrho_V(x_0)$, this implies that for all $n \in \mathbb{N}$,

$$\frac{1}{M_N} \int_\mathbb{R} f_N(x)^n K_V^N(x, x) dx \simeq \int_\mathbb{R} f(x)^n dx.$$

Thus, we obtain the condition (2.11) with $d\eta = dx$. Moreover, the Assumption (2.13) is given by (4.13). Then it just remains to prove the estimates (2.12). Let us observe that by a change of variables, we also have for all $n, m \geq 1$,

$$\int_{\substack{\mathbb{R}^n \\ u_0 = u_n}} \Upsilon_m^n[f_N](u) \prod_{1 \leq j \leq n} K_V^N(u_j, u_{j-1}) d\mu^n(u)$$

$$= \int_{\substack{\mathbb{R}^n \\ u_0 = u_n}} \Upsilon_m^n[f](u) \prod_{1 \leq j \leq n} \widetilde{K}_{V,x_0}^N(u_j, u_{j-1}) d\mu^n(u). \tag{4.14}$$

Exactly like the proof of (4.11)—which corresponds to the case $m = 0$ by formula (2.5)—we deduce from the proof of [38, Proposition 2.2] that for any $m \geq 1$,

$$\int_{\substack{\mathbb{R}^n \\ x_0 = x_n}} \Upsilon_m^n[f](x) \prod_{j \leq n} \widetilde{K}_{V,x_0}^N(x_j, x_{j-1}) d^n x \simeq \int_{\substack{\mathfrak{X}^n \\ x_0 = x_n}} \Upsilon_m^n[h_N](x) \prod_{1 \leq j \leq n} K_{\eta_N}^{\sin}(x_j, x_{j-1}) d^n x$$

$$\tag{4.15}$$

where $h_N = f \circ \zeta_N$ and ζ_N is given by (4.12). To finish the proof, we need the following estimates.

Proposition 4.1 *Suppose that* $\mathrm{supp}(f) \subset (-L, L)$. *There exists* $N_0 > 0$ *such that for all* $N \geq N_0$, *the functions* $h_N = f \circ \zeta_N$ *are well-defined on* $\mathbb{R}$, $h_N \in C_c^2([-L, L])$ *and for all* $u \in \mathbb{R}$,

$$\left|\widehat{h_N}(u)\right| \leq \|f\|_{C^2(\mathbb{R})} \frac{C}{1 + |u|^2}. \tag{4.16}$$

Proof When the potential V is analytic, the *bulk* $\mathscr{I}_V$ consists of finitely many open intervals and the equilibrium density ϱ_V is smooth on $\mathscr{I}_V$. Since $x_0 \in \mathscr{I}_V$, by formula (4.12), the function ζ_N is increasing and smooth on the interval $[-L, L]$ with

$$\zeta_N''(x) = \varrho_V(x_0)^2 G_V''\left(F_V(x_0) + \varrho_V(x_0) x N^{-\alpha}\right) N^{-\alpha}.$$

Moreover, since $\zeta_N(0) = 0$ and $\zeta_N'(0) = G_V'\big(F_V(x_0)\big)\varrho_V(x_0) = 1$, this implies that

$$\zeta_N(x) = x + \underset{N\to\infty}{O}(N^{-\alpha})$$

uniformly for all $x \in [-L, L]$. Since the open interval $(-L, L)$ contains the support of the test function f, this estimate shows that when the parameter N is large, we can define $h_N(x) = f\big(\zeta_N(x)\big)$ for all $x \in [-L, L]$ and extend it by 0 on $\mathbb{R}\backslash[-L, L]$. Then $h_N \in C_0^2(\mathbb{R})$ and

$$h_N''(x) = \zeta_N''(x)f'(\zeta_N(x)) + \zeta_N'(x)^2 f''(\zeta_N(x)) \tag{4.17}$$

for all $x \in [-L, L]$. Moreover, we can use the estimate

$$\big|\widehat{h_N}(u)\big| \le C\frac{\|h_N\|_\infty + \|h_N''\|_\infty}{1 + |u|^2} \tag{4.18}$$

to get the upper-bound (4.16). Plainly $\|h_N\|_\infty \le \|f\|_\infty$ and it is easy to deduce from formula (4.17) that

$$\|h_N''\|_\infty \le \|f'\|_\infty + C\|f''\|_\infty.$$

$\square$

To compute the limit of the RHS of (4.15), we also need the following asymptotics which come from the proof of Lemma 1 in Soshnikov's paper [48] on linear statistics of the CUE and Sine process.

Lemma 4.2 *Let $n \ge 2$ and let $\eta_N > 0$ such that $\eta_N \nearrow \infty$ as $\to +\infty$. Suppose that h_N is a sequence of integrable functions such that*

$$\lim_{N\to\infty} \int_{\substack{\mathbb{R}^{n-1}\\ u_1+\cdots+u_n=0\\ |u_1|+\cdots+|u_n|>\eta_N}} \big|\widehat{h_N}(u_1)\cdots\widehat{h_N}(u_n)\big||u_1|d^{n-1}\mathbf{u} = 0. \tag{4.19}$$

Then, for any map $\Upsilon : \mho \to \mathbb{R}$ such that $\sum_{\mathbf{k}\vdash n} \Upsilon(\mathbf{k}) = 0$, we have

$$\int_{\substack{\mathcal{X}^n\\ x_0=x_n}} \Upsilon^n[h_N](x) \prod_{1\le j\le n} K_{\eta_N}^{\sin}(x_j, x_{j-1})d^n x$$

$$\simeq - \int_{\substack{\mathbb{R}^{n-1}\\ u_1+\cdots+u_n=0}} \Re\left\{\prod_{j=1}^n \widehat{h_N}(u_j)\right\}\sum_{\mathbf{k}\vdash n}\Upsilon(\mathbf{k})\Psi_u(\mathbf{k})d^{n-1}u, \tag{4.20}$$

where for any $u \in \mathbb{R}^n$ and for any composition $\mathbf{k} \vdash n$,

$$\Psi_u(\mathbf{k}) = 2\max_{1\le j<\ell(\mathbf{k})}\{0, u_1 + \cdots + u_{k_1+\cdots+k_j}\}.$$

Proof Based on the formula

$$K_{\eta_N}^{\sin}(x, y) = \int_{\mathbb{R}} \mathbf{1}_{\{|u|<\eta_N/2\}}e^{2\pi iu(x-y)}du,$$

we obtain

$$\mathscr{T}_N := \int_{\substack{\mathcal{X}^n \\ x_0=x_n}} \Upsilon^n[h_N](x) \prod_{1 \leq j \leq n} K_{\eta_N}^{\sin}(x_j, x_{j-1}) d^n x$$

$$= \int_{\substack{\mathbb{R}^{n-1} \\ u_1+\cdots+u_n=0}} \prod_{j=1}^n \widehat{h_N}(u_j) \sum_{\mathbf{k} \vdash n} \Upsilon(\mathbf{k}) \max\left\{0, \eta_N - \Psi_u(\mathbf{k})/2 - \Psi_{-u}(\mathbf{k})/2\right\} d^{n-1}u.$$

Then, the condition $\sum_{\mathbf{k} \vdash n} \Upsilon(\mathbf{k}) = 0$ implies that

$$\left| \mathscr{T}_N + \int_{\substack{\mathbb{R}^{n-1} \\ u_1+\cdots+u_n=0}} \prod_{j=1}^n \widehat{h_N}(u_j) \sum_{\mathbf{k} \vdash n} \Upsilon(\mathbf{k}) \frac{\Psi_u(\mathbf{k}) + \Psi_{-u}(\mathbf{k})}{2} d^{n-1}u \right|$$

$$\leq \int_{\substack{\mathbb{R}^{n-1} \\ u_1+\cdots+u_n=0}} \left| \prod_{j=1}^n \widehat{h_N}(u_j) \right| \sum_{\mathbf{k} \vdash n} |\Upsilon(\mathbf{k})| \Psi_u(\mathbf{k}) \mathbf{1}_{\{\Psi_u(\mathbf{k}) + \Psi_{-u}(\mathbf{k}) \geq 2\eta_N\}} d^{n-1}u$$

Since $\left|\Psi_u(\mathbf{k})/2\right| \leq |u_1| + \cdots + |u_n|$ for any $\mathbf{k} \vdash n$, the condition (4.19) is sufficient to obtain the asymptotics (4.20). $\square$

We are now ready to complete the proof of Theorem 1.5. Using the estimate (4.16), we see that when the parameter N is sufficiently large, there exists a constant $C > 0$ so that

$$\left| \int_{\substack{\mathbb{R}^{n-1} \\ u_1+\cdots+u_n=0}} \Re\left\{ \prod_{j=1}^n \widehat{h_N}(u_j) \right\} \sum_{\mathbf{k} \vdash n} \Upsilon_m(\mathbf{k}) \Psi_u(\mathbf{k}) d^{n-1}u \right|$$

$$\leq C \int_{\substack{\mathbb{R}^{n-1} \\ u_1+\cdots+u_n=0}} \frac{|u_1| + \cdots + |u_n|}{(1 + |u_1|^2) \cdots (1 + |u_n|^2)} d^{n-1}u. \tag{4.21}$$

A similar upper-bound shows that the sequence $h_N = f \circ \zeta_N$ satisfies the condition (4.19) of Lemma 4.2. Hence, combining the asymptotics (4.15) and (4.20), we obtain

$$\int_{\substack{\mathbb{R}^n \\ x_0=x_n}} \Upsilon_m^n[f](x) \prod_{j \leq n} \widetilde{K}_{V,x_0}^N(x_j, x_{j-1}) d^n x \simeq$$

$$- \int_{\substack{\mathbb{R}^{n-1} \\ u_1+\cdots+u_n=0}} \Re\left\{ \prod_{j=1}^n \widehat{h_N}(u_j) \right\} \sum_{\mathbf{k} \vdash n} \Upsilon_m(\mathbf{k}) \Psi_u(\mathbf{k}) d^{n-1}u.$$

By (4.21), we see that the previous integral is uniformly bounded by a constant which depends only on the test function f and $n, m \in \mathbb{N}$. Hence, by (4.14), we obtain the estimates (2.12) with an error which is $O(1)$ and we can apply Theorem 2.2. This completes the proof of Theorem 1.5.

Appendix A: Incomplete Determinantal Processes

In this appendix, we review some background material on determinantal processes and we give an alternative proof of the results of Bohigas and Pato, [9], that the incomplete process is determinantal with correlation kernel $p_N K_V^N(z, w)$. There are many excellent surveys about determinantal processes and we refer to [11,33,36,47] for further details and an overview of the main examples.

Let $\mathfrak{X}$ be a complete separable metric space equipped with a Radon measure μ. The configuration space $\mathcal{Q}(\mathfrak{X})$ is the set of integer-valued locally finite Borel measure on $\mathfrak{X}$ equipped with the topology which is generated by the maps

$$\mathcal{Q}(\mathfrak{X}) \ni \Xi \mapsto \Xi_A := \int_A d\Xi$$

for every Borel set $A \subseteq \mathfrak{X}$. A point process $\mathbb{P}$ is a Borel probability measure on the configuration space $\mathcal{Q}(\mathfrak{X})$. A point process can be characterized by its correlation functions $(\rho_n)_{n=1}^\infty$. If it exists, ρ_n is a symmetric function on $\mathfrak{X}^n$ which satisfies the identity

$$\mathbb{E}\left[\binom{\Xi_{A_1}}{k_1} \cdots \binom{\Xi_{A_\ell}}{k_\ell}\right] = \frac{1}{k_1! \cdots k_\ell!} \int_{A_1^{k_1} \times \cdots \times A_\ell^{k_\ell}} \rho_n(x_1, \ldots, x_n) d\mu^n(x) \qquad \text{(A.1)}$$

for all compositions $\mathbf{k} \vdash n$ and for all disjoint Borel sets $A_1, \ldots, A_\ell \subseteq \mathfrak{X}$. A point process is called **determinantal** if all its correlation functions exist and are given by

$$\rho_n(x_1, \ldots, x_n) = \det_{n \times n}\left[K(x_i, x_j)\right]. \qquad \text{(A.2)}$$

The function $K : \mathfrak{X} \times \mathfrak{X} \to \mathbb{C}$ is called the **correlation kernel**. For instance, given random points $(\lambda_1, \ldots, \lambda_N)$ with the joint density $G_N(x) = e^{-\beta \mathcal{H}_V^N(x)}/Z_V^N$ on $\mathfrak{X}^N$, see (1.2), the random measure $\sum_{k=1}^N \delta_{\lambda_k}$ defines a point process and its correlation functions satisfy $\rho_N = N! G_N$ and

$$\rho_k(x_1, \ldots, x_k) = \frac{N!}{(N-k)!} \int G_N(x_1, \ldots, x_N) dx_{k+1} \cdots dx_N \qquad \text{(A.3)}$$

for all $k < N$. When $\beta = 2$, it is easy to verify that

$$G_N(x) = \frac{1}{N!} \det_{N \times N}\left[K_V^N(x_i, x_j)\right]$$

and that, by formula (A.3), the process Ξ is determinantal with the correlation kernel K_V^N given by (1.3). In general, if K is a continuous, Hermitian symmetric, function which satisfies property (2.1), then there exists a determinantal process on $\mathfrak{X}$ with correlation kernel K; c.f. [47, Theorem 3].

Let $q \in (0, 1)$ and $p = 1 - q$. Recall that X is a Binomial random variable with parameter p and $N \in \mathbb{N}$ if for all $k \in \mathbb{N}_0$,

$$\mathbb{E}\left[\binom{X}{k}\right] = p^k \binom{N}{k}. \qquad \text{(A.4)}$$

By convention, $\binom{N}{k} = 0$ if $k > N$. Let Ξ be a random (point) configuration and let $\widehat{\Xi}$ be the configuration obtained after performing a Bernoulli percolation on Ξ. By construction, for any disjoint Borel set $A \subseteq \mathfrak{X}$, the conditional distribution of the random variable $\widehat{\Xi}_A$ given

Ξ has a Binomial distribution with parameters p and Ξ_A and it is statistically independent of Ξ_B for any Borel set B disjoint of A. By formula (A.4), this implies that

$$\mathbb{E}\left[\binom{\widehat{\Xi}_{A_1}}{k_1}\cdots\binom{\widehat{\Xi}_{A_\ell}}{k_\ell}\Bigg|\Xi_A = N_1, \ldots, \Xi_A = N_\ell\right] = p^n\binom{N_1}{k_1}\cdots\binom{N_\ell}{k_\ell}$$

for any composition $\mathbf{k} \vdash n$ and for all disjoint Borel sets $A_1, \ldots, A_\ell \subseteq \mathfrak{X}$. Hence, we obtain

$$\mathbb{E}\left[\binom{\widehat{\Xi}_{A_1}}{k_1}\cdots\binom{\widehat{\Xi}_{A_\ell}}{k_\ell}\right] = p^n\mathbb{E}\left[\binom{\Xi_{A_1}}{k_1}\cdots\binom{\Xi_{A_\ell}}{k_\ell}\right]$$

so that, by formula (A.1), the correlation functions of *the incomplete process* $\widehat{\Xi}$ are given by $p^n\rho_n(x_1, \ldots, x_n)$ for all $n \geq 1$. In particular, we deduce from formula (A.2) that, if Ξ is a determinantal process with a correlation kernel K, then the point process $\widehat{\Xi}$ is also determinantal with kernel pK.

Appendix B: Off-Diagonal Decay of the Correlation Kernel K_V^N in Dimension 2

In this section, we review some classical estimates for the correlation kernel (1.3) which have been used in [2] to prove the CLT (1.10). Then, we prove Lemma 3.3 and an analogous result for the cumulants of the ∞-Ginibre process. For completeness, we also give the proof of Lemma 3.2. We will use the formulation of Sect. 5 in [2] but the estimates (B.1) and (B.2) go back to the papers [7] and [1]. Suppose that the potential $V : \mathbb{C} \to \mathbb{R}$ is real-analytic and satisfies the condition (1.1). Then, there a function $\phi_V : \mathbb{C} \to \mathbb{R}^+$ such that $\phi_V(z) \geq \nu \log|z|^2$ as $|z| \to \infty$ and some constants $C, c, \delta > 0$ such that

$$\left|K_V^N(z, w)\right| \leq CNe^{-c\sqrt{N}(|z-w|\wedge\delta)}, \tag{B.1}$$

and

$$K_V^N(z, z) \leq CNe^{-N\phi_V(z)}, \tag{B.2}$$

for all $w \in \mathscr{S}$ and $z \in \mathbb{C}$.

Proof of Lemma 3.3 We use the convention $z_0 = z_{n+1}$. Since the kernel K_V^N is reproducing, by the Cauchy-Schwartz inequality,

$$\left|K_V^N(z, w)\right| \leq \sqrt{K_V^N(z, z)K_V^N(w, w)},$$

for all $z, w \in \mathbb{C}$, so that

$$\left|\prod_{j=0}^n K_V^N(z_j, z_{j+1})\right| \leq \left|K_V^N(z_0, z_1)\right|\sqrt{K_V^N(z_0, z_0)K_V^N(z_1, z_1)}\prod_{j=2}^n K_V^N(z_j, z_j).$$

Since $\int_{\mathbb{C}} K_V^N(z, z)d\mathrm{A}(z) = N$ and $K_V^N(z, z) \leq CN$ [see the estimate (B.2)], we obtain

$$\int_{\substack{\mathbb{C}^n\times\mathscr{S}\\|z_1-z_{n+1}|>\epsilon_N}}\left|\prod_{j=0}^n K_V^N(z_j, z_{j+1})\right|d\mathrm{A}^{n+1}(z) \leq CN^n\int_{\substack{\mathscr{S}\times\mathbb{C}\\|z_1-z_0|>\epsilon_N}}\left|K_V^N(z_0, z_1)\right|d\mathrm{A}(z_0)d\mathrm{A}(z_1).$$

Then, it easy to check that the estimates (B.1) and (B.2) imply that

$$\int_{\substack{\mathscr{S}\times\mathbb{C}\\|z_1-z_0|>\epsilon_N}} \left|K_V^N(z_0,z_1)\right| d\mathrm{A}(z_0)d\mathrm{A}(z_1) \le CNe^{-c\sqrt{N}\epsilon_N}.$$

Hence, if $\epsilon_N = \kappa N^{-1/2}\log N$ and $\kappa \ge (n+1)/c$, we obtain

$$\int_{\substack{\mathbb{C}^n\times\mathscr{S}\\|z_1-z_{n+1}|>\epsilon_N}} \left|\prod_{j=0}^n K_V^N(z_j,z_{j+1})\right| d\mathrm{A}^{n+1}(z) = \underset{N\to\infty}{O}(N^{-1}). \tag{B.3}$$

Moreover, since $\sup\left\{|F_N(z_0,z)| : z \in \mathbb{C}^n, N \ge N_0\right\} \le C\mathbf{1}_{z_0\in\mathscr{S}}$ by (3.6), the estimate (B.3) implies that

$$\int_{\substack{\mathbb{C}^{n+1}\\z_0=z_{n+1}}} F_N(z)\prod_{j=0}^n K_V^N(z_j,z_{j+1})d\mathrm{A}^{n+1}(z)$$

$$= \int_{\substack{\mathbb{C}^n\times\mathscr{S}\\|z_1-z_{n+1}|\le\epsilon_N}} F_N(z)\prod_{j=0}^n K_V^N(z_j,z_{j+1})\, d\mathrm{A}^{n+1}(z) + \underset{N\to\infty}{O}(N^{-1}). \tag{B.4}$$

Now, we can proceed by induction to get formula (3.7). If $\mathscr{C}_N = \{z \in \mathbb{C}^{n+1} : z_{n+1} \in \mathscr{S}, |z_1 - z_{n+1}| \le \epsilon_N\}$, the next step is to show that

$$\int_{\substack{\mathscr{C}_N\\|z_2-z_1|>\epsilon_N}} \left|\prod_{j=0}^n K_V^N(z_j,z_{j+1})\right| d\mathrm{A}^{n+1}(z) = \underset{N\to\infty}{O}(N^{-1}). \tag{B.5}$$

Since the set $\mathscr{S}_V$ is open, there exists a compact set $\mathscr{S}' \subset \mathscr{S}_V$ such that $\mathscr{S} \subset \mathscr{S}'$ and $\mathscr{C}_N \subset \{z \in \mathbb{C}^{n+1} : z_{n+1}, z_1 \in \mathscr{S}'\}$ when the parameter N is sufficiently large. Then, as before, we obtain

$$\int_{\substack{\mathscr{C}_N\\|z_2-z_1|>\epsilon_N}} \left|\prod_{j=0}^n K_V^N(z_j,z_{j+1})\right| d\mathrm{A}^{n+1}(z) \le CN^n \int_{\substack{\mathscr{S}'\times\mathbb{C}\\|z_2-z_1|>\epsilon_N}} \left|K_V^N(z_1,z_2)\right| d\mathrm{A}(z_0)d\mathrm{A}(z_1),$$

and formula (B.5) also follows directly from the estimate (B.3). Hence, by formula (B.4), this implies that

$$\int_{\substack{\mathbb{C}^{n+1}\\z_0=z_{n+1}}} F_N(z)\prod_{j=0}^n K_V^N(z_j,z_{j+1})d\mathrm{A}^{n+1}(z)$$

$$= \int_{\substack{\mathbb{C}^n\times\mathscr{S}\\|z_1-z_{n+1}|\le\epsilon_N\\|z_2-z_1|\le\epsilon_N}} F_N(z)\prod_{j=0}^n K_V^N(z_j,z_{j+1})\, d\mathrm{A}^{n+1}(z) + \underset{N\to\infty}{O}(N^{-1}).$$

If we repeat this argument, we obtain (3.7). $\qquad\qquad\square$

Lemma 4.3 *Let* $n \in \mathbb{N}$, $w_0 = w_{n+1} = 0$, *and let* $H(w)$ *be a polynomial of degree at most 2 in the variables* $w_1, \ldots, w_n, \overline{w_1}, \ldots, \overline{w_n}$. *For any sequence* $\delta_N \ge k\rho_N^{-1/2}\sqrt{\log\rho_N}$ *with* $k > 0$, *we have*

$$\int_{\mathbb{C}^n} H(\mathrm{w}) \prod_{j=0}^{n} K_{\rho_N}^{\infty}(w_j, w_{j+1})\, d\mathrm{A}^n(\mathrm{w})$$

$$= \int_{\mathscr{A}(0;\delta_N)} H(\mathrm{w}) \prod_{j=0}^{n} K_{\rho_N}^{\infty}(w_j, w_{j+1})\, d\mathrm{A}^n(\mathrm{w}) + \underset{N\to\infty}{O}(\rho_N^{1-k^2/2})$$

where the set $\mathscr{A}(0; \delta_N)$ is given by formula (3.5).

Proof We will first show that

$$\int_{\substack{\mathbb{C}^n \\ |w_1|>\delta_N}} H(\mathrm{w}) \prod_{j=0}^{n} K_{\rho_N}^{\infty}(w_j, w_{j+1})\, d\mathrm{A}^n(\mathrm{w}) = \underset{N\to\infty}{O}\left(\rho_N e^{-\rho_N \delta_N^2/2}\right). \tag{B.6}$$

First, notice that if $H = 1$, since $w_0 = 0$, by formula (3.14), we have

$$\int_{\substack{\mathbb{C}^n \\ |w_1|>\delta_N}} \left| \prod_{j=0}^{n} K_{\rho_N}^{\infty}(w_j, w_{j+1}) \right| d\mathrm{A}^n(\mathrm{w}) \le e^{-\rho_N \delta_N^2/2} \rho_N^{n+1} \underbrace{\int_{\mathbb{C}^n} \prod_{j=1}^{n} e^{-\rho_N |v_j|^2/2} d\mathrm{A}^n(\mathrm{v})}_{=\, \rho_N^{-n}}.$$

In general, there exists a constant $C > 0$ which only depends on the polynomial H so that

$$\left| H(\mathrm{w}) \right| \le C\left\{ 1 + |w_1|^2 + \cdots + |w_n|^2 \right\}$$

or

$$\left| H(\mathrm{w}) \right| \le C\left\{ 1 + |w_2 - w_1|^2 + \cdots + |w_{n-1} - w_n|^2 + |w_n|^2 \right\}. \tag{B.7}$$

Since, for any $k = 1, \ldots, n$,

$$\int_{\substack{\mathbb{C}^n \\ |w_1|>\delta_N}} |w_k - w_{k+1}|^2 \left| \prod_{j=0}^{n} K_{\rho_N}^{\infty}(w_j, w_{j+1}) \right| d\mathrm{A}^n(\mathrm{w})$$

$$\le e^{-\rho_N \delta_N^2/2} \rho_N^{n+1} \underbrace{\int_{\mathbb{C}^n} |v_k|^2 \prod_{j=1}^{n} e^{-\rho_N |v_j|^2/2} d\mathrm{A}^n(\mathrm{v})}_{=\, 1},$$

the estimate (B.6) follows directly from (B.7) and the leading contribution comes from the constant term. If we use the estimate

$$\left| H(\mathrm{w}) \right| \le C\left\{ 1 + |w_1|^2 + \sum_{\substack{j=1 \\ j\ne k}}^{n} |w_{j+1} - w_j|^2 + |w_n|^2 \right\}$$

instead, the same argument shows that for any $k = 1, \ldots n$,

$$\int_{|w_k - w_{k+1}|>\delta_N} H(\mathrm{w}) \prod_{j=0}^{n} K_{\rho_N}^{\infty}(w_j, w_{j+1})\, d\mathrm{A}^n(\mathrm{w}) = \underset{N\to\infty}{O}\left(\rho_N e^{-\rho_N \delta_N^2/2}\right).$$

Hence, the Lemma follows from applying a union bound and from the choice of the sequence δ_N. $\qquad\square$

Proof of Lemma 3.2 The map $(z, w) \mapsto \Psi(z, w)$ is bi-holomorphic in a neighborhood of $(x_0, \overline{x_0})$, so there exists $0 < \epsilon < 1$ such that for all $|u|, |v| \leq \epsilon$,

$$\Psi(x_0 + u, \overline{x_0} + v) = \sum_{k, j \geq 0} a_{kj} u^k v^j.$$

By definition, $\overline{\Psi(z, w)} = \Psi(\overline{w}, \overline{z})$, so that the coefficients of the previous power series are Hermitian-symmetric: $a_{kj} = \overline{a_{jk}}$ for all $k, j \geq 0$. Moreover, by definition, we have

$$a_{11} = \partial_{\overline{z}} \partial_z V|_{z=x_0} = \Delta V(x_0) = \frac{b_0(x_0, \overline{x_0})}{2}. \tag{B.8}$$

Let

$$\mathfrak{h}(u) = i \sum_{k>0} \left\{ a_{k0} u^k - a_{0k} \overline{u}^k \right\} = -2 \Im \left\{ \sum_{k>0} a_{k0} u^k \right\}.$$

Since $V(x_0 + u) = \Psi(x_0 + u, \overline{x_0} + \overline{u})$, we see that for any $|u|, |v| \leq \epsilon$,

$$2 \Psi(x_0 + u, \overline{x_0} + \overline{v}) - V(x_0 + u) - V(x_0 + v)$$
$$= -i \{ \mathfrak{h}(u) - \mathfrak{h}(v) \} + a_{11} \left(2u\overline{v} - |\overline{u}|^2 - |\overline{v}|^2 \right) + O(\epsilon^3).$$

By formula (3.1), this implies that for any $|u|, |v| \leq \epsilon_N = \log(N^\kappa) N^{-1/2}$,

$$\frac{B^N(x_0 + u, x_0 + v) e^{iN\mathfrak{h}(u)}}{e^{iN\mathfrak{h}(v)}} = Nb_0(x_0, \overline{x_0}) e^{Na_{11}(2u\overline{v} - |\overline{u}|^2 - |\overline{v}|^2)} \left\{ 1 + \underset{N \to \infty}{O} \left((\log N)^2 \epsilon_N \right) \right\}.$$

By formula (B.8) and the definition of the ∞-Ginibre kernel, it completes the proof. $\square$

References

1. Ameur, Y., Hedenmalm, H., Makarov, N.: Berezin transform in polynomial Bergman spaces. Commun. Pure Appl. Math. **63**, 1533–1584 (2010)
2. Ameur, Y., Hedenmalm, H., Makarov, N.: Fluctuations of eigenvalues of random normal matrices. Duke Math. J. **159**, 31–81 (2011)
3. Ameur, Y., Hedenmalm, H., Makarov, N.: Random normal matrices and Ward identities. Ann. Probab. **43**, 1157–1201 (2015)
4. Anderson, G.W., Guionnet, A., Zeitouni, O.: An Introduction to Random Matrices. Cambridge Univ. Press, Cambridge (2010)
5. Bekerman, F., Lodhia, A.: Mesoscopic central limit theorem for general β-ensembles. Ann. Inst. Henri Poincar Probab. Stat. **54**(4), 1917–1938 (2018)
6. Berggren, T., Duits, M.: Mesoscopic fluctuations for the thinned circular unitary ensemble. Math. Phys. Anal. Geom. **20**(3), 19 (2017)
7. Berman, R.J.: Bergman kernels for weighted polynomials and weighted equilibrium measures of $\mathbb{C}^n$. Indiana U. Math. J. **58**, 1921–1946 (2009)
8. Bohigas, O., Pato, M.P.: Missing levels and correlated spectra. Phys. Lett. B **595**(036212), 171–176 (2004)
9. Bohigas, O., Pato, M.P.: Randomly incomplete spectra and intermediate statistics. Phys. Rev. E **3**(74), 036212 (2006)
10. Bauerschmidt, R., Bourgade, P., Nikula, M., Yau, H.-T.: The two-dimensional Coulomb plasma: quasi-free approximation and central limit theorem. arXiv:1609.08582
11. Borodin, A.: Determinantal point processes, in the Oxford handbook of random matrix theory, pp. 231–249. Oxford Univ. Press, Oxford (2011)
12. Borot, G., Guionnet, A.: Asymptotic expansion of beta matrix models in the one-cut regime. Commun. Math. Phys. **317**(2), 447–483 (2013)
13. Borot, G., Guionnet, A.: Asymptotic expansion of beta matrix models in the multi-cut regime. arXiv:1303.1045

14. Bothner, T., Deift, P., Its, A., Krasovsky, I.: On the asymptotic behavior of a log gas in the bulk scaling limit in the presence of a varying external potential I. Commun. Math. Phys. **337**, 1397–1463 (2015)

15. Bothner, T., Deift, P., Its, A., Krasovsky, I.: On the asymptotic behavior of a log gas in the bulk scaling limit in the presence of a varying external potential II. In Large truncated Toeplitz matrices, Toeplitz operators, and related topics, pp. 213–234. Birkhuser/Springer, Cham (2017)

16. Breuer, J., Duits, M.: Universality of mesoscopic fluctuations for orthogonal polynomial ensembles. Commun. Math. Phys. **342**, 491–531 (2016)

17. Breuer, J., Duits, M.: Central limit theorems for biorthogonal ensembles and asymptotics of recurrence coefficients. J. Am. Math. Soc. **30**(1), 27–66 (2017)

18. Chafaï, D., Hardy, A., Maïda, M.: Concentration for Coulomb gases and Coulomb transport inequalities. J. Funct. Anal. **275**(6), 1447–1483 (2018)

19. Charlier, C., Claeys, T.: Thinning and conditioning of the circular unitary ensemble. Random Matrices **6**(2), 1750007 (2017)

20. Costin, O., Lebowitz, J.: Gaussian fluctuations in random matrices. Phys. Rev. Lett. **75**, 69–72 (1995)

21. Dean, D.S., Doussal, P.L., Majumdar, S.N., Schehr, G.: Finite temperature free fermions and the Kardar–Parisi–Zhang equation at finite time. Phys. Rev. Lett. **114**(11), 110402 (2015)

22. Deift, P.: Orthogonal Polynomials and Random Matrices: A Riemann–Hilbert approach. Courant Lecture Notes in Mathematics, vol. 3. Courant Institute of Mathematical Sciences; American Mathematical Society, New York; Providence (1999)

23. Deift, P.: Some open problems in random matrix theory and the theory of integrable systems. II. SIGMA **13**, 016 (2017)

24. 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. Commun. Pure Appl. Math. **52**(11), 1335–1425 (1999)

25. Duits, M., Johansson, K.: On mesoscopic equilibrium for linear statistics in Dyson's Brownian motion. Mem. Am. Math. Soc. **255**(1222), 118 (2018)

26. Dyson, F.J.: The Coulomb fluid and the fifth Painlevé transcendent. In: Liu, C.S., Yau, S.T. (eds.) Chen Ning Yang: A Great Physicist of the Twentieth Century, pp. 131–146. International Press, Cambridge (1995)

27. Erdős, L., Knowles, A.: The Altshuler-Shklovskii formulas for random band matrices I: the unimodular case. Commun. Math. Phys. **333**, 1365–1416 (2015)

28. Erdős, L., Knowles, A.: The Altshuler–Shklovskii formulas for random band matrices II: the general case. Ann. Henri Poincaré **16**, 709–799 (2015)

29. He, Y., Knowles, A.: Mesoscopic eigenvalue statistics of Wigner matrices. Ann. Appl. Probab. **27**(3), 1510–1550 (2017)

30. Hough, B., Krishnapur, M., Peres, Y., Virág, B.: Determinantal processes and independence. Probab. Surv. **3**, 206–229 (2006)

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

32. Johansson, K.: Universality of the local spacing distribution in certain ensembles of Hermitian Wigner matrices. Commun. Math. Phys. **215**, 683–705 (2001)

33. Johansson, K.: Random Matrices and Determinantal Processes. Mathematical Statistical Physics, pp. 1–55. Elsevier B. V, Amsterdam (2006)

34. Johansson, K.: From Gumble to Tracy–Widom. Probab. Theory Relat. Fields **138**, 75–112 (2007)

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

36. König, W.: Orthogonal polynomial ensembles in probability theory. Probab. Surv. **2**, 385–447 (2005)

37. Kuijlaars, A.B.J.: Universality in the Oxford Handbook of Random Matrix Theory, pp. 103–134. Oxford Univ. Press, Oxford (2011)

38. Lambert, G.: Mesoscopic fluctuations for unitary invariant ensembles. Electron. J. Probab. **23**(7), 33 (2018)

39. Lambert, G.: CLT for biorthogonal ensembles and related combinatorial identities. Adv. Math. **329**, 590–648 (2018)

40. Lavancier, F., Møller, J., Rubak, E.: Determinantal point process models and statistical inference. J. R. Stat. Soc. Ser. B **77**(4), 853–877 (2015)

41. Leblé, T., Serfaty, S.: Fluctuations of two-dimensional Coulomb gases. Geom. Funct. Anal. **28**(2), 443–508 (2018)

42. Pastur, L.A.: Limiting laws of linear eigenvalue statistics for Hermitian matrix models. J. Math. Phys. **47**, 103303 (2003)

43. Pastur, L.A., Shcherbina, M.: Eigenvalue Distribution of Large Random Matrices Mathematical Surveys and Monographs, vol. 171. American Mathematical Society, Providence, RI (2011)
44. Rider, B., Virág, B.: Complex determinantal processes and H^1 noise. Electron. J. Probab. **12**, 1238–1257 (2007)
45. Rider, B., Virág, B (2007) The noise in the circular law and the Gaussian free field. Int. Math. Res. Notes. https://doi.org/10.1093/imrn/rnm006 (2007)
46. Shcherbina, M.: Fluctuations of linear eigenvalue statistics of β-matrix models in the multi-cut regime. J. Stat. Phys. **151**, 1004–1034 (2013)
47. Soshnikov, A.: Determinantal random point fields. Russ. Math. Surv. **55**, 923–975 (2000)
48. Soshnikov, A.: The Central Limit Theorem for local linear statistics in classical compact groups and related combinatorial identities. Ann. Probab. **28**, 1353–1370 (2000)
49. Soshnikov, A.: Gaussian limit for determinantal random point fields. Ann. Probab. **30**, 1–17 (2001)
50. Spencer, T.: Random Banded and Sparse Matrices in the Oxford Handbook of Random Matrix Theory, pp. 471–488. Oxford Univ. Press, Oxford (2011)