GAUSSIAN ASYMPTOTICS OF DISCRETE β-ENSEMBLES

by Alexei BORODIN, Vadim GORIN, *and* Alice GUIONNET

ABSTRACT

We introduce and study stochastic N-particle ensembles which are discretizations for general-β log-gases of random matrix theory. The examples include random tilings, families of non-intersecting paths, (z, w)-measures, etc. We prove that under technical assumptions on general analytic potential, the global fluctuations for such ensembles are asymptotically Gaussian as $N \to \infty$. The covariance is universal and coincides with its counterpart in random matrix theory.

Our main tool is an appropriate discrete version of the Schwinger-Dyson (or loop) equations, which originates in the work of Nekrasov and his collaborators.

CONTENTS

1. Introduction

1.1. *Continuous log-gases.* — A general-β log-gas is a probability distribution on N-tuples of reals $x_1 < x_2 < \cdots < x_N$ with density proportional to

$$(\mathbf{1}) \qquad \prod_{1 \le i \le j \le N} (x_j - x_i)^\beta \prod_{i=1}^{N} \exp\big(-NV(x_i)\big),$$

where $V(x)$ is a continuous function called *potential*. For $V(x) = x^2$ and $\beta = 1, 2, 4$, the density (1) describes the joint distribution of the eigenvalues of random matrices from Gaussian Orthogonal/Unitary/Symplectic Ensemble; much more general potentials $V(x)$ are widespread and extensively studied in random matrix theory and beyond, see the books [Me], [Fo], [AGZ], [Ox], [PS].

Under weak assumptions on the potential, the ensembles (1) exhibit a Law of Large Numbers as $N \to \infty$, which means that the (random) empirical measure μ_N defined via

$$\mu_N = \frac{1}{N} \sum_{i=1}^{N} \delta_{x_i}$$

converges (weakly, in probability) to a non-random *equilibrium measure* μ. For $\beta = 1, 2, 4$ and $V(x) = x^2$, this statement dates back to the work of Wigner [Wi], and μ is known in this case as the Wigner *semicircle law*. The results for generic $V(x)$ were established much later, see [BPS], [BeGu], [J1].

The next order asymptotics asks about *global fluctuations*, i.e. how the functional $\mu_N - \mu$ behaves as $N \to \infty$. A natural approach here is to take (sufficiently smooth) functions $f(x)$ and consider the asymptotic behavior of random variables

$$(\textbf{2}) \qquad N\left(\int f(x)\mu_N(dx) - \int f(x)\mu(dx) \right), \quad N \to \infty.$$

For quite general potentials $V(x)$ the limits of (2) are Gaussian with *universal* covariance depending only on the support of the equilibrium measure μ. In the breakthrough paper [J1] Johansson proved such a statement for general $\beta > 0$ and wide class of potentials under the assumption that μ has a single interval of support. Further developments have led to establishing such results for *generic* analytic potentials, see [KS], [BoGu1], [Shch], [BoGu2]. Note that when the support of μ has several intervals one needs to be careful as an additional *discrete* component might appear. This does not happen if one deterministically fixes the *filling fractions*, which are the numbers of particles in each interval of the support. In the one-interval case the limiting covariance can be identified with that of a $1d$ section of the two-dimensional Gaussian Free Field, see [B2], [BoGo2] for the details.

While for certain specific choices of potentials $V(x)$ as well as for special values $\beta = 1, 2, 4$ there are several different methods for establishing central limit theorems for global fluctuations, all the developments for generic β and $V(x)$ rely on the analysis of *loop equations* (also known as Schwinger-Dyson equations). These are equations for certain observables of the log-gases (1), which originated and have been widely used in physics literature, cf. [Mi], [AM], [Ey1], [CE] and references therein. More precisely, the observables of interest are the Stieltjes transforms of the empirical measure μ_N and the Schwinger-Dyson equations involve their cumulants. *Formally assuming* that these cumulants have a converging N^{-1} expansion as $N \to \infty$, one can use the equations to derive recursively

the asymptotics of the cumulants starting from the equilibrium measure. These recursive relations are sometimes called the *topological recursion* because, at least when $\beta = 2$, they mimic the relations between maps of different genus, see e.g. [EO] and allow to recover the result from 't Hooft and Brézin-Itzykson-Parisi-Zuber [BIPZ] showing that matrix integrals can be seen as generating functions of maps; for the general β recursion see [CE].

Loop equations were introduced to the mathematical community by Johansson in [J1] to derive the Gaussian behavior; his results were significantly extended in the later work [KS], [BoGu1], [Shch], [BoGu2]. The loop equations, or rather "their spirit", also have further applications far beyond the central limit theorems for global fluctuations, e.g. they were used in the recent work on *local universality* for random matrices, see [BEY], [BFG]. We would like to emphasize an important distinction between the approaches of physics and mathematics literature: the latter operates not with formal expansions, but with converging asymptotic expansions. This requires *a priori* estimates, whose derivations rely on a set of tools different from the loop equations themselves.

The Schwinger-Dyson equations for general-β log-gases are obtained by integration by parts and derivation with respect to the potential. In fact, they can be derived in a much more general continuous setting, for instance when the underlying measure is the Haar measure on a compact Lie group, see e.g. [AGZ, (5.4.29)], and further used to derive the topological asymptotic expansions of related matrix models, see [CGM], [GN]. For a recent application of the Schwinger-Dyson equations to the lattice gauge theory see [Cha].

1.2. *Integrable discretization of log-gases.* — The *discrete* versions of the distribution (1) (with x_i's living on a lattice) at $\beta = 2$ arise in numerous problems of $2d$ statistical mechanics. Examples include random tilings (cf. Figure 1 and [CLP], [J4], [G], [BKMM]), stochastic systems of non-intersecting paths (cf. [KOR], [BBDT]), last passage percolation (cf. [J2]), interacting particle systems (cf. [J3], [BF]).

The law of large numbers for such systems (even for general values of β) can be established by essentially the same methods as for the continuous ones, cf. [J3], [J4], [Fe]. However, the situation is drastically different for the study of the asymptotics for global fluctuations. While for some very specific *integrable* choices of the potential at $\beta = 2$ the central limit theorem was proven (cf. [BF], [P2], [BD]), until now no general approach that would work for generic β and V(x) existed. We note however the works [HO, Chapter 12], [DF], [Mo] where the global Gaussian asymptotics was proven for certain (very concrete) discrete general β probabilistic models originating in the representation theory. Moreover, it was not clear whether the CLT in the case of general potential should be the same as in the continuous case, or there should be some significant differences (indeed, for instance the *local* limits in the bulk must be different). The main technical difficulty lied in the absence of a nice generalization of the loop equations to the discrete setting.

In [Ey2], [Ey3], Eynard proposed to interpret laws of random partitions, which include discrete $\beta = 2$ analogues of (3), as matrix models. One of the key points in his
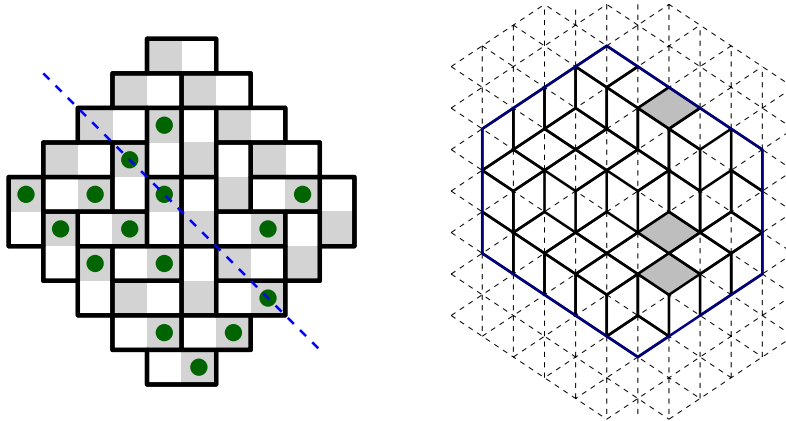
FIG. 1. — The discrete particle systems arising in uniformly random *domino tilings* of the Aztec diamond (*left panel*) and uniformly random *lozenge tilings* of a hexagon (*right panel*). The distributions of 3 particles on both pictures have the form (1) with N = 3, $\beta = 2$ and suitable (different) potentials V(x)

interpretation is to view discrete sums as highly oscillatory continuous integrals. Based on this identification, he conjectured that the asymptotic expansions in this case are described by the same Schwinger-Dyson equations with initial step given by the equilibrium measure of the model, see e.g. [Ey2, Section 2.4, Section 2.7.1]. In particular, the fluctuations in the context of the central limit theorem should be universal. However, so far this approach has not yet progressed much beyond the predictions.

The central goal of this article is to study global fluctuations of the empirical measure for discrete analogues of general-β log-gases. One outcome is that indeed these fluctuations are universal and described by the same covariance as for their continuous counterparts. Our analysis is based on appropriate discrete versions of the Schwinger-Dyson equations, which (unlike in the continuous setting) do not appear as a direct consequence of integration by parts or perturbative arguments.

The search for the discrete loop equations starts with identifying a good discrete analogue of the general β distribution (1). For that we fix a parameter $\theta > 0$ and a positive real-valued function $w(x; N)$.[1] Consider the probability distribution

$$(3) \qquad \mathbf{P}_N(\ell_1, \ldots, \ell_N) = \frac{1}{Z_N} \prod_{1 \leq i < j \leq N} \frac{\Gamma(\ell_j - \ell_i + 1)\Gamma(\ell_j - \ell_i + \theta)}{\Gamma(\ell_j - \ell_i)\Gamma(\ell_j - \ell_i + 1 - \theta)} \prod_{i=1}^{N} w(\ell_i; N)$$

on ordered N-tuples $\ell_1 < \ell_2 < \cdots < \ell_N$, such that $\ell_i = \lambda_i + \theta i$ and $\lambda_1 \leq \lambda_2 \cdots \leq \lambda_N$ are integers. We refer to ℓ_i's as the positions of N particles. Let us note that if $\theta \neq 1$, then ℓ_i do not sit on the fixed lattice.

_____

[1] $w(x; N)$ should decay at least as $|x|^{-(2N-2)\theta - 1 - \text{const}}$ with const > 0 as $|x| \to \infty$.

Note that if $\theta = 1$, then (3) has the same form as (1) with $\beta = 2$. Similarly, $\theta = 1/2$ leads to (1) with $\beta = 1$.[2] More generally, if we set $\ell_i = \mathrm{L}x_i$, then as $\mathrm{L} \to \infty$, the ratio of Gamma-functions in (3) behaves as $(\ell_j - \ell_i)^{2\theta}$ and mimics (1) with $\beta = 2\theta$.

The most important reason to view (3) as a correct *integrable* discretization of the continuous log-gas (1) is the following observation which is the starting point of the results in the present paper.

Theorem **1.1** *(Nekrasov's equation). — Consider the probability distribution* (3), *and assume that*

$$\frac{w(x; \mathrm{N})}{w(x - 1; \mathrm{N})} = \frac{\phi_{\mathrm{N}}^{+}(x)}{\phi_{\mathrm{N}}^{-}(x)}$$

and define

$$(\mathbf{4}) \qquad \mathrm{R_N}(\xi) = \phi_{\mathrm{N}}^{-}(\xi) \cdot \mathbf{E}_{\mathbf{P_N}} \left[\prod_{i=1}^{\mathrm{N}} \left(1 - \frac{\theta}{\xi - \ell_i} \right) \right]$$

$$+ \phi_{\mathrm{N}}^{+}(\xi) \cdot \mathbf{E}_{\mathbf{P_N}} \left[\prod_{i=1}^{\mathrm{N}} \left(1 + \frac{\theta}{\xi - \ell_i - 1} \right) \right].$$

If $\phi_{\mathrm{N}}^{\pm}(\xi)$ are holomorphic in a domain $\mathcal{M}_{\mathrm{N}} \subset \mathbf{C}$, then so is $\mathrm{R_N}(\xi)$. Moreover, if $\phi_{\mathrm{N}}^{\pm}(\xi)$ are polynomials of degree at most d, then so is $\mathrm{R_N}(\xi)$.

Theorem 1.1 is essentially due to Nekrasov and his collaborators, as it is a variant of similar statements in [N], [NP], [NS]. For the proof see Theorem 4.1 below.

It is reasonable to ask how one could guess the form of the measure (3), which would lead to Theorem 1.1. The integrability properties of such measures, including the product of the ratios of Gamma functions as in (3), can be traced to their connections to representation theory and symmetric functions. The same products of Gamma functions appear in the evaluation formulas for Jack symmetric polynomials (cf. [Ma, Chapter VI, Section 10]) and in problems of asymptotic representation theory (cf. [O2]). Another trace of integrability is the existence of *discrete Selberg integrals*, which are evaluation formulas for the partition function $\mathrm{Z_N}$ in (3) for special choices of the weight $w(x; \mathrm{N})$. Examples of such evaluations can be found in [O2, Section 2], [GS, Section 2.2]. The latter reference also explains the degeneration to the conventional Selberg integral, which is the computation of the normalization constants for continuous log-gases (1) with specific choices of the potential $\mathrm{V}(x)$, see [Me, Chapter 17], [Fo, Chapter 4]. On the other hand, for a "naive" discrete version, when one takes the same formula (1) as a definition of a discrete log-gas, we are not aware of the existence of similar evaluation formulas (outside $\beta = 2$, when (1) and (3) coincide).

---

[2] However, observe that the lattice where particles sit at $\theta = 1/2$ is not $\mathbf{Z}$.

1.3. *Global fluctuations of discrete log-gases: one cut case.* — We proceed to our results on global fluctuations for the distributions (3) as $N \to \infty$. For that we need to postulate how the weight $w(x; N)$ changes as $N \to \infty$. Our methods work for a variety of possibilities here, but to simplify the exposition we stick to the following assumption in this section:

$$(\mathbf{5}) \qquad w(x; N) = \exp\left(-NV\left(\frac{x}{N}\right)\right),$$

where $V(z)$ is an analytic function of real argument z such that for $|z|$ large enough, $V(z)$ is monotone and satisfies

$$(\mathbf{6}) \qquad V(z) > c\ln(|z|), \quad \text{where } c > 2\theta.$$

Our first result is the law of large numbers for the empirical measures μ_N defined via

$$\mu_N = \frac{1}{N}\sum_{i=1}^{N}\delta_{\ell_i/N}, \quad (\ell_1, \ldots, \ell_N) \text{ is } \mathbf{P}_N\text{-distributed.}$$

Theorem **1.2.** — *There exists a deterministic absolutely continuous compactly supported probability measure $\mu(x)dx$ with $0 \le \mu(x) \le \theta^{-1}$, such that μ_N converges to $\mu(x)dx$ as $N \to \infty$, in the sense that for any compactly supported Lipshitz function $f(x)$ the following convergence in probability holds:*

$$\lim_{N\to\infty}\int_{\mathbf{R}}f(x)\mu_N(dx) = \int_{\mathbf{R}}f(x)\mu(x)dx.$$

In fact, we prove a more general statement where, in particular, $V(x)$ does not have to be analytic, see Theorems 5.3 and 10.1 below. The measure $\mu(x)dx$ is the equilibrium measure, and it can be found as a solution to a variational problem. At $\theta = 1$, Theorem 5.3 reduces to results of [J3, J4, Fe]. For general values of θ, additional arguments are required, and we present them.

Note the condition $\mu(x) \le \theta^{-1}$, which is not present in the continuous log-gases, and arises from the fact that the minimal distance between adjacent particles in (3) is at least θ. This is a specific feature of the discrete models.

At this moment we need to make certain assumptions on the equilibrium measure. A *band*[3] of $\mu(x)$ is a maximal interval (α, β) such that $0 < \mu(x) < \theta^{-1}$ on (α, β). From the random matrix literature (cf. [BDE], [Shch], [BoGu2]) one expects that the global fluctuations are qualitatively different depending on whether $\mu(x)$ has one or more bands. Here we stick to the *one band* case.

---

[3] We follow the terminology from [BKMM].

Theorem **1.3.** — *Assume* (5), (6), *that $\mu(x)$ has a unique band (a_-, a_+), and that (technical) Assumption 4 from Section 3.2 holds. Take any $m \geq 1$ bounded analytic functions $f_1, \ldots, f_m$ on **R**. Then the m random variables*

$$\mathcal{L}_{f_j} = \sum_{i=1}^{N} \big(f_j(\ell_i) - \mathbf{E}_{\mathbf{P}_N} f_j(\ell_i) \big), \quad (\ell_1, \ldots, \ell_N) \text{ is } \mathbf{P}_N\text{-distributed,}$$

converge (in distribution and in the sense of moments) to centered Gaussian random variables with explicit covariance depending on $a_\pm$ and θ, and given in (127) *below. In particular, if $f_i(x) = (z_i - x)^{-1}$, $z_i \in \mathbf{C} \setminus \mathbf{R}$, $i = 1, \ldots, m$, then*

$$(\mathbf{7}) \qquad \lim_{N \to \infty} \mathbf{E}_{\mathbf{P}_N} \mathcal{L}_{f_i} \mathcal{L}_{f_j} = -\frac{\theta^{-1}}{2(z_i - z_j)^2}$$

$$\times \left(1 - \frac{z_i z_j - \frac{1}{2}(a_- + a_+)(z_+ z_j) + a_+ a_-}{\sqrt{(z_i - a_-)(z_i - a_+)}\sqrt{(z_j - a_-)(z_j - a_+)}} \right),$$

Remark **1.4.** — We prove below in Theorem 10.1 that with exponentially high probability all particles ℓ_i are inside an interval $[-DN, DN]$. Further, it is enough to assume in Theorem 1.3 that f_i are analytic only in $[-D, D]$. Moreover, we believe (but do not prove) that the analyticity assumption can be replaced by sufficient smoothness.

Remark **1.5.** — The second order asymptotic expansion of the *mean* $\mathbf{E}_{\mathbf{P}_N} (\sum_{i=1}^{N} f_j(\ell_i))$ can be also analyzed, cf. Theorem 6.1 with $k = 1$, $m = 0$.

The proof of Theorem 1.3 is a combination of Theorem 7.1 and Theorem 10.1 in the main text.

The technical Assumption 4 from Section 3.2 is a statement that a certain function produced from the equilibrium measure $\mu(x)dx$ has no zeros. In Section 9.3 we show that this assumption always holds when $V(x)$ is a convex function. A somewhat similar assumption appears in the work of Johansson [J1] on the central limit theorem for global fluctuations of continuous log-gases.

Let us emphasize that the limiting covariance in Theorem 1.3 depends only on the support of the equilibrium measure, but it is not sensitive to other features. The same phenomenon is known in the random matrix setting. Moreover, comparing (7) with expressions in [J1, Theorem 4.2], [PS, Chapter 3], we conclude that the covariance is *precisely the same* as for the continuous log-gases. Thus, the discreteness of the model is invisible on the level of the central limit theorem.

The covariance in Theorem 1.3 can be related to that of a section of the $2d$ Gaussian Free Field in the upper half-plane with Dirichlet boundary conditions. One way to predict that is by noticing that sections in lozenge tilings models with several specific boundary conditions yield distributions of the form (3) with $\theta = 1$, see Section 9.2 for one

example. On the other hand, there exist several results on the appearance of the Gaussian Free Field in the asymptotics for global fluctuations in lozenge tilings, cf. [K], [BF], [P2], [BuGo2].

Subsequent work [KS], [BoGu1] for continuous log-gases showed that Johansson's technical assumption on the equilibrium measure can be relaxed and replaced by certain weaker assumptions, which hold for *generic* analytic potentials. We hope that a similar thing can also be done in the present discrete setting, but this would require further investigations.

1.4. *Weight supported on several finite intervals.* — In several applications the weight $w(x; N)$ is not defined on the whole real line, but instead it is supported by a union of several disjoint intervals $(a_i(N), b_i(N))$, $i = 1, \ldots, k$. In other words, the particles ℓ_i are now confined to the union of these intervals $\bigcup_{i=1}^{N} (a_i(N), b_i(N))$. For instance, this happens in tilings models, see Section 9.2.

At this point we have an additional choice: One could either *fix* the filling fractions $n_i(N)$, which are the numbers of particles ℓ_i in each of the intervals $(a_i(N), b_i(N))$, or not. In the present paper we stick to the former and fix $n_i(N)$.

We refer to Sections 3 and 5 below for details of the exact assumptions that we impose on the weight $w(x; N)$, intervals $(a_i(N), b_i(N))$, and filling fractions $n_i(N)$. Here we will only briefly summarize the results obtained in such a framework.

The first result is the law of large numbers, which is the exact analogue of Theorem 1.2, see Theorem 5.3 for the details. As before, for the asymptotics of the global fluctuations we need to have some information about the bands of the equilibrium measure $\mu(x)dx$. We assume that there is *one band per interval*.

Theorem **1.6** *(Theorem 7.1).* — *Assume that all the data specifying the model satisfies Assumptions 1–5 in Section 3.2 below. In particular, the equilibrium measure $\mu(x)dx$ has k bands (α_i, β_i), $i = 1, \ldots, k$, one per interval of the support of the model. Take $m \geq 1$ functions $f_1(z), \ldots, f_m(z)$, which are analytic in a neighborhood of $\bigcup_{i=1}^{k} [a_i(N)/N, b_i(N)/N]$ for large N. Then as $N \to \infty$ the joint moments of the m random variables*

$$\mathcal{L}_{f_j} = \sum_{i=1}^{N} \big(f_j(\ell_i) - \mathbf{E}_{\mathbf{P}_N} f_j(\ell_i) \big), \quad (\ell_1, \ldots, \ell_N) \text{ is } \mathbf{P}_N\text{-distributed,}$$

approximate those of centered Gaussian random variables with asymptotic covariance depending only on $\theta, \alpha_1, \ldots, \alpha_k, \beta_1, \ldots, \beta_k$ and given by (116).

For $k = 1$, the limiting covariance coincides with that of Theorem 1.3. For $k > 1$ it is no longer expressible through elementary functions, e.g. for $k = 2$ elliptic functions appear in the formulas. We again emphasize that the covariance depend only on the end-points of the bands and is not sensitive to other features of the equilibrium measure. Furthermore, the covariance is the same as for the continuous log-gases, cf. [Shch], [BoGu2].

Similarly to the $k = 1$ case, one could try to identify the covariance with that of a canonically defined random field. We believe that there should be a link to sections of the Gaussian Free Field in a domain with $k - 1$ holes, but we postpone this discussion till a future publication.

1.5. *Organization of the paper.* — In Section 2 we present our approach and results: Nekrasov's equation, Law of Large Numbers and Central Limit Theorem—in the simplest yet non-trivial case when $\theta = 1$ and $\phi_N^{\pm}$ in Theorem 1.1 are linear functions. The measure $\mathbf{P}_N$ in this case is known as the Krawtchouk orthogonal polynomial ensemble; it appeared and was studied in numerous previous articles. In particular, the Central Limit Theorem for global fluctuations in this specific case can be also obtained by other methods: see [CJY], [BD], and [BuGo2].[4]

In Sections 3–8 we explain a much more general framework, in which the same ideas work and lead to the Law of Large Numbers and Central Limit Theorem. In these sections θ is an arbitrary positive number, and the weight $w(x; N)$ and functions $\phi_N^{\pm}$ are general (subject to certain technical assumptions).

In Section 9 we specialize the general framework to certain specific examples, which include lozenge tilings and (z, w)-measures from asymptotic representation theory. We explain how all the technical assumptions are checked in all these examples.

Finally, Section 10 explains how the case of the infinite support of $w(x; N)$ can be reduced to the case of the bounded support (which is studied in Sections 3–8) through large deviations estimates.

2. Toy example: $\theta = 1$ binomial weight

The aim of this section is to describe our method in the simplest, yet non-trivial case of the binomial weight and $\theta = 1$. The resulting N-particle ensemble is known as the *Krawtchouk orthogonal polynomial ensemble*. It appeared in the literature before in the connection with uniformly random domino tilings of the Aztec diamond (cf. [J4]), last passage percolation (cf. [J2]), stochastic systems of non-intersecting paths (cf. [KOR], [BBDT]), and with representation theory of the infinite-dimensional unitary group $U(\infty)$ (cf. [BO2, Section 5], [B1, Section 4]).

In the subsequent sections we will extend our approach to more general cases, but the methodology and many key ideas remain the same.

Fix two integers $0 < N \leq M$ and consider the space $\mathbf{W}_N$ of N-tuples of integers

$$0 \leq \ell_1 < \ell_2 < \cdots < \ell_N \leq M.$$

---

[4] We remark that while the approaches of these articles admit various generalizations, to the best of our knowledge, they do not extend further in the direction of the present paper.

We define a probability distribution $\mathbf{P}_N$ on $\mathbf{W}_N$ through

$$(8) \qquad \mathbf{P}_N(\ell_1, \ldots, \ell_N) = \frac{1}{Z(N, M)} \prod_{1 \le i < j \le N} (\ell_i - \ell_j)^2 \prod_{i=1}^{N} \binom{M}{\ell_i}.$$

We remark that the partition function $Z(N, M)$ is explicitly known in this case:

$$Z(N, M) = 2^{N(M-N+1)}(M!)^N \prod_{j=0}^{N-1} \frac{j!}{(M-j)!}.$$

However, in a generic situation there are typically no simple formulas for the partition function and we are not going to use its explicit form.

Our analysis of the distribution $\mathbf{P}_N$ is based on the following $\theta = 1$ version of Theorem 4.1. This is essentially due to [N], [NP], [NS] and we will call the main statement the *Nekrasov's equation*. It should be viewed as a discrete space analogue of the Schwinger-Dyson (also known as "loop") equations.

Proposition **2.1.** *— Let* $\mathbf{P}_N$ *be a probability distribution on* N*-tuples* $0 \le \ell_1 < \cdots < \ell_N \le M$ *such that*

$$\mathbf{P}_N(\ell_1, \ldots, \ell_N) = \frac{1}{Z} \prod_{1 \le i < j \le N} (\ell_i - \ell_j)^2 \prod_{i=1}^{N} w(\ell_i),$$

where

$$\frac{w(x)}{w(x-1)} = \frac{\phi_N^+(x)}{\phi_N^-(x)},$$

and $\phi_N^\pm(x)$ *are analytic functions in a complex neighborhood of* $[-1, M+1]$ *which are positive on* $[0, M]$ *and satisfy* $\phi_N^+(M+1) = \phi_N^-(0) = 0$. *Define*

$$(9) \qquad R_N(\xi) = \phi_N^-(\xi) \cdot \mathbf{E}_{\mathbf{P}_N}\left[\prod_{i=1}^{N} \left(1 - \frac{1}{\xi - \ell_i} \right) \right]$$

$$+ \phi_N^+(\xi) \cdot \mathbf{E}_{\mathbf{P}_N}\left[\prod_{i=1}^{N} \left(1 + \frac{1}{\xi - \ell_i - 1} \right) \right].$$

Then $R_N(\xi)$ *is analytic in the same complex neighborhood of* $[-1, M+1]$. *If* $\phi_N^\pm(\xi)$ *are polynomials of degree at most* d, *then* $R_N(\xi)$ *is also a polynomial of degree at most* d.

Proof. — The possible singularities of $R_N(\xi)$ are simple poles at points $m = 0, 1, 2, \ldots, M$. Let us compute a residue at such a point.

The expectation $\mathbf{E}_{\mathbf{P}_N}$ in (9) is a sum over all possible configurations $(\ell_1, \ldots, \ell_N) \in \mathbf{W}_N$. Such a configuration contributes to the residue at m if either $\ell_i = m$ or $\ell_i + 1 = m$ for some $i = 1, \ldots, N$.

We separately analyze the contributions appearing from each $i = 1, \ldots, N$, which we now fix. Given a particle configuration $\ell = (\ell_1, \ldots, \ell_N)$, let ℓ^+ denote the configuration with ith coordinate increased by 1 and let ℓ^- denote the configuration with ith coordinate decreased by 1. Note that, in principle ℓ^+ (similarly ℓ^-) might fail to be in $\mathbf{W}_N$, as the coordinates might coincide. However, in this case the formula for $\mathbf{P}_N(\ell^+)$ still applies and gives zero.[5]

The contribution to the residue of R_N at $m \in \{0, \ldots, M\}$, arising from the ith coordinate of $(\ell_1, \ldots, \ell_N)$ is

$$(\mathbf{10}) \qquad - \sum_{\ell \in \mathbf{W}_N | \ell_i = m} \phi_N^-(m) \mathbf{P}_N(\ell_1, \ldots, \ell_N) \prod_{j \neq i} \left(1 - \frac{1}{m - \ell_j} \right)$$

$$+ \sum_{\ell \in \mathbf{W}_N | \ell_i = m-1} \phi_N^+(m) \mathbf{P}_N(\ell_1, \ldots, \ell_N) \prod_{j \neq i} \left(1 + \frac{1}{m - \ell_j - 1} \right).$$

But using the definition of $\mathbf{P}_N$ we see that

$$\phi_N^-(m) \mathbf{P}_N(\ell_1, \ldots \ell_{i-1}, m, \ell_{i+1}, \ldots, \ell_N) \prod_{j \neq i} \left(1 - \frac{1}{m - \ell_j} \right)$$

$$= \phi_N^+(m) \mathbf{P}_N(\ell_1, \ldots \ell_{i-1}, m-1, \ell_{i+1}, \ldots, \ell_N) \prod_{j \neq i} \left(1 + \frac{1}{m - \ell_j - 1} \right).$$

We conclude that for each μ, the terms with $\ell = \mu$ and $\ell = \mu^+$ (or $\ell = \mu^-$ and $\ell = \mu$) in the first and second sum in (10) cancel out and the total residue is zero.

For the polynomiality statement it suffices to notice that if $\phi_N^\pm(\xi)$ are polynomials of degree at most d, then $R_N(\xi)$ is an entire function which grows as $O(\xi^d)$ as $\xi \to \infty$. Hence, by Liouville's theorem $R_N(\xi)$ is a polynomial. $\qquad\square$

Note that for the distribution (8) the functions $\phi_N^\pm(x)$ can be chosen to be linear and we set

$$(\mathbf{11}) \qquad \phi_N^-(x) = \frac{x}{N}, \qquad \phi_N^+(x) = \frac{M+1}{N} - \frac{x}{N}.$$

Thus, in a sense, (8) is one of the simplest possible distribution in the framework of Proposition 2.1.

In this section we aim to study the asymptotics of the distributions $\mathbf{P}_N$ as $N \to \infty$. The parameter M will also depend on N. We fix $\mathfrak{m} > 1$ and set $M = \lfloor \mathfrak{m} N \rfloor$.

---

[5] This is where we need the condition of $\phi_N^+(M+1) = \phi_N^-(0) = 0$, which translates into $w(M+1) = w(-1) = 0$.

2.1. *Law of large numbers.* — As $N \to \infty$, a certain law of large numbers holds for the measures $\mathbf{P}_N$. Let us introduce a random probability measure μ_N on $\mathbf{R}$ via

(**12**) $$\mu_N = \frac{1}{N} \sum_{i=1}^{N} \delta\left(\frac{\ell_i}{N}\right), \quad (\ell_1, \ldots, \ell_N) \text{ is } \mathbf{P}_N\text{-distributed}.$$

The measure μ_N is often referred to as the *empirical measure* of the point configuration $\ell_1, \ldots, \ell_N$. Note that our definitions imply the condition $\ell_{i+1} - \ell_i \geq 1$, which shows that for any interval $[p, q]$, its μ_N-measure is bounded from above by $(q - p + 1/N)$.

The idea for the proof of the law of large numbers for μ_N as $N \to \infty$ is to establish the large deviations principle for the measures $\mathbf{P}_N$, which would show that the measure is concentrated on ℓ_i which maximize the probability density (8). This was done rigorously in [J3], [J4], [Fe], see also Section 2.4 below for some details. The explicit formula for $\lim_{N \to \infty} \mu_N$ is then obtained as a solution of a variational problem, this solution for our weight was found earlier in [DS, Example 4.2]. The developments of these articles are summarized in the following proposition.

Proposition **2.2.** — *The measures μ_N converge (weakly, in probability) to a deterministic absolutely continuous measure $\mu_{\mathfrak{m}}(x)dx$, which is called* equilibrium measure. *For $\mathfrak{m} \geq 2$, the density of the measure $\mu_{\mathfrak{m}}(x)dx$ is*

$$\mu_{\mathfrak{m}}(x) = \begin{cases} \dfrac{1}{\pi} \operatorname{arccot}\left(\dfrac{\mathfrak{m} - 2}{2\sqrt{\mathfrak{m} - 1 - (x - \mathfrak{m}/2)^2}}\right), & |x - \tfrac{\mathfrak{m}}{2}| < \sqrt{\mathfrak{m} - 1}, \\ 0, & |x - \tfrac{\mathfrak{m}}{2}| \geq \sqrt{\mathfrak{m} - 1}, \end{cases}$$

and for $1 < \mathfrak{m} \leq 2$,

$$\mu_{\mathfrak{m}}(x) = \begin{cases} \dfrac{1}{\pi} \operatorname{arccot}\left(\dfrac{\mathfrak{m} - 2}{2\sqrt{\mathfrak{m} - 1 - (x - \mathfrak{m}/2)^2}}\right), & |x - \tfrac{\mathfrak{m}}{2}| < \sqrt{\mathfrak{m} - 1}, \\ 1, & \tfrac{\mathfrak{m}}{2} \geq |x - \tfrac{\mathfrak{m}}{2}| \geq \sqrt{\mathfrak{m} - 1}, \\ 0, & |x - \tfrac{\mathfrak{m}}{2}| \geq \tfrac{\mathfrak{m}}{2}, \end{cases}$$

where $\operatorname{arccot}(y)$ *is the inverse cotangent function.*

A convenient way of working with the equilibrium measure $\mu_{\mathfrak{m}}(x)$ is through its Stieltjes transform $G_{\mathfrak{m}}(z)$ defined through

(**13**) $$G_{\mathfrak{m}}(z) = \int_{-\infty}^{\infty} \frac{\mu_{\mathfrak{m}}(x)dx}{z - x}.$$

Observe that (13) makes sense for all z outside the support of $\mu_{\mathfrak{m}}(x)$, and $G_{\mathfrak{m}}(z)$ is holomorphic there. Further, as $z \to \infty$, we have $G_{\mu}(z) \sim z^{-1}$. An explicit formula for $G_{\mathfrak{m}}(z)$ can be readily extracted from Proposition 2.3 below.

We introduce two functions

$$R_\mathfrak{m}(z) = z\exp(-G_\mathfrak{m}(z)) + (\mathfrak{m} - z)\exp(G_\mathfrak{m}(z)),$$
$$Q_\mathfrak{m}(z) = z\exp(-G_\mathfrak{m}(z)) - (\mathfrak{m} - z)\exp(G_\mathfrak{m}(z)),$$

for $z \in \mathbf{C} \setminus [0, \mathfrak{m}]$ and consider their analytic continuations.

Proposition **2.3.** — *For any* $\mathfrak{m} > 1$,

$$R_\mathfrak{m}(z) = \mathfrak{m} - 2, \qquad Q_\mathfrak{m}(z) = 2\sqrt{(z - \mathfrak{m}/2)^2 - (\mathfrak{m} - 1)},$$

with the branch of the square root chosen so that $Q_\mathfrak{m}(z) \sim 2z$ *as* $z \to \infty$.

Proof. — We first observe that $R_\mathfrak{m}(z)$ is, in fact, a linear polynomial. Indeed, this is the $N \to \infty$ limit of $R_N(Nz)$ in Proposition 2.1 for the measure $\mathbf{P}_N$. The fact that $G_\mathfrak{m}(z) \sim z^{-1}$ as $z \to \infty$, implies $\lim_{z\to\infty} R_\mathfrak{m}(z) = \mathfrak{m} - 2$, and thus $R_\mathfrak{m}(z) = \mathfrak{m} - 2$. Further, observe that the definition of $Q_\mathfrak{m}$ and $R_\mathfrak{m}$ implies

$$Q_\mathfrak{m}(z)^2 - R_\mathfrak{m}(z)^2 = -4z(\mathfrak{m} - z).$$

Therefore,

$$Q_\mathfrak{m}(z)^2 = 4z^2 - 4z\mathfrak{m} + (\mathfrak{m} - 2)^2 = 4(z - \mathfrak{m}/2)^2 - 4(\mathfrak{m} - 1). \qquad \square$$

2.2. *Second order expansion.* — Define the Stieltjes transform $G_N(z)$ of the prelimit empirical measure (12) through

$$(\mathbf{14}) \qquad G_N(z) = \int_{-\infty}^{\infty} \frac{1}{z - x}\mu_N(dx) = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{z - \ell_i/N},$$

$(\ell_1, \ldots, \ell_N)$ is $\mathbf{P}_N$-distributed.

We aim to study how $G_N(z)$ approximate $G_\mathfrak{m}(z)$ as $N \to \infty$. For that we introduce a deformed version of the same function. Take $2k$ parameters $\mathbf{t} = (t_1, \ldots, t_k)$, $\mathbf{v} = (v_1, \ldots, v_k)$ such that $v_a + t_a - y \neq 0$ for all $a = 1, \ldots, k$ and all $y \in [0, \mathfrak{m}]$ and let the deformed distribution $\mathbf{P}_N^{\mathbf{t},\mathbf{v}}$ be defined through

$$(\mathbf{15}) \qquad \mathbf{P}_N^{\mathbf{t},\mathbf{v}}(\ell_1, \ldots, \ell_N)$$

$$= \frac{1}{Z(N, M; \mathbf{t}, \mathbf{v})}\prod_{1\leq i<j\leq N}(\ell_i - \ell_j)^2\prod_{i=1}^{N}\left[\binom{M}{\ell_i}\prod_{a=1}^{k}\left(1 + \frac{t_a}{v_a - \ell_i/N}\right)\right].$$

If $k = 0$, then $\mathbf{P}_N^{\mathbf{t},\mathbf{v}} = \mathbf{P}_N$ is the undeformed measure. In general, $\mathbf{P}_N^{\mathbf{t},\mathbf{v}}$ may be a complex-valued measure. The normalizing constant $Z(N, M; \mathbf{t}, \mathbf{v})$ in (15) is chosen so that the

total mass of $\mathbf{P}_N^{\mathbf{t,v}}$ is 1, i.e. $\sum_{\ell \in \mathbf{W}_N} \mathbf{P}_N^{\mathbf{t,v}}(\ell) = 1$. Let us note that the numbers t_a are always chosen small enough, which guarantees that $Z(N, M; t, v) \neq 0$.

Observe that $\mathbf{P}_N^{\mathbf{t,v}}$ satisfies the assumptions of Proposition 2.1 with

$$(\mathbf{16}) \qquad \phi_N^+(x) = \left(\frac{M+1}{N} - \frac{x}{N} \right) \cdot \prod_{a=1}^{k} \left(v_a - \frac{x}{N} + \frac{1}{N} \right) \left(t_a + v_a - \frac{x}{N} \right),$$

$$(\mathbf{17}) \qquad \phi_N^-(x) = \frac{x}{N} \cdot \prod_{a=1}^{k} \left(v_a - \frac{x}{N} \right) \left(t_a + v_a - \frac{x}{N} + \frac{1}{N} \right).$$

As above, we set $M = \lfloor \mathfrak{m} N \rfloor$, and omit it from the notations. We also define

$$\psi_N^+(z) = \left(\lfloor \mathfrak{m} N \rfloor + 1 - \mathfrak{m} N \right) + \sum_{a=1}^{k} \frac{\mathfrak{m} - z}{v_a - z}, \qquad \psi_N^-(z) = \sum_{a=1}^{k} \frac{z}{t_a + v_a - z}.$$

Clearly,

$$\phi_N^+(Nz) = \left(\mathfrak{m} - z + \frac{\psi_N^+(z)}{N} + O\left(\frac{1}{N^2} \right) \right) \cdot \prod_{a=1}^{k} (v_a - z)(t_a + v_a - z),$$

$$\phi_N^-(Nz) = \left(z + \frac{\psi_N^-(z)}{N} + O\left(\frac{1}{N^2} \right) \right) \cdot \prod_{a=1}^{k} (v_a - z)(t_a + v_a - z).$$

We further define $\mu_N^{\mathbf{t,v}}$ as the empirical distribution of $\mathbf{P}_N^{\mathbf{t,v}}$ and set

$$(\mathbf{18}) \qquad G_N(z) = \int_{-\infty}^{\infty} \frac{1}{z-x} \mu_N^{\mathbf{t,v}}(dx) = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{z - \ell_i/N},$$

$$(\ell_1, \dots, \ell_N) \text{ is } \mathbf{P}_N^{\mathbf{t,v}}\text{-distributed}.$$

Define

$$(\mathbf{19}) \qquad \Delta G_N(z) = N\big(G_N(z) - G_{\mathfrak{m}}(z) \big).$$

Note that we often omit the dependence on $\mathbf{t}, \mathbf{v}$ from the notations, to keep them concise.

The definition of the deformed measure $\mathbf{P}_N^{\mathbf{t,v}}$ is motivated by the following observation. It was used before in the related context in random matrix theory, cf. [Mi], [Ey1].

Lemma 2.4. — *For any $k \geq 1$, the kth mixed derivative*

$$(\mathbf{20}) \qquad \left. \frac{\partial^k}{\partial t_1 \partial t_2 \dots \partial t_k} \mathbf{E}_{\mathbf{P}_N^{\mathbf{t,v}}}\big(\Delta G_N(u) \big) \right|_{t_a = 0, 1 \leq a \leq k}$$

is the joint cumulant of $k+1$ random variables $NG_N(u), NG_N(v_1), \dots, NG_N(v_k)$ with respect to the measure $\mathbf{P}_N$.

Proof. — We first note that the joint cumulants are invariant under addition of constants. Therefore, we can replace $\mathrm{NG_N}(u)$ by $\Delta\mathrm{G_N}(u)$ for the purpose of computing the cumulants.

Further recall that one way to define the joint cumulant of $k+1$ bounded random variables $\xi_0, \ldots, \xi_k$ is through

$$\frac{\partial^{k+1}}{\partial t_0 \partial t_1 \ldots \partial t_k} \ln\left(\mathbf{E}\exp\left(\sum_{i=0}^{k} t_i \xi_i\right)\right)\Bigg|_{t_0=t_1=\cdots=t_k=0}.$$

Taking the derivative with respect to t_0 explicitly, we can rewrite this also as

$$\frac{\partial^k}{\partial t_1 \ldots \partial t_k} \frac{\mathbf{E}(\xi_0 \exp(\sum_{i=1}^{k} t_i \xi_i))}{\mathbf{E}(\exp(\sum_{i=1}^{k} t_i \xi_i))}\Bigg|_{t_1=\cdots=t_k=0}.$$

Setting $\xi_0 = \mathrm{N}\Delta\mathrm{G_N}(u)$, $\xi_i = \mathrm{NG_N}(v_i)$, $i = 1, \ldots, k$ and observing that

$$\exp\big(t\mathrm{NG_N}(v)\big) = \prod_{i=1}^{\mathrm{N}}\left(1 + \frac{t}{v - \ell_i/\mathrm{N}}\right) + \mathrm{O}\big(t^2\big), \quad t \to 0,$$

we get the desired statement. $\square$

Theorem **2.5.** — *Fix $k = 0, 1, \ldots$ and choose k complex numbers $\mathbf{v} = (v_1, \ldots, v_k) \in (\mathbf{C} \setminus [0, \mathfrak{m}])^k$. Then, as $\mathrm{N} \to \infty$,*

$$(\mathbf{21}) \qquad \frac{\partial^k}{\partial t_1 \cdots \partial t_k} \mathbf{E}_{\mathbf{P}_\mathrm{N}^{\mathbf{t},\mathbf{v}}}\big(\Delta\mathrm{G_N}(u)\big)\Bigg|_{t_a=0, 1\leq a\leq k}$$

$$= o(1) + \frac{\partial^k}{\partial t_1 \cdots \partial t_k}\left[\frac{1}{2\pi\mathbf{i}\cdot 2\sqrt{(u-\mathfrak{m}/2)^2 - (\mathfrak{m}-1)}}\right.$$

$$\times \oint_{\gamma_{[0,\mathfrak{m}]}} \frac{dz}{(u-z)}\cdot\left(\psi_\mathrm{N}^-(z)e^{-\mathrm{G}_\mathfrak{m}(z)} + \psi_\mathrm{N}^+(z)e^{\mathrm{G}_\mathfrak{m}(z)}\right.$$

$$\left.\left. + \sqrt{(z-\mathfrak{m}/2)^2 - (\mathfrak{m}-1)}\cdot\frac{\partial}{\partial z}\mathrm{G}_\mathfrak{m}(z)\right)\right]_{t_a=0, 1\leq a\leq k},$$

where $\gamma_{[0,\mathfrak{m}]}$ is a simple positively-oriented contour enclosing the segment $[0, \mathfrak{m}]$ (the points u and $v_1, \ldots, v_k$ are outside the contour). The remainder $o(1)$ is uniform over $u, v_1, \ldots, v_k$ in compact subsets of the unbounded component of $\mathbf{C} \setminus \gamma_{[0,\mathfrak{m}]}$. The $k = 0$ case is that we take no derivatives in (21).

Remark **2.6.** — The only cases, where the right-hand side of (21) is meaningful are $k = 0, 1$, since its $\mathrm{N} \to \infty$ limit is zero for $k \geq 2$. Indeed, only ψ_N^- depends on t_a, and moreover it is a sum of functions of single variables t_a, $a = 1, \ldots, k$; therefore, all mixed

partial derivatives vanish. However, we present Theorem 2.5 in this way, since that's the form which appears in our proofs. For $k = 1$ we will compute the limit of (21) below in Theorem 2.8. For $k = 0$ and generic $\mathfrak{m}$ the right-hand side of (21) has no limit, as $(\lfloor \mathfrak{m}N \rfloor + 1 - \mathfrak{m}N)$ in the definition of ψ_N^+ oscillates.

Remark **2.7.** — It might seem a bit unexpected that for $k = 0$ the right-hand side of (21) does not vanish when $\theta = 1$. Indeed, in the context of random matrices this corresponds to $\beta = 2$ case, where the mean is known to vanish (cf. [J1]). For our model the non-zero mean can be traced back to two features. First, due to discreteness we can not adjust $M = \lfloor \mathfrak{m}N \rfloor$ so that $(\lfloor \mathfrak{m}N \rfloor + 1 - \mathfrak{m}N)$ becomes zero for all N. Second, the logarithm of the weight, $\ln \left(\binom{M}{x} \right)$, has a non-trivial $N \to \infty$ asymptotic expansion, which affects the result.

Before providing a proof of Theorem 2.5, let us give its important corollaries. Note that if we set $k = 0$ in Theorem 2.5, then it gives the limit behavior for $\mathbf{E}_{\mathbf{P}_N} \Delta G_N(z)$. In fact, it also gives the Central Limit Theorem for $G_N(z)$.

Theorem **2.8.** — *The random field* $N(G_N(z) - \mathbf{E}_{\mathbf{P}_N} G_N(z))$, $z \in \mathbf{C} \setminus [0, \mathfrak{m}]$, *converges as* $N \to \infty$ *(in the sense of joint moments, uniformly in z in compact subsets of* $\mathbf{C} \setminus [0, \mathfrak{m}]$*) to a centered complex Gaussian random field with second moment*

$$(22) \qquad \lim_{N \to \infty} N^2 \big(\mathbf{E}_{\mathbf{P}_N} \big(G_N(u) G_N(v) \big) - \mathbf{E}_{\mathbf{P}_N} G_N(u) \mathbf{E}_{\mathbf{P}_N} G_N(v) \big) =: \mathcal{C}(u, v)$$

$$= -\frac{1}{2(u-v)^2} \left(1 - \frac{uv - \frac{1}{2}(a_- + a_+)(u + v) + a_+ a_-}{\sqrt{(u - a_-)(u - a_+)} \sqrt{(v - a_-)(v - a_+)}} \right),$$

where $a_\pm = \frac{\mathfrak{m}}{2} \pm \sqrt{\mathfrak{m} - 1}$.

Remark **2.9.** — Note that the covariance $\mathcal{C}(u, v)$ has no singularity at $u = v$, since the right-hand side of (22) has a finite $u \to v$ limit.

Remark **2.10.** — Since $\overline{G_N(u)} = G_N(\overline{u})$, the formula (22) is sufficient for determining the asymptotic covariance of the random field $G_N(u)$.

Proof of Theorem 2.8. — We start from the result of Lemma 2.4 for the joint cumulant of $k + 1$ random variables $NG_N(u), NG_N(v_1), \ldots, NG_N(v_k)$. In particular, if $k = 1$, then we get the covariance of $NG_N(u)$ and $NG_N(v_1)$.

For $k > 1$, differentiating (21) we see that the result vanishes as $N \to \infty$, see Remark 2.6. This implies the asymptotic Gaussianity of the random field $N(G_N(z) - \mathbf{E}_{\mathbf{P}_N} G_N(z))$.

In the case $k = 1$, differentiating (21) we see that the covariance given by (20) is (recall that u and v_1 lie outside the integration contour)

$$(23) \qquad \frac{1}{2\pi\mathbf{i} \cdot 2\sqrt{(u - \mathfrak{m}/2)^2 - (\mathfrak{m} - 1)}} \oint_{\gamma_{[0,\mathfrak{m}]}} \frac{dz}{(u - z)} \cdot \left(-\frac{z}{(v_1 - z)^2} e^{-G_{\mathfrak{m}}(z)} \right)$$

$$= \frac{1}{2\pi\mathbf{i} \cdot 4\sqrt{(u - \mathfrak{m}/2)^2 - (\mathfrak{m} - 1)}} \oint_{\gamma_{[0,\mathfrak{m}]}} \frac{dz}{(z - u)(v_1 - z)^2}$$

$$\times \big(R_{\mathfrak{m}}(z) + Q_{\mathfrak{m}}(z) \big).$$

The term with $R_{\mathfrak{m}}(z)$ integrates to 0, as it has no singularities inside $\gamma_{[0,\mathfrak{m}]}$. The term with $Q_{\mathfrak{m}}$ is computed as the sum of the residues at $z = u$ and at $z = v_1$, which gives the desired covariance formula. $\qquad\square$

Theorem 2.8 implies the central limit theorem for general analytic linear statistics.

Corollary **2.11.** — *Let the distribution* $\mathbf{P}_N$ *be given by* (8) *with* $M = \lfloor \mathfrak{m}N \rfloor$. *Take k real valued functions* $f_1(z), \ldots, f_k(z)$ *on* $[0, \mathfrak{m}]$, *which can be extended to holomorphic functions in a complex neighborhood* $\mathcal{B}$ *of* $[0, \mathfrak{m}]$. *Then as* $N \to \infty$ *the k random variables*

$$\mathcal{L}_{f_j} = \sum_{i=1}^{N} \big(f_j(\ell_i) - \mathbf{E}_{\mathbf{P}_N} f_j(\ell_i) \big), \quad (\ell_1, \ldots, \ell_N) \text{ is } \mathbf{P}_N\text{-distributed,}$$

converge in the sense of moments to centered Gaussian random variables with covariance

$$(24) \qquad \lim_{N \to \infty} \mathbf{E}_{\mathbf{P}_N} \mathcal{L}_{f_i} \mathcal{L}_{f_j} = \frac{1}{(2\pi\mathbf{i})^2} \oint_{\gamma_{[0,\mathfrak{m}]}} \oint_{\gamma_{[0,\mathfrak{m}]}} f_i(u) f_j(v) \mathcal{C}(u, v) \, du \, dv,$$

where $\gamma_{[0,\mathfrak{m}]}$ *is a positively oriented contour in* $\mathcal{B}$ *which encloses* $[0, \mathfrak{m}]$, *and* $\mathcal{C}(u, v)$ *is given by* (22).

Remark **2.12.** — The covariance (24) has the same form as for random matrices and log-gases in the one cut regime. It depends only on the restrictions of functions f_j onto the interval $[a_-, a_+]$ and can be rewritten in several other equivalent forms, cf. [J1, Theorem 4.2], [PS, Chapter 3], [AGZ, Section 4.3.3].

Proof of Corollary 2.11. — Observe that

$$\mathcal{L}_f = \frac{N}{2\pi\mathbf{i}} \oint_{\gamma_{[0,\mathfrak{m}]}} f(z) \big(G_N(z) - \mathbf{E}_{\mathbf{P}_N} G_N(z) \big) dz.$$

Therefore, all the moments of $\mathcal{L}_f$ are obtained from the centered moments of $G_N(z)$ by integration. Since the latter converge uniformly in z on the integration contour, so do the former. It remains to use the fact that the integrals of jointly Gaussian random variables are also Gaussian. $\qquad\square$

Remark **2.13.** — Similarly to Corollary 2.11, Theorem 2.5 can be used to obtain the first two terms in the $N \to \infty$ asymptotic expansion of $\mathbf{E}_{\mathbf{P}_N} \sum_{i=1}^{N} f(\ell_i)$ for functions f holomorphic in a neighborhood of $[0, \mathfrak{m}]$.

The rest of this section is devoted to the proof of Theorem 2.5.

2.3. *Heuristic argument for Theorem 2.5.* — In this section we present a sketch of the proof for Theorem 2.5 in which we omit crucial bounds on remainders in the asymptotic formulas. These bounds will be established further on.

We start from the statement of Proposition 2.1 for the measures $\mathbf{P}_N^{\mathbf{t}, \mathbf{v}}$. Making the change of variables $\xi = Nz$, we have for z outside $[0, \mathfrak{m}]$

$$
(25) \qquad \prod_{i=1}^{N} \left(1 - \frac{1}{\xi - \ell_i} \right)
$$

$$
= \exp\left(\sum_{i=1}^{N} \ln\left(1 - \frac{1}{N} \cdot \frac{1}{z - \ell_i/N} \right) \right)
$$

$$
= \exp\left(-G_N(z) + \frac{1}{2N} \cdot \frac{\partial}{\partial z} G_N(z) + O\left(\frac{1}{N^2} \right) \right)
$$

$$
= \exp\left(-G_{\mathfrak{m}}(z) - \frac{1}{N} \Delta G_N(z) + \frac{1}{2N} \cdot \frac{\partial}{\partial z} G_N(z) + O\left(\frac{1}{N^2} \right) \right),
$$

where the remainder is uniform over z in compact subsets of $\mathbf{C} \setminus [0, \mathfrak{m}]$. Similarly,

$$
(26) \qquad \prod_{i=1}^{N} \left(1 + \frac{1}{\xi - \ell_i - 1} \right) = \exp\left(\ln\left(1 + \frac{1}{N} \cdot \frac{1}{z - \ell_i/N - 1/N} \right) \right)
$$

$$
= \exp\left(G_{\mathfrak{m}}(z) + \frac{1}{N} \Delta G_N(z) - \frac{1}{2N} \cdot \frac{\partial}{\partial z} G_N(z) + O\left(\frac{1}{N^2} \right) \right).
$$

Recalling the definition of $R_{\mathfrak{m}}(z)$, we conclude that the function $R_N(Nz)$ from Proposition 2.1 can be written in the following form

$$
(27) \qquad R_N(Nz) = \prod_{a=1}^{k} (v_a - z)(t_a + v_a - z)
$$

$$
\times \left[R_{\mathfrak{m}}(z) + z e^{-G_{\mathfrak{m}}(z)} \mathbf{E}_{\mathbf{P}_N^{\mathbf{t},\mathbf{v}}}\left(\exp\left(\frac{1}{N} \Delta G_N(z) \right) - 1 \right) \right.
$$

$$
\left. + (\mathfrak{m} - z) e^{G_{\mathfrak{m}}(z)} \mathbf{E}_{\mathbf{P}_N^{\mathbf{t},\mathbf{v}}}\left(\exp\left(-\frac{1}{N} \Delta G_N(z) \right) - 1 \right) \right.
$$

$$+ \frac{\psi_N^-(z)}{N} e^{-G_\mathfrak{m}(z)} + \frac{\psi_N^+(z)}{N} e^{G_\mathfrak{m}(z)} + \frac{z e^{-G_\mathfrak{m}(z)}}{N} \left(\frac{1}{2} \frac{\partial}{\partial z} G_\mathfrak{m}(z) \right)$$

$$+ \frac{(\mathfrak{m} - z) e^{G_\mathfrak{m}(z)}}{N} \left(-\frac{1}{2} \frac{\partial}{\partial z} G_\mathfrak{m}(z) \right) \Bigg] + o\left(\frac{1}{N} \right),$$

where $\psi_N^\pm$ appeared from the second terms in $1/N$ expansion of $\phi_N^\pm$ from (16), (17). We further want to simplify the expression in the second line of (27), by replacing $e^h - 1$ by h under expectations. For that we note a simple inequality, which we will use with $n = 2$:

$$(28) \qquad \left| e^h - \sum_{j=0}^{n-1} \frac{h^j}{j!} \right| \leq |h|^n e^{|h|}, \quad h \in \mathbf{C}, \ n = 1, 2, \ldots.$$

We will later establish in Section 2.5 that uniformly in z in compact subsets of $\mathbf{C} \setminus [0, \mathfrak{m}]$ we have

$$(29) \qquad \mathbf{E}_{\mathbf{P}_N^{t,v}} \left(\left| \frac{1}{N} \Delta G_N(z) \right|^2 \exp \left(\left| \frac{1}{N} \Delta G_N(z) \right| \right) \right) = o\left(\frac{1}{N} \right), \quad N \to \infty.$$

We therefore can rewrite (27) as

$$(30) \qquad R_N(Nz) = \prod_{a=1}^{k} (v_a - z)(t_a + v_a - z) \cdot \left[R_\mathfrak{m}(z) + \frac{Q_\mathfrak{m}(z)}{N} \mathbf{E}_{\mathbf{P}_N^{t,v}} \big(\Delta G_N(z) \big) \right.$$

$$+ \frac{\psi_N^-(z)}{N} e^{-G_\mathfrak{m}(z)} + \frac{\psi_N^+(z)}{N} e^{G_\mathfrak{m}(z)} + \frac{z e^{-G_\mathfrak{m}(z)}}{N} \left(\frac{1}{2} \frac{\partial}{\partial z} G_\mathfrak{m}(z) \right)$$

$$\left. + \frac{(\mathfrak{m} - z) e^{G_\mathfrak{m}(z)}}{N} \left(-\frac{1}{2} \frac{\partial}{\partial z} G_\mathfrak{m}(z) \right) \right] + o\left(\frac{1}{N} \right),$$

where the estimate of the remainder is uniform over z in compact subsets of $\mathbf{C} \setminus [0, \mathfrak{m}]$.

Let us now fix u outside the contour $\gamma_{[0,\mathfrak{m}]}$ enclosing the interval $[0, \mathfrak{m}]$, divide (30) by

$$2 \cdot 2\pi \mathbf{i} \cdot (u - z) \cdot \prod_{a=1}^{k} (v_a - z)(t_a + v_a - z)$$

and integrate over $\gamma_{[0,\mathfrak{m}]}$. Since both $R_N(Nz)$ and $R_\mathfrak{m}(z)$ are holomorphic inside the contour, the integrals of the corresponding terms vanish. From the rest we get, with the help of Proposition 2.3,

$$(31) \qquad \frac{1}{2\pi \mathbf{i}} \oint_{\gamma_{[0,\mathfrak{m}]}} \frac{\sqrt{(z - a_-)(z - a_+)}}{u - z} \cdot \mathbf{E}_{\mathbf{P}_N^{t,v}} \big(\Delta G_N(z) \big) dz$$

$$= -\frac{1}{2\pi \mathbf{i}} \oint_{\gamma_{[0,\mathfrak{m}]}} \frac{dz}{2(u-z)} \cdot \left(\psi_{\mathrm{N}}^-(z) e^{-\mathrm{G}_{\mathfrak{m}}(z)} + \psi_{\mathrm{N}}^+(z) e^{\mathrm{G}_{\mathfrak{m}}(z)} \right.$$

$$\left. + z e^{-\mathrm{G}_{\mathfrak{m}}(z)} \left(\frac{1}{2} \frac{\partial}{\partial z} \mathrm{G}_{\mathfrak{m}}(z) \right) + (\mathfrak{m} - z) e^{\mathrm{G}_{\mathfrak{m}}(z)} \left(-\frac{1}{2} \frac{\partial}{\partial z} \mathrm{G}_{\mathfrak{m}}(z) \right) \right) + o(1).$$

For the first line of (31), note that $\mathbf{E}_{\mathbf{P}_{\mathrm{N}}^{\mathbf{t,v}}}(\Delta \mathrm{G}_{\mathrm{N}}(z))$ is analytic outside the contour of integration and decays as $1/z^2$ when $z \to \infty$. Therefore, we can compute the integral as (minus) the residue at $z = u$, which is

$$\sqrt{(u - a_-)(u - a_+)} \cdot \mathbf{E}_{\mathbf{P}_{\mathrm{N}}^{\mathbf{t,v}}}\big(\Delta \mathrm{G}_{\mathrm{N}}(u)\big).$$

This leads, via differentiation in t's, to the desired formula (21) of Theorem 2.5.

Let us point out the parts of the above argument that are not yet rigorous and whose justification is necessary to complete the proof of Theorem 2.5:

- We need to show that the bound (29) is valid.
- We need to prove that all the remainders remain small when we differentiate with respect to variables t_a, $a = 1, \ldots, k$.

2.4. *A weak a priori estimate.* — The following lemma is a first step for establishing the desired estimates.

Lemma 2.14. — *Fix a positive integer n and take a compact set* $\mathrm{A} \subset \mathbf{C} \setminus [0, \mathfrak{m}]$. *Then for every* $\varepsilon > 0$ *there exists a constant* $\mathrm{C} > 0$ *such that for every* $z_1, \ldots, z_n \in \mathrm{A}$ *and every* $\mathrm{N} = 1, 2, \ldots$ *we have*

$$(32) \qquad \mathbf{E}_{\mathbf{P}_{\mathrm{N}}} \left| \prod_{i=1}^{n} \Delta \mathrm{G}_{\mathrm{N}}(z_i) \right| \le \mathrm{C} \cdot \mathrm{N}^{n(\frac{1}{2}+\varepsilon)},$$

with $\Delta \mathrm{G}_{\mathrm{N}}(z)$ *defined by* (19) *with* $k = 0$.

Remark 2.15. — We will show in the next section that $\mathrm{C} \cdot \mathrm{N}^{n(\frac{1}{2}+\varepsilon)}$ in the right-hand side of (32) can be replaced by a constant. However, to produce such a sharp estimate we need to start from this weaker one.

The proof of Lemma 2.14 that we now present is similar to the argument of [BoGu2, Lemma 3.4.5], which follows the ideas of [1].

Before we proceed we need to recall the characterization of the equilibrium measure $\mu_{\mathfrak{m}}$. Introduce the functional $\mathrm{I}[\mu]$ of a measure μ on $[0, \mathfrak{m}]$ via

$$(33) \qquad \mathrm{I}[\mu] = \iint_{\substack{0 \le x, y \le \mathfrak{m} \\ x \ne y}} \ln |x - y| d\mu(x) d\mu(y) - \int_0^{\mathfrak{m}} \mathrm{V}(x) d\mu(x),$$

where

$$V(x) = x\ln(x) + (\mathfrak{m} - x)\ln(\mathfrak{m} - x).$$

Note that if the measure μ has atoms, then it is important to exclude the diagonal in the integral (33), i.e. integrate only over $x \neq y$.

The variational characterization of the measure $\mu_{\mathfrak{m}}(x)dx^6$ (see [J3], [J4], [Fe]) yields that $\mu_{\mathfrak{m}}(x)dx$ is the unique minimizer of $I[\mu]$ among the probability measures of density at most 1.

Define the functions $F_{\mathfrak{m}}(x)$ through

(**34**) $$F_{\mathfrak{m}}(x) = 2\int_0^{\mathfrak{m}} \ln|x - t|\mu_{\mathfrak{m}}(t)dt - V(x).$$

Then varying the functional $I[\mu]$ at $\mu_{\mathfrak{m}}$, one proves (cf. [DS], [Fe], [ST]) that there exists a real number f such that $F_{\mathfrak{m}}(x) - f = 0$ on $S = \{0 \leq x \leq \mathfrak{m} \mid 0 < \mu_{\mathfrak{m}}(x) < 1\}$, $F_{\mathfrak{m}}(x) - f < 0$ on the complement of the support of $\mu_{\mathfrak{m}}$, and $F_{\mathfrak{m}}(x) - f > 0$ when the density $\mu_{\mathfrak{m}}(x)$ is equal to 1.

Now take any two compactly supported absolutely continuous probability measures with uniformly bounded densities $\nu(dx) = \nu(x)dx$ and $\rho(dx) = \rho(x)dx$ and define $\mathcal{D}(\nu, \rho)$ through

(**35**) $$\mathcal{D}^2(\nu, \rho) = -\int_{\mathbf{R}}\int_{\mathbf{R}} \ln|x - y|\big(\nu(x) - \rho(x)\big)\big(\nu(y) - \rho(y)\big)dxdy.$$

There is an alternative formula for $\mathcal{D}(\nu(x), \rho(x))$ in terms of Fourier transforms, cf. [BeGu]:

(**36**) $$\mathcal{D}(\nu, \rho) = \sqrt{\int_0^\infty \frac{1}{t}\left|\int_{\mathbf{R}} e^{\mathbf{i}tx}\big(\nu(x) - \rho(x)\big)dx\right|^2 dt}.$$

Fix a parameter $p > 2$ and let $\tilde{\mu}_N$ denote the convolution of the empirical measure μ_N given by (12) with uniform measure on the interval $[0, N^{-p}]$.

Proposition **2.16.** — *There exists* $C \in \mathbf{R}$ *such that for all* $\gamma > 0$ *and all* $N \geq 1$ *we have*

$$\mathbf{P}_N\big(\mathcal{D}(\tilde{\mu}_N, \mu_{\mathfrak{m}}) \geq \gamma\big) \leq \exp\big(CN\ln(N) - \gamma^2 N^2\big).$$

Proof. — Observe that for every N-tuple $0 \leq \ell_1 < \ell_2 < \cdots < \ell_N \leq M$ we have, using Stirling's formula for factorials,

(**37**) $$\mathbf{P}_N(\ell_1, \ldots, \ell_N)$$
$$= \frac{\exp(2N(N - 1)\ln(N) + N^2 I[\text{mes}[\ell_1, \ldots, \ell_N]] + O(N\ln(N)))}{Z(N, \lfloor\mathfrak{m}N\rfloor)},$$

---

[6] Throughout the paper the density of a measure μ is denoted $\mu(x)$.

where

$$\mathrm{mes}[\ell_1, \ldots, \ell_N] = \frac{1}{N} \sum_{i=1}^{N} \delta\left(\frac{\ell_i}{N}\right).$$

Let us obtain a lower bound for the partition function $Z(N, \lfloor \mathfrak{m}N \rfloor)$ in (37). For that let x_i, $i = 1, \ldots, N$ be quantiles of $\mu_{\mathfrak{m}}$ defined through

$$\int_0^{x_i} \mu_{\mathfrak{m}}(x)\,dx = \frac{i - 1/2}{N}, \quad 1 \le i \le N.$$

Since $\mu_{\mathfrak{m}}(x) \le 1$, $x_{i+1} - x_i \ge 1/N$ and therefore, the numbers $\lfloor Nx_i \rfloor$, $1 \le i \le N$ are all distinct. We can then write

(**38**) $$Z(N, \lfloor \mathfrak{m}N \rfloor) \ge \exp\left(2N(N-1)\ln(N) + N^2 I\left[\mathrm{mes}\left[\lfloor Nx_1 \rfloor, \ldots, \lfloor Nx_N \rfloor\right]\right]\right.$$
$$\left. + O\left(N\ln(N)\right)\right).$$

We claim that (38) can be transformed into

$$\exp\left(2N(N-1)\ln(N) + N^2 I[\mu_{\mathfrak{m}}] + O\left(N\ln(N)\right)\right).$$

Indeed, for the double-integral part of $I[\cdot]$ we write using the monotonicity of logarithm

$$\sum_{i<j} \ln\left(\frac{\lfloor Nx_j \rfloor}{N} - \frac{\lfloor Nx_i \rfloor}{N}\right)$$

$$\le \sum_{i<j} \ln\left(x_j - x_i + \frac{1}{N}\right)$$

$$\le N^2 \sum_{i<j} \int_{x_j}^{x_{j+1}} \int_{x_{i-1}}^{x_i} \ln\left(t - s + \frac{1}{N}\right) \mu_{\mathfrak{m}}(t)\mu_{\mathfrak{m}}(s)\,dt\,ds + O\left(N\ln(N)\right)$$

$$= N^2 \iint_{s<t} \ln(t - s)\mu_{\mathfrak{m}}(t)\mu_{\mathfrak{m}}(s)\,dt\,ds + O\left(N\ln(N)\right),$$

and similarly for the opposite inequality. For the single-integral part of $I[\cdot]$ we have

$$N \sum_{i=1}^{N} V\left(\frac{\lfloor Nx_i \rfloor}{N}\right) = N^2 \sum_{i=2}^{N-2} \int_{x_i}^{x_{i+1}} \left(V(t) + \left(t - \frac{\lfloor Nx_i \rfloor}{N}\right) V'\left(\kappa(t)\right)\right) \mu_{\mathfrak{m}}(t)\,dt$$
$$+ O(N),$$

where $\kappa(t)$ is a point inside $[x_1, x_N]$. Observe that $|V'(\kappa(t))| = O(\ln(N))$, and thus

$$N \sum_{i=1}^{N} V\left(\frac{\lfloor N x_i \rfloor}{N}\right)$$

$$= N^2 \sum_{i=2}^{N-2} \int_{x_i}^{x_{i+1}} \left(V(t) + \left(x_{i+1} - x_i + \frac{1}{N}\right) O(\ln(N))\right) \mu_{\mathfrak{m}}(t) dt + O(N)$$

$$= N^2 \int_0^{\mathfrak{m}} V(t) \mu_m(t) dt + N^2 \frac{O(\ln(N))}{N} \sum_{i=2}^{N-2} \left(x_{i+1} - x_i + \frac{1}{N}\right)$$

$$= N^2 \int_0^{\mathfrak{m}} V(t) \mu_m(t) dt + O(N \ln(N)).$$

The next step is to replace $\mathrm{mes}[\ell_1, \ldots, \ell_N]$ in (37) by its convolution with the uniform measure on $[0, N^{-p}]$, that we denote $\widetilde{\mathrm{mes}}[\ell_1, \ldots, \ell_N]$. For that take two independent random variables u, $\tilde{u}$ uniformly distributed on $[0, N^{-p}]$, where $p > 2$ as above. Then

(**39**) $I[\widetilde{\mathrm{mes}}[\ell_1, \ldots, \ell_N]]$

$$= \mathbf{E}_{u, \tilde{u}} \int_0^{\mathfrak{m}} \int_0^{\mathfrak{m}} \ln |x - y + u - \tilde{u}| \mathrm{mes}[\ell_1, \ldots, \ell_N](dx) \mathrm{mes}[\ell_1, \ldots, \ell_N](dy)$$

$$- \mathbf{E}_u \int_0^{\mathfrak{m}} V(x + u) \mathrm{mes}[\ell_1, \ldots, \ell_N](dx)$$

$$= I[\mathrm{mes}[\ell_1, \ldots, \ell_N]] + \frac{1}{N} \mathbf{E}_{u, \tilde{u}} \int_0^{\mathfrak{m}} \ln |u - \tilde{u}| \mathrm{mes}[\ell_1, \ldots, \ell_N](dx)$$

$$+ \mathbf{E}_{u, \tilde{u}} \iint_{x \neq y} \ln \left|1 + \frac{u - \tilde{u}}{x - y}\right| \mathrm{mes}[\ell_1, \ldots, \ell_N](dx) \mathrm{mes}[\ell_1, \ldots, \ell_N](dy)$$

$$+ \mathbf{E}_u \int_0^{\mathfrak{m}} \left(V(x + u) - V(x)\right) \mathrm{mes}[\ell_1, \ldots, \ell_N](dx)$$

$$= I[\mathrm{mes}[\ell_1, \ldots, \ell_N]] + O\left(\frac{\ln(N)}{N}\right).$$

We conclude that there exists a constant C such that

$$\mathbf{P}_N(\ell_1, \ldots, \ell_N) \leq \exp\left(CN \ln(N)\right) \exp\left(N^2 \left(I[\widetilde{\mathrm{mes}}[\ell_1, \ldots, \ell_N]] - I[\mu_{\mathfrak{m}}]\right)\right).$$

Further, completing the square we get

(**40**) $I[\widetilde{\mathrm{mes}}[\ell_1, \ldots, \ell_N]] - I[\mu_{\mathfrak{m}}]$

$$= -\mathcal{D}^2\left(\mu_{\mathfrak{m}}, \widetilde{\mathrm{mes}}[\ell_1, \ldots, \ell_N]\right) + \int_0^{\mathfrak{m}} F_{\mathfrak{m}}(x) \left(\widetilde{\mathrm{mes}}[\ell_1, \ldots, \ell_N] - \mu_{\mathfrak{m}}\right)(dx)$$

$$= -\mathcal{D}^2\big(\mu_\mathfrak{m}, \widetilde{\mathrm{mes}}[\ell_1, \ldots, \ell_N]\big)$$
$$+ \int_0^\mathfrak{m} \big(F_\mathfrak{m}(x) - f\big)\big(\widetilde{\mathrm{mes}}[\ell_1, \ldots, \ell_N] - \mu_\mathfrak{m}\big)(dx),$$

with $F_\mathfrak{m}(x)$ and f defined in (34) and directly below that formula, respectively. Let us analyze the last term in (40). On the set $S = \{x \mid 0 < \mu_\mathfrak{m}(x) < 1\}$ the function $F_\mathfrak{m}(x) - f$ vanishes. On the complement of the support of $\mu_\mathfrak{m}(x)$ we have $F_\mathfrak{m}(x) - f < 0$ and the corresponding part of the last integral in (40) is negative. Finally, on the set $S' = \{x \mid \mu_\mathfrak{m}(x) = 1\}$, we have $F_\mathfrak{m}(x) - f > 0$. Since all the points ℓ_i/N are at least $1/N$ apart, for any a such that $[a, a + 1/N] \subset S'$ we have

$$\big(\widetilde{\mathrm{mes}}[\ell_1, \ldots, \ell_N] - \mu_\mathfrak{m}\big)\big([a, a + 1/N]\big) \le 0.$$

Thus,

(**41**)
$$\int_a^{a+1/N} \big(F_\mathfrak{m}(x) - f\big)\big(\widetilde{\mathrm{mes}}[\ell_1, \ldots, \ell_N] - \mu_\mathfrak{m}\big)(dx)$$
$$\le \int_a^{a+1/N} \big(F_\mathfrak{m}(x) - F_\mathfrak{m}(a)\big)\big(\widetilde{\mathrm{mes}}[\ell_1, \ldots, \ell_N] - \mu_\mathfrak{m}\big)(dx)$$
$$\le \frac{2}{N} \sup_{\substack{x,y \in S' \\ |x-y| \le 1/N}} \big|F_\mathfrak{m}(x) - F_\mathfrak{m}(y)\big|.$$

Observe that the last sup is at most $O(\ln(N)/N)$. Now partition S' into segments of the form $[a, a + 1/N]$ and note that for boundary segments (which have to be shorter than $1/N$) the bound for the integral of the form (41) is still valid as $F_\mathfrak{m}(x)$ is equal to f in one of the end-points of such segment. Summing the bounds over all segments, we get a bound on the integral over S'. It follows that as $N \to \infty$

$$\int_0^\mathfrak{m} \big(F_\mathfrak{m}(x) - f\big)\big(\widetilde{\mathrm{mes}}[\ell_1, \ldots, \ell_N] - \mu_\mathfrak{m}\big)(dx) \le O\left(\frac{\ln(N)}{N}\right).$$

Therefore, we finally obtain

$$\mathbf{P}_N(\ell_1, \ldots, \ell_N) \le \exp\big(C'N\ln(N)\big)\exp\big(-N^2\mathcal{D}^2\big(\widetilde{\mathrm{mes}}[\ell_1, \ldots, \ell_N], \mu_\mathfrak{m}\big)\big).$$

Since the total number of N-tuples $0 \le \ell_1 < \ldots \ell_N \le \lfloor \mathfrak{m}N \rfloor$ is $\binom{\lfloor \mathfrak{m}N \rfloor + 1}{N} = \exp(O(N\ln(N)))$, the proof is complete. $\square$

Corollary **2.17.** — *For a compactly supported Lipschitz function g define*

$$\|g\|_{1/2} = \left(\int_{-\infty}^\infty |s| \left|\int_{-\infty}^\infty e^{\mathbf{i}sx} g(x)\, dx\right|^2 ds\right)^{1/2}, \quad \|g\|_{\mathrm{Lip}} = \sup_{x \ne y}\left|\frac{g(x) - g(y)}{x - y}\right|.$$

Fix any $p > 2$. Then there exists $C \in \mathbf{R}$ such that for all $\gamma > 0$, all $N \geq 1$ and all g we have

$$(\mathbf{42}) \qquad \mathbf{P}_N\left(\left|\int_0^{\mathfrak{m}} g(x)\mu_N(dx) - \int_0^{\mathfrak{m}} g(x)\mu_{\mathfrak{m}}(dx)\right| \geq \gamma \|g\|_{1/2} + \frac{\|g\|_{\mathrm{Lip}}}{N^p}\right)$$

$$\leq \exp\left(CN\ln(N) - \frac{\gamma^2 N^2}{2}\right).$$

Proof. — We have

$$\left|\int_0^{\mathfrak{m}} g(x)\mu_N(dx) - \int_0^{\mathfrak{m}} g(x)\mu_{\mathfrak{m}}(dx)\right|$$

$$\leq \left|\int_0^{\mathfrak{m}} g(x)\mu_N(dx) - \int_0^{\mathfrak{m}} g(x)\tilde{\mu}_N(dx)\right|$$

$$+ \left|\int_0^{\mathfrak{m}} g(x)\tilde{\mu}_N(dx) - \int_0^{\mathfrak{m}} g(x)\mu_{\mathfrak{m}}(dx)\right|.$$

The first term is bounded by $\frac{\|g\|_{\mathrm{Lip}}}{N^p}$ and corresponds to such term in (42). Therefore, it remains to work out the second term. Since scalar products are preserved under Fourier transform, with the notation

$$\hat{\phi}(s) = \int_{-\infty}^{\infty} e^{\mathbf{i}sx}\phi(x)dx$$

we write the Plancherel formula (note that $g(x)$ and $(\tilde{\mu}_N(x) - \mu_{\mathfrak{m}}(x))$ are bounded and belong to $L^1[0, \mathfrak{m}] \cap L^2[0, \mathfrak{m}]$) and use the Cauchy-Schwartz inequality

$$\left|\int_0^{\mathfrak{m}} g(x)\left(\tilde{\mu}_N(x) - \mu_{\mathfrak{m}}(x)\right)dx\right| = \left|\int_{-\infty}^{\infty}\left(\sqrt{|t|}\hat{g}(t)\right)\frac{\hat{\tilde{\mu}}_N(t) - \hat{\mu}_{\mathfrak{m}}(t)}{\sqrt{|t|}}dt\right|$$

$$\leq \|g\|_{1/2}\sqrt{\int_{-\infty}^{\infty}\frac{|\hat{\tilde{\mu}}_N(t) - \hat{\mu}_{\mathfrak{m}}(t)|^2}{|t|}dt}$$

$$= \sqrt{2}\|g\|_{1/2}\mathcal{D}(\tilde{\mu}_N, \mu_{\mathfrak{m}}).$$

It remains to use Proposition 2.16. □

Proof of Lemma 2.14. — Choose a small $\eta > 0$ and take an infinitely differentiable function $h(x)$, whose support is inside $[-\eta, \mathfrak{m} + \eta]$ and such that $h(x) = 1$ for $0 \leq x \leq \mathfrak{m}$.

Since both μ_N and $\mu_{\mathfrak{m}}$ are supported on $[0, \mathfrak{m}]$, we can replace $1/(z-x)$ in the definition of $G_N(z)$ and $G_{\mathfrak{m}}(z)$ by a nice smooth compactly supported function $h(x)/(z-x)$ without changing $G_N(z)$ and $G_{\mathfrak{m}}(z)$. Now choose $\gamma = q \cdot N^{-1/2+\varepsilon}$, $q > 0$, in Corollary 2.17

and note that for this choice the right-hand side of (42) still exponentially decays as $N \to \infty$. This readily implies the bound of Lemma 2.14. $\qquad\square$

2.5. *Self-improving estimates and the proof of Theorem 2.5.* — The aim of this section is to finish the proof of Theorem 2.5. The key part is establishing the following statement.

Proposition **2.18.** *— For any $k \geq 1$ and any $u_1, \ldots, u_k \in \mathbf{C} \setminus [0, \mathfrak{m}]$, the joint moments*

$$\mathbf{E}_{\mathbf{P}_N} \prod_{a=1}^{k} \big| \Delta G_N(u_a) \big|$$

are uniformly (in N and in $u_1, \ldots, u_k$ in compact subsets of $\mathbf{C} \setminus [0, \mathfrak{m}]$) bounded.

The idea for getting such estimates is to start from Lemma 2.14 and then recursively feed the existing estimates into the argument of Section 2.3 to obtain the stronger ones. This is very similar to the argument of [BoGu1, Section 4.3].

Proof of Proposition 2.18. — Fix $n = 0, 1, \ldots$ and $\mathbf{v} = (v_1, \ldots, v_n)$. Arguing as in Lemma 2.4, we prove that for any bounded random variable ξ we have

$$(43) \qquad \frac{\partial^n}{\partial t_1 \cdots \partial t_n} \big(\mathbf{E}_{\mathbf{P}_N^{\mathbf{t;v}}} \xi \big) \bigg|_{t_1 = \cdots = t_n = 0} = M_c \big(\xi, NG_N(v_1), \ldots, NG_N(v_n) \big),$$

where M_c is the joint cumulant. Since the cumulants are unchanged under shifts, we can also replace $NG_N(v_a)$ by $\Delta G_N(v_a)$ in (43).

We aim to differentiate the formulas of Section 2.3 with respect to t_a at $t_a = 0$. For that we need to examine the remainders. The remainder $o(N^{-1})$ in (27) comes from three sources: from the expansions (25), (26); from the N^{-1} expansion of $\phi_N^{\pm}(z)$; from the replacement of $\frac{\partial}{\partial z} G_N(z)$ by $\frac{\partial}{\partial z} G_{\mathfrak{m}}(z)$. The remainder can then be written as a sum corresponding to these three sources:

$$(44) \qquad \prod_{a=1}^{n} (v_a - z)(t_a + v_a - z) \left[\frac{1}{N^2} \mathbf{E}_{\mathbf{P}_N^{\mathbf{t;v}}} \xi_N(z) + \frac{1}{N^2} \mathfrak{c}(z; \mathbf{t}, \mathbf{v}) \mathbf{E}_{\mathbf{P}_N^{\mathbf{t;v}}} \xi_N'(z) \right.$$

$$\left. + \frac{1}{N^2} \mathbf{E}_{\mathbf{P}_N^{\mathbf{t;v}}} \left(\xi_N''(z) \frac{\partial}{\partial z} \Delta G_N(z) \right) \right]$$

where $\xi_N(z), \xi_N'(z), \xi_N''(z)$ are random variables, which are bounded uniformly in N, in z belonging to compact subsets of $\mathbf{C} \setminus [0, \mathfrak{m}]$, and which do not depend on $\mathbf{t}, \mathbf{v}$. The function $\mathfrak{c}(z; \mathbf{t}, \mathbf{v})$ arises from the large N expansions of (16), (17), and it is uniformly bounded in z belonging to compact subsets of $\mathbf{C} \setminus \{v_1, \ldots, v_n; t_1 + v_1; \ldots t_n + v_n\}$. The dependence on z is holomorphic in all the terms.

Further, when we pass from (27) to (30), we need to expand $\mathbf{E}_{\mathbf{P}_N}(\exp(\frac{1}{N}\Delta G_N(z)) - 1)$. Since $\frac{1}{N}\Delta G_N(z)$ is a bounded random variable, we can use the Taylor expansion to get

$$(45) \qquad \mathbf{E}_{\mathbf{P}_N^{t,v}}\left(\exp\left(\frac{1}{N}\Delta G_N(z)\right) - 1\right)$$

$$= \mathbf{E}_{\mathbf{P}_N^{t,v}}\left(\frac{1}{N}\Delta G_N(z)\right) + \frac{1}{2}\mathbf{E}_{\mathbf{P}_N^{t,v}}\left(\frac{1}{N}\Delta G_N(z)\right)^2$$

$$+ \frac{1}{6}\mathbf{E}_{\mathbf{P}_N^{t,v}}\left(\frac{1}{N}\Delta G_N(z)\right)^3 + \dots$$

Thus, after reconstructing the remainders, (31) is replaced by

$$(46) \qquad \frac{1}{4\pi\mathbf{i}}\oint_{\gamma_{[0,m]}}\frac{1}{u-z}\cdot\left(Q_m(z)\mathbf{E}_{\mathbf{P}_N^{t,v}}\big(\Delta G_N(z)\big) + \frac{R_m(z)}{2N}\mathbf{E}_{\mathbf{P}_N^{t,v}}\big(\Delta G_N(z)^2\big)\right.$$

$$+ \frac{Q_m(z)}{6N^2}\mathbf{E}_{\mathbf{P}_N^{t,v}}\big(\Delta G_N(z)^3\big) + \dots$$

$$+ \frac{1}{N}\left[\mathbf{E}_{\mathbf{P}_N^{t,v}}\xi_N(z) + \mathfrak{c}(z; \mathbf{t}, \mathbf{v})\mathbf{E}_{\mathbf{P}_N^{t,v}}\xi_N'(z)\right.$$

$$\left.\left. + \mathbf{E}_{\mathbf{P}_N^{t,v}}\left(\xi_N''(z)\frac{\partial}{\partial z}\Delta G_N(z)\right)\right]\right)dz$$

$$= -\frac{1}{2\pi\mathbf{i}}\oint_{\gamma_{[0,m]}}\frac{dz}{2(u-z)}$$

$$\times\left(\psi_N^-(z)e^{-G_m(z)} + \psi_N^+(z)e^{G_m(z)} + \frac{Q_m(z)}{2}\frac{\partial}{\partial z}G_m(z)\right).$$

We now differentiate (46) with respect to all t_a at $t_a = 0$. For all the terms involving random variables we use (43) to rewrite the result as a cumulant. Note that when we differentiate $\mathfrak{c}(z; \mathbf{t}, \mathbf{v})\mathbf{E}_{\mathbf{P}_N^{t,v}}\xi_N''(z)$, we need to apply the Leibnitz rule and therefore to differentiate each of the two factors and get a sum. We also compute the integral of $\frac{1}{u-z}Q_m(z)\mathbf{E}_{\mathbf{P}_N^{t,v}}(\Delta G_N(z))$ as minus the residue at $z = u$ to get

$$(47) \qquad \frac{Q_m(u)}{2}M_c\big(\Delta G_N(u), \Delta G_N(v_1), \dots, \Delta G_N(v_n)\big)$$

$$+ \frac{1}{4\pi\mathbf{i}}\oint_{\gamma_{[0,m]}}\frac{1}{u-z}\left(\frac{R_m(z)}{4N}M_c\big((\Delta G_N(z))^2, \Delta G_N(v_1), \dots, \Delta G_N(v_n)\big)\right.$$

$$\left. + \frac{Q_m(z)}{12N^2}M_c\big((\Delta G_N(z))^3, \Delta G_N(v_1), \dots, \Delta G_N(v_n)\big) + \dots\right)dz$$

$$+ \frac{1}{N} \cdot \frac{1}{4\pi \mathbf{i}} \oint_{\gamma_{[0,\mathfrak{m}]}} \frac{1}{u-z} M_c$$

$$\times \left(\xi_N(z) + \xi_N''(z) \frac{\partial}{\partial z} \Delta G_N(z), \Delta G_N(v_1), \dots, \Delta G_N(v_n) \right) dz$$

$$+ \frac{1}{N} \cdot \frac{1}{4\pi \mathbf{i}} \sum_{A \subset \{1,\dots,n\}} \oint_{\gamma_{[0,\mathfrak{m}]}} \frac{1}{u-z} \mathfrak{c}_A(z; \mathbf{t}, \mathbf{v}) M_c\big(\xi_N'(z), \Delta G_N(v_a), a \in A \big) dz$$

$$= \frac{\partial^n}{\partial t_1 \cdots \partial t_n} \left[\frac{1}{4\pi \mathbf{i}} \oint_{\gamma_{[0,\mathfrak{m}]}} \frac{dz}{(u-z)} \cdot \left(\psi_N^-(z) e^{-G_\mathfrak{m}(z)} + \psi_N^+(z) e^{G_\mathfrak{m}(z)} \right. \right.$$

$$\left. \left. + \frac{Q_\mathfrak{m}(z)}{2} \cdot \frac{\partial}{\partial z} G_\mathfrak{m}(z) \right) \right]_{t_a=0, 1 \le a \le n},$$

where the fifth line is the result of the differentiation of $\mathfrak{c}(z; \mathbf{t}, \mathbf{v}) \mathbf{E}_{\mathbf{P}_N^{\mathbf{t};\mathbf{v}}} \xi_N''(z)$, i.e. $\mathfrak{c}_A(z; \mathbf{t}, \mathbf{v})$ is the mixed derivative of $\mathfrak{c}(z; \mathbf{t}, \mathbf{v})$ with respect to t_a, $a \in \{1, \dots, n\} \setminus A$. Let us analyze the resulting expression (47) as $N \to \infty$. At this moment all we need from the right-hand side of (47) is that it is $O(1)$ as $N \to \infty$.

Consider the infinite sum over growing powers of $\Delta G_N(z)$ in (47). All the terms starting from the Hth one can be combined into

$$(\mathbf{48}) \qquad M_c\left(\sum_{h=H}^{\infty} (Q/R)_h \frac{(\Delta G_N(z))^h}{2h! N^{h-1}}, \Delta G_N(v_1), \dots, \Delta G_N(v_n) \right),$$

where $(Q/R)_h$ is either $Q_\mathfrak{m}(z)$ or $R_\mathfrak{m}(z)$ depending on the parity of h. Expanding the cumulants in terms of centered moments and using Holder's inequality, we observe that if ζ is a bounded random variable, then for any fixed $k \ge 1$ there is a constant C_k such that

$$(\mathbf{49}) \qquad \left| M_c\big(\zeta, \Delta G_N(v_1), \dots, \Delta G_N(v_k) \big) \right|$$

$$\le C_k \sqrt{\mathbf{E}|\zeta|^2} \prod_{i=1}^{k} \big[\mathbf{E}|\Delta G_N(v_i) - \mathbf{E}[\Delta G_N(v_i)]|^{2k} \big]^{1/2k}$$

Observe that the random variables $G_N(v_a)$ are uniformly bounded, same is true about $(Q/R)_h$ (for z and v_a in compact subsets of $\mathbf{C} \setminus [0, \mathfrak{m}]$), and

$$\zeta = \sum_{h=H}^{\infty} (Q/R)_h \frac{(\Delta G_N(z))^h}{2h! N^{h-1}}$$

satisfies (uniformly over z in compact subsets of $\mathbf{C} \setminus [0, \mathfrak{m}]$) the point-wise bound

$$|\zeta| \le \text{const} \cdot N \cdot \left| \frac{\Delta G_N(z)}{N} \right|^H \exp \left| \frac{\Delta G_N(z)}{N} \right|.$$

Thus, we can bound (48) by

$$
(\mathbf{50}) \qquad \text{const} \cdot \mathrm{N}^{n+1} \sup_z \sqrt{\mathbf{E}_{\mathbf{P}_\mathrm{N}}\left(\left|\frac{\Delta \mathrm{G}_\mathrm{N}(z)}{\mathrm{N}}\right|^{2\mathrm{H}} \exp\left|\frac{2\Delta \mathrm{G}_\mathrm{N}(z)}{\mathrm{N}}\right|\right)}
$$

$$
\leq \text{const} \cdot \mathrm{N}^{n+1} \sup_z \sqrt{\mathbf{E}_{\mathbf{P}_\mathrm{N}}\left(\left|\frac{\Delta \mathrm{G}_\mathrm{N}(z)}{\mathrm{N}}\right|^{2\mathrm{H}}\right)},
$$

where in the last inequality we used the uniform boundedness of $\frac{\Delta \mathrm{G}_\mathrm{N}(u)}{\mathrm{N}}$, and $\sup_z$ is to be taken over any set which includes both $\gamma_{[0,\mathfrak{m}]}$ and all points v_a, $a = 1, \ldots, n$.

We can use Lemma 2.14 to bound the expectation in (50) and conclude that if $\mathrm{H} > 3n$, then (48) is $o(1)$. Therefore, renaming u into v_0, (47) is rewritten as

$$
(\mathbf{51}) \qquad \frac{\mathrm{Q}_{\mathfrak{m}}(v_0)}{2} \mathrm{M}_c\big(\Delta \mathrm{G}_\mathrm{N}(v_0), \Delta \mathrm{G}_\mathrm{N}(v_1), \ldots, \Delta \mathrm{G}_\mathrm{N}(v_n)\big)
$$

$$
+ \frac{1}{4\pi \mathbf{i}} \oint_{\gamma_{[0,\mathfrak{m}]}} \frac{1}{v_0 - z} \left(\sum_{h=2}^{3n} \frac{1}{\mathrm{N}^{h-1}}\right.
$$

$$
\left. \times \frac{(\mathrm{Q}/\mathrm{R})_h}{2h!} \mathrm{M}_c\big((\Delta \mathrm{G}_\mathrm{N}(z))^h, \Delta \mathrm{G}_\mathrm{N}(v_1), \ldots, \Delta \mathrm{G}_\mathrm{N}(v_n)\big)\right) dz
$$

$$
+ \frac{1}{\mathrm{N}} \cdot \frac{1}{4\pi \mathbf{i}} \oint_{\gamma_{[0,\mathfrak{m}]}} \frac{1}{v_0 - z} \mathrm{M}_c\bigg(\xi_\mathrm{N}(z)
$$

$$
+ \xi_\mathrm{N}''(z) \frac{\partial}{\partial z} \Delta \mathrm{G}_\mathrm{N}(z), \Delta \mathrm{G}_\mathrm{N}(v_1), \ldots, \Delta \mathrm{G}_\mathrm{N}(v_n)\bigg) dz
$$

$$
+ \frac{1}{\mathrm{N}} \cdot \frac{1}{4\pi \mathbf{i}} \sum_{A \subset \{1,\ldots,n\}} \oint_{\gamma_{[0,\mathfrak{m}]}} \frac{1}{v_0 - z} \mathfrak{c}_A(z; \mathbf{t}, \mathbf{v})
$$

$$
\times \mathrm{M}_c\big(\xi_\mathrm{N}'(z), \Delta \mathrm{G}_\mathrm{N}(v_a), a \in A\big) dz
$$

$$
= \mathrm{O}(1),
$$

where $(\mathrm{Q}/\mathrm{R})_h$ is either $\mathrm{Q}_{\mathfrak{m}}(z)$ or $\mathrm{R}_{\mathfrak{m}}(z)$ depending on the parity of h, and we replaced u by v_0.

At this moment we claim that (51) for all $n = 0, 1, 2\ldots$ together with the bound of Lemma 2.14 implies Proposition 2.18. Indeed, take any two *disjoint* compact sets $\mathcal{U}, \mathcal{V} \subset \mathbf{C} \setminus [0, \mathfrak{m}]$, which are invariant under conjugation, and suppose that $\gamma_{[0,\mathfrak{m}]} \subset \mathcal{U}$. Expanding cumulants in terms of centered moments and using Lemma 2.14, we obtain

$$
\text{(52)} \quad \sup_{v_0,\dots,v_n\in\mathcal{V}} \frac{1}{N}\cdot\frac{1}{4\pi\mathbf{i}}\oint_{\gamma_{[0,\mathfrak{m}]}}\frac{1}{v_0-z}
$$

$$
\times M_c\!\left(\xi_N(z)+\xi_N''(z)\frac{\partial}{\partial z}\Delta G_N(z),\,\Delta G_N(v_1),\dots,\Delta G_N(v_n)\right)dz
$$

$$
= O\!\left(N^{(n+1)(1/2+\varepsilon)-1}\right),
$$

$$
\text{(53)} \quad \sup_{v_0,\dots,v_n\in\mathcal{V}} \frac{1}{N}\cdot\frac{1}{4\pi\mathbf{i}}\sum_{A\subset\{1,\dots,n\}}\oint_{\gamma_{[0,\mathfrak{m}]}}\frac{1}{v_0-z}\mathfrak{c}_A(z;\mathbf{t},\mathbf{v})
$$

$$
\times M_c\!\left(\xi_N'(z),\,\Delta G_N(v_a),\,a\in A\right)dz
$$

$$
= O\!\left(N^{n(1/2+\varepsilon)-1}\right),
$$

$$
\text{(54)} \quad \sup_{v_0,\dots,v_n\in\mathcal{V}} \frac{1}{4\pi\mathbf{i}}\oint_{\gamma_{[0,\mathfrak{m}]}}\frac{1}{v_0-z}\left(\frac{1}{N^{h-1}}\right.
$$

$$
\left.\times\frac{(Q/R)_h}{2h!}M_c\!\left(\left(\Delta G_N(z)\right)^h,\,\Delta G_N(v_1),\dots,\Delta G_N(v_n)\right)\right)dz
$$

$$
= O\!\left(N^{(h+n)(1/2+\varepsilon)+1-h}\right).
$$

One might be cautious about (52), as it involves the derivative of $\Delta G_N(z)$ (instead of $\Delta G_N(z)$ itself as in Lemma 2.14), yet for analytic functions a uniform bound for a function implies a uniform bound for its derivative, therefore, the bound is valid.

We plug (52), (53), (54) into (51) to get

$$
\text{(55)} \quad \sup_{v_0,\dots,v_n\in\mathcal{V}} \frac{Q_{\mathfrak{m}}(v_0)}{2}M_c\!\left(\Delta G_N(v_0),\Delta G_N(v_1),\dots,\Delta G_N(v_n)\right)
$$

$$
+\left[\sum_{h=2}^{3n}O\!\left(N^{(h+n)(1/2+\varepsilon)+1-h}\right)\right]+O\!\left(N^{(n+1)(1/2+\varepsilon)-1}\right)+O\!\left(N^{n(1/2+\varepsilon)-1}\right)
$$

$$
= O(1),
$$

which implies that for each $n=0,1,2\dots$

$$
\text{(56)} \quad \sup_{v_0,v_1,\dots,v_n\in\mathcal{V}}\left|M_c\!\left(\Delta G_N(v_0),\dots,\Delta G_N(v_{n-1})\right)\right|
$$

$$
= O\!\left(N^{(n+1)(1/2+\varepsilon)-(1/2-\varepsilon)}\right)+O(1),
$$

where the remainder depends on the choice of the compact set $\mathcal{V}\subset\mathbf{C}\setminus[0,\mathfrak{m}]$. As centered moments are linear combinations of products of joint cumulants, we deduce from

(56) that for all $k = 1, 2, \ldots,$

$$(\mathbf{57}) \quad \sup_{v_1,\ldots,v_k\in\mathcal{V}} \mathbf{E}\left[\prod_{a=1}^{k}\left(\Delta G_N(v_a) - \mathbf{E}\left[\Delta G_N(v_a)\right]\right)\right] = O\left(N^{k(1/2+\varepsilon)-(1/2-\varepsilon)}\right) + O(1).$$

Combining $n = 0$ version of (56) with (57) we finally conclude that

$$(\mathbf{58}) \quad \sup_{v_1,\ldots,v_k\in\mathcal{V}} \mathbf{E}\left[\prod_{a=1}^{k}\Delta G_N(v_a)\right] = O\left(N^{k(1/2+\varepsilon)-(1/2-\varepsilon)}\right) + O(1), \quad k = 1, 2, \ldots.$$

If k is even, then we can choose $v_1 = \cdots = v_{k/2} = v$ and $v_{k/2+1} = \cdots = v_k = \bar{v}$ in (58) and get

$$(\mathbf{59}) \quad \sup_{v\in\mathcal{V}} \mathbf{E}\left[\left|\Delta G_N(v)\right|^k\right] = O\left(N^{k(1/2+\varepsilon)-(1/2-\varepsilon)}\right) + O(1), \quad k = 2, 4, 6, \ldots.$$

The estimate for odd k is reduced to the even k case with the use of Jensen's inequality in the form

$$\mathbf{E}\left(\xi^k\right) \le \left(\mathbf{E}\left(\xi^{k+1}\right)\right)^{\frac{k}{k+1}}, \quad k > 0,$$

which leads to the following bound for all k:

$$(\mathbf{60}) \quad \sup_{v\in\mathcal{V}} \mathbf{E}\left[\left|\Delta G_N(v)\right|^k\right] = O\left(N^{k(1/2+\varepsilon)-\frac{k}{k+1}(1/2-\varepsilon)}\right) + O(1), \quad k = 1, 2, 3, \ldots.$$

We finally use the Holder inequality to get

$$(\mathbf{61}) \quad \sup_{v_1,\ldots,v_k\in\mathcal{V}} \mathbf{E}\left[\prod_{a=1}^{k}\left|\Delta G_N(v_a)\right|\right] \le \sup_{v\in\mathcal{V}} \mathbf{E}\left[\left|\Delta G_N(v)\right|^k\right]$$

$$= O\left(N^{k(1/2+\varepsilon)-\frac{k}{k+1}(1/2-\varepsilon)}\right) + O(1)$$

$$= O\left(N^{k(1/2+\varepsilon)-1/6}\right) + O(1), \quad k = 1, 2, 3, \ldots,$$

where we silently assumed that $\varepsilon > 0$ is small enough for the last equality to hold.

At this moment we can iterate the argument. Expanding cumulants in terms of moments, we deduce from (61) the following three bounds:

$$(\mathbf{62}) \quad \sup_{v_0,\ldots,v_n\in\mathcal{V}} \frac{1}{N} \cdot \frac{1}{4\pi\mathbf{i}} \oint_{\gamma[0,\mathrm{m}]} \frac{1}{v_0 - z}$$

$$\times M_c\left(\xi_N(z) + \xi_N''(z)\frac{\partial}{\partial z}\Delta G_N(z), \Delta G_N(v_1), \ldots, \Delta G_N(v_n)\right)dz$$

$$= O\left(N^{(n+1)(1/2+\varepsilon)-1-1/6}\right),$$

$$(\mathbf{63}) \qquad \sup_{v_0,\dots,v_n\in\mathcal{V}} \frac{1}{N}\cdot\frac{1}{4\pi\mathbf{i}}$$

$$\times \sum_{A\subset\{1,\dots,n\}} \oint_{\gamma_{[0,\mathfrak{m}]}} \frac{1}{v_0-z}\mathfrak{c}_A(z;\mathbf{t},\mathbf{v})M_c\big(\xi'_N(z),\Delta G_N(v_a), a\in A\big)dz$$

$$= O\big(N^{n(1/2+\varepsilon)-1-1/6}\big),$$

$$(\mathbf{64}) \qquad \sup_{v_0,\dots,v_n\in\mathcal{V}} \frac{1}{4\pi\mathbf{i}}\oint_{\gamma_{[0,\mathfrak{m}]}} \frac{1}{v_0-z}$$

$$\times \left(\frac{1}{N^{h-1}}\cdot\frac{(Q/R)_h}{2h!}M_c\big(\big(\Delta G_N(z)\big)^h, \Delta G_N(v_1),\dots,\Delta G_N(v_n)\big)\right)dz$$

$$= O\big(N^{(h+n)(1/2+\varepsilon)+1-h-1/6}\big).$$

The only difference between (52), (53), (54) and (62), (63), (64) is that the degree of N in the bound decreased by $1/6$ (which is because the bound of Lemma 2.14 is replaced by (61)).

We then plug (62), (63), (64) into (51) to get

$$(\mathbf{65}) \qquad \sup_{v_0,\dots,v_n\in\mathcal{V}} \frac{Q_{\mathfrak{m}}(v_0)}{2}M_c\big(\Delta G_N(v_0), \Delta G_N(v_1),\dots,\Delta G_N(v_n)\big)$$

$$+ \left[\sum_{h=2}^{3n} O\big(N^{(h+n)(1/2+\varepsilon)+1-h-1/6}\big)\right] + O\big(N^{(n+1)(1/2+\varepsilon)-1-1/6}\big)$$

$$+ O\big(N^{n(1/2+\varepsilon)-1-1/6}\big) = O(1).$$

In the same way as (55) implied (61), the bound (65) implies

$$(\mathbf{66}) \qquad \sup_{v_1,\dots,v_k\in\mathcal{V}} \mathbf{E}\left[\prod_{a=1}^{k}\big|\Delta G_N(v_a)\big|\right] = O\big(N^{k(1/2+\varepsilon)-\frac{k}{k+1}(1/2-\varepsilon)-1/6}\big) + O(1)$$

$$= O\big(N^{k(1/2+\varepsilon)-2\cdot 1/6}\big) + O(1), \quad k=1,2,3,\dots.$$

Repeating the same argument $m-2$ more times, we improve (66) to

$$(\mathbf{67}) \qquad \sup_{v_1,\dots,v_k\in\mathcal{V}} \mathbf{E}\left[\prod_{a=1}^{k}\big|\Delta G_N(v_a)\big|\right] = O\big(N^{k(1/2+\varepsilon)-m/6}\big) + O(1), \quad k=1,2,3,\dots.$$

Since m is arbitrary, this implies

$$(\mathbf{68}) \qquad \sup_{v_1,\dots,v_k\in A} \mathbf{E}\left[\prod_{a=1}^{k}\big|\Delta G_N(v_a)\big|\right] = O(1), \quad k=1,2,\dots,$$

and finishes the proof of Proposition 2.18. $\qquad\qquad\square$

Proof of Theorem 2.5. — Take (47) and observe that the bounds of Section 2.5 imply that all the terms except for the second line and the last line are negligible as $N \to \infty$. The second line of (47) is precisely the left-hand side of (21), while the last line of (47) is the right-hand side of (21). ☐

3. General setup

3.1. *Definition of the system.* — Our next goal is to generalize the arguments of the previous section to a much more general setting of a multi-cut fixed filling fractions model with fixed parameter $\theta > 0$ and general weight $w(x)$. Informally, we want to consider measures of the form

$$\prod_{1 \le i < j \le N} \frac{\Gamma(\ell_j - \ell_i + 1)\Gamma(\ell_j - \ell_i + \theta)}{\Gamma(\ell_j - \ell_i)\Gamma(\ell_j - \ell_i + 1 - \theta)} \prod_{i=1}^{N} w(\ell_i; N)$$

on ordered N-tuples $\ell_1 < \ell_2 < \cdots < \ell_N$ referred to as positions of N particles and satisfying two additional constraints. First, the particles are separated into k groups, and particles in each group must belong to its own interval of the real line. Second, if ith and $(i+1)$st particles are in the same group, then $\ell_{i+1} - \ell_i \in \{\theta, \theta+1, \theta+2, \dots\}$.

For instance, if we have a single group ($k = 1$), then after defining λ_i through $\ell_i = \lambda_i + \theta i$, the constraint boils to down to saying that all λ_i are integers and they satisfy $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_N$.

Let us give precise definitions for general k. The model depends on an integer parameter $N = 1, 2, \dots$ and amounts to fixing for each N a probability distribution on certain N-point subsets of $\mathbf{R}$.

We fix an integer $k = 1, 2, \dots$, whose meaning is the number of segments in the support of the measure. For each $N = 1, 2, \dots$ we take k integers $n_1(N), \dots, n_k(N)$, such that $\sum_{i=1}^{k} n_i(N) = N$ and k disjoint intervals $(a_1(N), b_1(N)), \dots, (a_k(N), b_k(N))$ of the real line ordered from left to right.[7]

We assume that $b_i(N) + \theta \le a_{i+1}(N)$ for $i = 1, \dots, k-1$. The numbers $a_i(N), b_i(N)$ must also satisfy the conditions

(**69**) $b_i(N) - \theta n_i(N) - a_i(N) \in \mathbf{Z}.$

Further, the number $n_i(N)$ counts the number of the particles in the ith interval; to make this statement precise we define the sets of indices $I_j \subset \{1, \dots, N\}, j = 1, \dots, k$, via

$$I_j = \left\{ i \in \mathbf{Z} \;\middle|\; \sum_{m=1}^{j-1} n_m(N) < i \le \sum_{m=1}^{j} n_m(N) \right\}.$$

We also set I_j^+ and I_j^- to be the maximal and minimal elements of I_j, respectively.

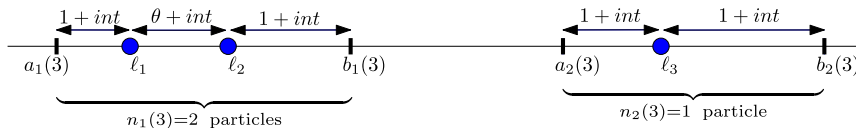[7] For a generalization to the case of infinite support see Sections 9.3, 9.4.

FIG. 2. — The state space for $N = 3$, $k = 2$. Numbers *int* indicate various (possibly different) nonnegative integers

Definition **3.1.** — *The state space* $\mathbf{W}_N^\theta$ *consists of* N-*tuples* $\ell_1 < \ell_2 < \cdots < \ell_N$ *such that for each* $j = 1, \ldots, k$:

(1) *If* $i = I_j^-$, *then* $\ell_i - a_i(N) \in \mathbf{Z}_{>0}$.
(2) *If* $i = I_j^+$, *then* $b_i(N) - \ell_i \in \mathbf{Z}_{>0}$.
(3) *If* $i \in I_j$, *but* $i \neq I_j^+$, *then* $\ell_{i+1} - \ell_i \in \{\theta, \theta + 1, \theta + 2, \ldots\}$.

Note that the conditions of Definition 3.1 imply that for every i from I_j, we have $\ell_i \in [a_i(N) + 1, b_i(N) - 1]$. An example of a configuration from $\mathbf{W}_N^\theta$ is shown in Figure 2.

We also take a *positive* weight function $w(x; N)$ for $x \in \bigcup_{i=1}^k [a_i(N) + 1, b_i(N) - 1]$ and define a probability measure $\mathbf{P}_N$ on $\mathbf{W}_N^\theta$ given by

$$(\mathbf{70}) \qquad \mathbf{P}_N(\ell_1, \ldots, \ell_N) = \frac{1}{Z_N} \prod_{1 \leq i < j \leq N} \frac{\Gamma(\ell_j - \ell_i + 1)\Gamma(\ell_j - \ell_i + \theta)}{\Gamma(\ell_j - \ell_i)\Gamma(\ell_j - \ell_i + 1 - \theta)} \prod_{i=1}^N w(\ell_i; N),$$

where Z_N is a normalizing constant which we will refer to as the *partition function*.

3.2. *List of regularity assumptions.* — Our ultimate goal is to study the asymptotics of the measures $\mathbf{P}_N$ as $N \to \infty$. For that we need to assume that the weights $w(x; N)$, as well as all other data specified in Section 3.1, depend on N in a regular way. Let us present all the technical assumptions that we impose on the data.

Assumption **1.** — *We require that for each* $i = 1, \ldots, k$, *as* $N \to \infty$

$$a_i(N) = N\hat{a}_i + O(\ln(N)), \quad b_i(N) = N\hat{b}_i + O(\ln(N)),$$

$$\hat{a}_1 < \hat{b}_1 < \hat{a}_2 < \cdots < \hat{a}_k < \hat{b}_k.$$

We require that $w(x; N)$ *in the intervals* $[a_i(N) + 1, b_i(N) - 1]$, $i = 1, \ldots, k$, *has the form*

$$w(x; N) = \exp\left(-NV_N\left(\frac{x}{N}\right)\right)$$

for a function V_N *that is continuous in the intervals* $[a_i(N) + 1, b_i(N) - 1]$ *and such that*

$$(\mathbf{71}) \qquad V_N(u) = V(u) + O\left(\frac{\ln(N)}{N}\right)$$

uniformly over $x = u \cdot \mathrm{N}$ *in the intervals* $[a_i(\mathrm{N}) + 1, b_i(\mathrm{N}) - 1]$. *The function* $\mathrm{V}(u)$ *is differentiable and the following bound holds for a constant* $\mathrm{C} > 0$

$$(\textbf{72}) \qquad \left| \mathrm{V}'(u) \right| \leq \mathrm{C} \left[1 + \sum_{i=1}^{k} \left(\left| \ln(u - \hat{a}_i) \right| + \left| \ln(u - \hat{b}_i) \right| \right) \right].$$

Remark **3.2.** — We believe that the assumption on the remainders can be weakened with minor changes in all the further statements and proofs. However, we do not pursue this direction due to lack of natural examples.

For the filling fractions $n_i(\mathrm{N})$ we make a weaker assumption, as it might be important for future applications, cf. [BoGu2].

Assumption **2.** — *There exists a constant* $\mathrm{C} > 0$ *such that for all* $1 \leq i \leq k$ *and all large enough* N *we have*

$$\mathrm{C} < \frac{n_i(\mathrm{N})}{\mathrm{N}} < \theta^{-1}(\hat{b}_i - \hat{a}_i) - \mathrm{C}.$$

Note that our definition of the state space $\mathbf{W}_{\mathrm{N}}^{\theta}$ implies that $\frac{n_i(\mathrm{N})}{\mathrm{N}} < \theta^{-1}(\hat{b}_i - \hat{a}_i) + o(1)$.

Introduce the notation

$$\frac{n_i}{\mathrm{N}} = \hat{n}_i, \quad i = 1, \dots, k.$$

Note that the numbers $\hat{n}_i$ *still depend on* N. However, we will hide this dependence from our notations. It is important here that all the limiting values as well as all the remainders in what follows will be uniform over $\hat{n}_i$ satisfying Assumption 2.

The next two assumptions deal with analytic properties of the weight $w(x; \mathrm{N})$ and the equilibrium measure μ. We fix an open set $\mathcal{M}_{\mathrm{N}} \subset \mathbf{C}$, such that $\bigcup_{i=1}^{k}[a_i(\mathrm{N}), b_i(\mathrm{N})] \subset \mathcal{M}_{\mathrm{N}}$.

Assumption **3.** — *There exist a pair of analytic in* $x \in \mathcal{M}_{\mathrm{N}}$ *functions* $\phi_{\mathrm{N}}^{+}(x)$, $\phi_{\mathrm{N}}^{-}(x)$ *such that*

$$\frac{w(x; \mathrm{N})}{w(x - 1; \mathrm{N})} = \frac{\phi_{\mathrm{N}}^{+}(x)}{\phi_{\mathrm{N}}^{-}(x)}.$$

Moreover,

$$\phi_{\mathrm{N}}^{\pm}(x) = \phi^{\pm}\left(\frac{x}{\mathrm{N}}\right) + \frac{1}{\mathrm{N}}\varphi_{\mathrm{N}}^{\pm}\left(\frac{x}{\mathrm{N}}\right) + \mathrm{O}\left(\frac{1}{\mathrm{N}^2}\right)$$

uniformly over x/N *in compact subsets of an open set* $\mathcal{M}_{\mathrm{N}}$, *which contains the union of the intervals* $[\hat{a}_i, \hat{b}_i]$, $i = 1, \dots, k$. *All the aforementioned functions are holomorphic in* $\mathcal{M}_{\mathrm{N}}$ *and functions* $\varphi_{\mathrm{N}}^{\pm}$ *are uniformly bounded as* $\mathrm{N} \to \infty$.

Remark **3.3.** — In the case when $V_N(x)$ in Assumption 1 is smooth and *uniformly* converges to V together with its derivative in a neighborhood of a point x, we have

$$(\textbf{73}) \qquad \exp\left(-\frac{\partial}{\partial x}V(x)\right) = \frac{\phi^+(x)}{\phi^-(x)}.$$

Indeed, in this case,

$$\begin{aligned}
\frac{\phi^+(x)}{\phi^-(x)} &= \lim_{N\to\infty} \frac{w(Nx; N)}{w(Nx-1; N)} \\
&= \lim_{N\to\infty} \exp\left(-N\left(V_N(x) - V_N(x-1/N)\right)\right) \\
&= \lim_{N\to\infty} \exp\left(-N\left(\int_{x-1/N}^{x} V_N'(y)dy\right)\right) = \exp\left(-\frac{\partial}{\partial x}V(x)\right).
\end{aligned}$$

Recall that the equilibrium measure μ with the density $\mu(x)$ encodes the Law of Large Numbers for $\mathbf{P}_N$ stated in Theorem 1.2 and discussed in more details in Section 5. A convenient way of working with the equilibrium measure is through its Stieltjes transform $G_\mu(z)$ defined through

$$(\textbf{74}) \qquad G_\mu(z) = \int_{-\infty}^{\infty} \frac{\mu(x)dx}{z-x}.$$

The following two functions $R_\mu(z)$, $Q_\mu(z)$ are important for our asymptotic study, cf. Section 2.1:

$$(\textbf{75}) \qquad R_\mu(z) := \phi^-(z)\exp\left(-\theta G_\mu(z)\right) + \phi^+(z)\exp\left(\theta G_\mu(z)\right),$$

$$(\textbf{76}) \qquad Q_\mu(z) := \phi^-(z)\exp\left(-\theta G_\mu(z)\right) - \phi^+(z)\exp\left(\theta G_\mu(z)\right).$$

We explain in Section 5 that $R_\mu(z)$ is analytic, while $Q_\mu(z)$ is a branch of a two-valued analytic function which is the square-root of a function holomorphic in $\mathcal{M}_N$. An important technical ingredient of our method is a restriction on its zeros, as is summarized in the following assumption.

Assumption **4.** — *We require that for each large enough* N *and corresponding* Q_μ *(which depends on* N *through the filling fractions* $\hat{n}_i$ *in the definition of* μ*), there exists a function* $H(z)$ *holomorphic in* $\mathcal{M}_N$ *and numbers* $\{\alpha_i, \beta_i\}_{i=1}^k$ *such that*

- $\hat{a}_i \le \alpha_i < \beta_i \le \hat{b}_i$, $i = 1, \dots, k$;
- $Q_\mu(z) = H(z)\prod_{i=1}^{k} \sqrt{(z-\alpha_i)(z-\beta_i)}$, *where the branch of the square root is such that* $\sqrt{(z-\alpha_i)(z-\beta_i)} \sim z$ *when* $z \to \infty$;
- $H(z) \ne 0$ *for all* $z \in \bigcup_{i=1}^{k}[\hat{a}_i, \hat{b}_i]$.

We remark that Assumption 4 does not describe a generic case. In particular, it implies that there is precisely one interval of support of $\mu(x)$ in each interval $[\hat{a}_i, \hat{b}_i]$. The authors are not aware of simple ways to check such property by examining the potential $V(x)$. Nevertheless, many natural models arising in the applications satisfy Assumption 4. We demonstrate this general principle by considering several examples in Section 9.

Finally, we need a simple vanishing assumption. It is convenient to work with it, yet we later show in Section 8 how it can be relaxed.

Assumption 5. — *For all $i = 1, \ldots, k$, we have $\phi_N^-(a_i(N) + 1) = \phi_N^+(b_i(N)) = 0$.*

4. Nekrasov's equation

The main tool for our study of the probability distributions $\mathbf{P}_N$ from the last section is a statement, which is essentially due to Nekrasov [N], [NP], [NS]. Its affine and q-versions are given in the following two theorems.

Theorem 4.1. — *Let $\mathbf{P}_N$ be a distribution on N-tuples $(\ell_1, \ldots, \ell_N) \in \mathbf{W}_N^\theta$ as in the previous section. Suppose that*

$$(\mathbf{77}) \qquad \frac{w(x; N)}{w(x - 1; N)} = \frac{\phi_N^+(x)}{\phi_N^-(x)},$$

and for all $i = 1, \ldots, k$, we have $\phi_N^-(a_i(N) + 1) = \phi_N^+(b_i(N)) = 0$. Define

$$(\mathbf{78}) \qquad R_N(\xi) = \phi_N^-(\xi) \cdot \mathbf{E}_{\mathbf{P}_N}\left[\prod_{i=1}^N \left(1 - \frac{\theta}{\xi - \ell_i}\right)\right]$$

$$+ \phi_N^+(\xi) \cdot \mathbf{E}_{\mathbf{P}_N}\left[\prod_{i=1}^N \left(1 + \frac{\theta}{\xi - \ell_i - 1}\right)\right].$$

If $\phi_N^\pm(\xi)$ are holomorphic in a domain $\mathcal{M}_N \subset \mathbf{C}$, then so is $R_N(\xi)$. Moreover, if $\phi_N^\pm(\xi)$ are polynomials of degree at most d, then so is $R_N(\xi)$.

Proof. — The possible singularities of $R_N(\xi)$ are simple poles arising from the denominator of the expression under expectation $\mathbf{E}_{\mathbf{P}_N}$ in (78). Let us compute a residue at such a pole m.

The expectation $\mathbf{E}_{\mathbf{P}_N}$ in (78) is a sum over all $(\ell_1, \ldots, \ell_N) \in \mathbf{W}_N^\theta$. Such a configuration contributes to the residue if $\ell_i = m$ or $\ell_i = m + 1$ for some $i = 1, \ldots, N$.

We separately analyze the contributions appearing from each $i = 1, \ldots, N$, which we now fix. According to definitions, the possible values for ℓ_i are $\{A, A+1, A+2, \ldots, B\}$

for certain A and B. Given a particle configuration $\ell = (\ell_1, \ldots, \ell_N)$, let ℓ^+ denote the configuration with ith coordinate increased by 1, and let ℓ^- denote the configuration with ith coordinate decreased by 1. Note that in principle, ℓ^+ (similarly ℓ^-) may fail to be in $\mathbf{W}_N^\theta$. However, in this case the formula for $\mathbf{P}_N(\ell^+)$ still applies, but gives zero.

Let us explain how the weight (70) of a configuration $(\ell_1, \ldots, \ell_N)$ changes when one coordinate is changed from $\ell_i = x$ to $\ell_i = x - 1$, i.e. we compute the ratio of the weights at $\ell_i = x - 1$ and at $\ell_i = x$ (all other coordinates are unchanged). The double product over $i < j$ in (70) produces factors (we denote $\ell_j = r$ here)

$$
\begin{aligned}
(\mathbf{79}) \qquad & \frac{\Gamma(r - x + 1)\Gamma(r - x + \theta)}{\Gamma(r - x)\Gamma(r - x + 1 - \theta)} \cdot \frac{\Gamma(r - x + 1)\Gamma(r - x + 2 - \theta)}{\Gamma(r - x + 2)\Gamma(r - x + 1 + \theta)} \\
& = \frac{(r - x)(r - x + 1 - \theta)}{(r - x + 1)(r - x + \theta)},
\end{aligned}
$$

if $i < j$, and the factor

$$
\begin{aligned}
(\mathbf{80}) \qquad & \frac{\Gamma(x - r + 1)\Gamma(x - r + \theta)}{\Gamma(x - r)\Gamma(x - r + 1 - \theta)} \cdot \frac{\Gamma(x - r - 1)\Gamma(x - r - \theta)}{\Gamma(x - r)\Gamma(x - r + \theta - 1)} \\
& = \frac{(x - r)(x - r + \theta - 1)}{(x - r - 1)(x - r - \theta)},
\end{aligned}
$$

if $i > j$. Note that (79) and (80) are two forms of the same rational expression.

Now take $m \in \{A, A + 1, \ldots, B\}$. The contribution to the residue of R_N at $z = m$, arising from the ith coordinate of configurations $(\ell_1, \ldots, \ell_N)$ is

$$
\begin{aligned}
(\mathbf{81}) \qquad & -\theta \sum_{\ell \in \mathbf{W}_N^\theta | \ell_i = m} \phi_N^-(m)\mathbf{P}_N(\ell_1, \ldots, \ell_N) \prod_{j \neq i} \left(1 - \frac{\theta}{m - \ell_j} \right) \\
& + \theta \sum_{\ell \in \mathbf{W}_N^\theta | \ell_i = m - 1} \phi_N^+(m)\mathbf{P}_N(\ell_1, \ldots, \ell_N) \prod_{j \neq i} \left(1 + \frac{\theta}{m - \ell_j - 1} \right).
\end{aligned}
$$

Note the difference between two summation sets in (81). If ℓ_i is *not* the smallest particle in an interval $[a_h(N) + 1, b_h(N) - 1]$, then the first sum contains the terms with $\ell_{i-1} = m - \theta$, while the second one does not. However, each such term in the first sum is actually zero. Also if ℓ_i is the smallest particle and $m = A = a_h(N) + 1$, then the second sum is empty as ℓ_i is never m. But we also know that $\phi_h^-(m) = \phi_h^-(a_h(N) + 1) = 0$ and the first sum vanishes as well. Similar considerations apply to two cases whether ℓ_i is the largest particle or not. We conclude, that it suffices to study the case when there is one-to-one correspondence between terms of two sums in (81).

Using (79), (80), and (77) we see that

$$\phi_N^-(m)\mathbf{P}_N(\ell_1, \ldots \ell_{i-1}, m, \ell_{i+1}, \ldots, \ell_N)\prod_{j\neq i}\left(1 - \frac{\theta}{m - \ell_j}\right)$$

$$= \phi_N^+(m)\mathbf{P}_N(\ell_1, \ldots \ell_{i-1}, m - 1, \ell_{i+1}, \ldots, \ell_N)\prod_{j\neq i}\left(1 + \frac{\theta}{m - \ell_j - 1}\right).$$

We conclude that for each μ, the terms with $\ell = \mu$ and $\ell = \mu^+$ (or $\ell = \mu^-$ and $\ell = \mu$) in the first and second sum in (81) cancel out and the total residue is zero.

For the polynomiality statement it suffices to notice that if $\phi_N^\pm(\xi)$ are polynomials of degree at most d, then $R_N(\xi)$ is an entire function which grows as $O(\xi^d)$ as $\xi \to \infty$. Hence, by Liouville's theorem $R_N(\xi)$ is a polynomial. $\qquad\square$

The proof of Theorem 4.1 reveals that it admits a natural q-deformation. Recall the definition of q-Gamma function Γ_q:

$$\Gamma_q(x) = (1 - q)^{1-x}\frac{(q; q)_\infty}{(q^x; q)_\infty},$$

where

$$(a; q)_\infty = \prod_{n=0}^{\infty}\left(1 - aq^n\right).$$

In the same framework of Section 3, define a q-deformation of $\mathbf{P}_N$ through

$$(82) \qquad \mathbf{P}_N^q(\ell_1, \ldots, \ell_N) = \frac{1}{Z_N^q}\prod_{1\leq i<j\leq N}q^{-\theta(\ell_j-\ell_i)}\frac{\Gamma_q(\ell_j - \ell_i + 1)\Gamma_q(\ell_j - \ell_i + \theta)}{\Gamma_q(\ell_j - \ell_i)\Gamma_q(\ell_j - \ell_i + 1 - \theta)}$$

$$\times \prod_{i=1}^{N}w(\ell_i; N).$$

Theorem **4.2.** — *Let* $\mathbf{P}_N^q$ *be a distribution on* N-*tuples* $(\ell_1, \ldots, \ell_N) \in \mathbf{W}_N^\theta$ *as above. Suppose that*

$$\frac{w(x; N)}{w(x - 1; N)} = \frac{\phi_N^+(x)}{\phi_N^-(x)},$$

and for all $i = 1, \ldots, k$, *we have* $\phi_N^-(a_i(N) + 1) = \phi_N^+(b_i(N)) = 0$. *Define*

$$(83) \qquad R_N^q(\xi) = \phi_N^-(\xi) \cdot \mathbf{E}_{\mathbf{P}_N^q}\left[\prod_{i=1}^{N}\left(q^{\frac{\theta}{2}}\frac{1 - q^{\xi-\ell_i-\theta}}{1 - q^{\xi-\ell_i}}\right)\right]$$

$$+ \phi_N^+(\xi) \cdot \mathbf{E}_{\mathbf{P}_N^q}\left[\prod_{i=1}^{N}\left(q^{-\frac{\theta}{2}}\frac{1 - q^{\xi-\ell_i-1+\theta}}{1 - q^{\xi-\ell_i-1}}\right)\right].$$

If $\phi_N^\pm(\xi)$ are holomorphic in a domain $\mathcal{M}_N \subset \mathbf{C}$, then so is $R_N^q(\xi)$. Moreover, if $\phi_N^\pm(\xi)$ are polynomials of degree at most d, then so is $R_N(\xi)$.

Proof. — The proof is the same as for Theorem 4.1. The only new ingredient is a q-deformation of (79), (80), which now reads

$$
\textbf{(84)} \qquad q^{-\theta(r-x)} \frac{\Gamma_q(r-x+1)\Gamma_q(r-x+\theta)}{\Gamma_q(r-x)\Gamma_q(r-x+1-\theta)}
$$

$$
\times\, q^{\theta(r-x+1)} \frac{\Gamma_q(r-x+1)\Gamma_q(r-x+2-\theta)}{\Gamma_q(r-x+2)\Gamma_q(r-x+1+\theta)}
$$

$$
= q^\theta \frac{(1-q^{r-x})(1-q^{r-x+1-\theta})}{(1-q^{r-x+1})(1-q^{r-x+\theta})}
$$

and

$$
\textbf{(85)} \qquad q^{-\theta(x-r)} \frac{\Gamma_q(x-r+1)\Gamma_q(x-r+\theta)}{\Gamma_q(x-r)\Gamma_q(x-r+1-\theta)}
$$

$$
\times\, q^{\theta(x-r-1)} \frac{\Gamma_q(x-r-1)\Gamma_q(x-r-\theta)}{\Gamma_q(x-r)\Gamma_q(x-r+\theta-1)}
$$

$$
= q^{-\theta} \frac{(1-q^{x-r})(1-q^{x-r+\theta-1})}{(1-q^{x-r-1})(1-q^{x-r-\theta})}.
$$

One readily recognizes the same expression in (84) and (85). We further conclude that

$$
\phi_N^-(m)\mathbf{P}_N^q(\ell_1, \ldots \ell_{i-1}, m, \ell_{i+1}, \ldots, \ell_N) \prod_{j\neq i}\left(q^{\frac{\theta}{2}} \frac{1-q^{m-\ell_j-\theta}}{1-q^{m-\ell_j}} \right)
$$

$$
= \phi_N^+(m)\mathbf{P}_N^q(\ell_1, \ldots \ell_{i-1}, m-1, \ell_{i+1}, \ldots, \ell_N) \prod_{j\neq i}\left(q^{-\frac{\theta}{2}} \frac{1-q^{m-\ell_j-1+\theta}}{1-q^{m-\ell_j-1}} \right),
$$

and therefore there is a cancellation of the poles. $\qquad\square$

5. Law of large numbers

5.1. *Limit shape.* — Assumptions 1 and 2 guarantee that a certain law of large numbers for the measures $\mathbf{P}_N$ holds as $N \to \infty$. To state it we introduce a random probability measure μ_N on $\mathbf{R}$ via

$$
\textbf{(86)} \qquad \mu_N = \frac{1}{N} \sum_{i=1}^{N} \delta\left(\frac{\ell_i}{N}\right), \quad (\ell_1, \ldots, \ell_N) \text{ is } \mathbf{P}_N\text{-distributed.}
$$

The measure μ_N is often referred to as the *empirical measure* of point configuration $\ell_1, \ldots, \ell_N$, cf. (12). Note that our definitions imply the condition $\ell_{i+1} - \ell_i \geq \theta$, which shows that for any interval $[p, q]$, its μ_N-measure is bounded from above by $\theta^{-1}(q - p + N^{-1})$.

The first order asymptotic behavior of measures μ_N can be understood through a variational problem. For a probability measure ρ supported on $\bigcup_i [\hat{a}_i, \hat{b}_i]$, define

$$(\mathbf{87}) \qquad I_V[\rho] = \theta \iint_{x \neq y} \ln|x - y| \rho(dx)\rho(dy) - \int_{-\infty}^{\infty} V(x)\rho(dx).$$

*Lemma **5.1.*** — *Let Θ be the set of absolutely continuous probability measures $\rho(x)dx$ supported on $\bigcup_i [\hat{a}_i, \hat{b}_i]$, whose density is between 0 and θ^{-1} and such that*

$$\int_{\hat{a}_i}^{\hat{b}_i} \rho(x)dx = \hat{n}_i, \quad 1 \leq i \leq k,$$

where $0 < \hat{n}_i < \theta^{-1}(\hat{b}_i - \hat{a}_i)$, $i = 1, \ldots, k$, are such that $\sum_{i=1}^{k} \hat{n}_i = 1$. Then the functional I_V has a unique maximum $\mu(x)dx$ on Θ.

*Remark **5.2.*** — The maximizer of I_V is called the *equilibrium measure*.

Proof of Lemma 5.1. — Let us equip Θ with weak topology, i.e. the topology of pointwise convergence for the distribution functions. Then Θ is compact.

Observe that the functional I_V is continuous in the weak topology on Θ; here it is crucial that the measures in Θ have density between 0 and θ^{-1}. Therefore, I_V attains its maximum on Θ. It remains to show that such a maximum is unique.

For that we note that I_V is strictly concave, i.e. for any $\rho, \rho' \in \Theta$ and any $0 < t < 1$

$$I_V\big[t\rho + (1 - t)\rho'\big] > tI_V[\rho] + (1 - t)I_V\big[\rho'\big].$$

Indeed, the linear part of I_V is concave by the definition. The quadratic part is negatively-definite due to the formula (36) for it, therefore, it is strictly concave.

Since a strictly concave functional cannot have more than one maximum on a convex set, we are done. $\qquad\square$

*Theorem **5.3.*** — *Under Assumptions 1 and 2 the random measures μ_N concentrate (in probability) near $\mu(x)dx$ of Lemma 5.1. More precisely, for each Lipshitz function $f(x)$ defined in a real neighborhood of $\bigcup_{i=1}^{k} [\hat{a}_i, \hat{b}_i]$, and each $\varepsilon > 0$ the random variables*

$$N^{1/2 - \varepsilon} \left| \int_{-\infty}^{\infty} f(x)\mu_N(dx) - \int_{-\infty}^{\infty} f(x)\mu(x)dx \right|$$

converge to 0 in probability and in the sense of moments.

Remark **5.4.** — Since $\mu(x)$ depends on $\hat{n}_i$, and the latter depend on N, so is the former. When $\hat{n}_i$ converge as $N \to \infty$, this dependence can be removed by using the continuity of $\mu(x)$ in $\hat{n}_i$, established below in Proposition 5.8.

Analogues of Theorem 5.3 are known in the literature, cf. [1], [AGZ], [Fe]. Our proof of Theorem 5.3 relies on a different characterization for μ maximizing the functional $I_V[\cdot]$.

The restriction $0 \le \mu(x) \le \theta^{-1}$ leads to the subdivision of **R** into three types of regions (we borrow the terminology from [BKMM]):

- Maximal (with respect to inclusion) closed connected intervals where $\mu(x) = 0$ are called *voids*.
- Maximal open connected intervals where $0 < \mu(x) < \theta^{-1}$ are called *bands*.
- Maximal closed connected intervals where $\mu(x) = \theta^{-1}$ are called *saturated regions*.

In a related context of random tilings and periodically-weighted dimers the voids and saturated regions are usually called *frozen*, while bands are *liquid regions*.

We further define the *effective potential* $F_V(x)$ through (cf. (34))

$$F_V(x) = 2\theta \int_{-\infty}^{\infty} \ln|x - y|\mu(y)dy - V(x).$$

Lemma **5.5.** — *There exist k constants $f_1, \ldots, f_k$, such that*

$$(\mathbf{88}) \qquad F_V(x) - f_i \le 0, \quad \text{for all } x \text{ in voids in } \left[\hat{a}_i, \hat{b}_i\right];$$

$$(\mathbf{89}) \qquad F_V(x) - f_i \ge 0, \quad \text{for all } x \text{ in saturated regions in } \left[\hat{a}_i, \hat{b}_i\right];$$

$$(\mathbf{90}) \qquad F_V(x) - f_i = 0, \quad \text{for all } x \text{ in bands in } \left[\hat{a}_i, \hat{b}_i\right].$$

Proof. — This characterization is readily obtained from the variation of the functional $I_V[\cdot]$, cf. [ST], [DS]. $\qquad \square$

Further fix a parameter $p > 2$ and let $\tilde{\mu}_N$ denote the convolution of the empirical measure μ_N with uniform measure on the interval $[0, N^{-p}]$. The following statement is a generalization of Proposition 2.16. Let us emphasize that the measure μ here and below depends on the filling fractions $\hat{n}_i = n_i(N)/N$, and thus changes with N.

Proposition **5.6.** — *Let μ be the maximizer of Lemma 5.1. Then there exists $C \in \mathbf{R}$ such that for all $\gamma > 0$ and all N we have*

$$\mathbf{P}_N\big(\mathcal{D}(\tilde{\mu}_N, \mu) \ge \gamma\big) \le \exp\big(CN\ln^2(N) - \gamma^2 N^2\big),$$

where $\mathcal{D}$ was defined in (35), (36).

Proof. — We start by analyzing the asymptotic of the formula for $\mathbf{P}_N$. We observe that the definition of $I_V[\cdot]$ by (87) makes sense for discrete measures. It is important here that we integrate only over $x \neq y$ in (87) as otherwise the integral would be infinite. Using for the double product in (70) the following corollary of the Stirling's formula

(**91**)
$$\frac{\Gamma(h+\theta)}{\Gamma(h)} = h^\theta\left(1 + O\left(h^{-1}\right)\right), \quad h \to +\infty,$$

we can write for every $(\ell_1, \ell_2, \ldots, \ell_N) \in \mathbf{W}_N^\theta$

(**92**)
$$\mathbf{P}_N(\ell_1, \ldots, \ell_N)$$
$$= \frac{\exp(2\theta N(N-1)\ln(N) + N^2 I_V[\mathrm{mes}[\ell_1, \ldots, \ell_N]] + O(N\ln(N)))}{Z_N},$$

where

$$\mathrm{mes}[\ell_1, \ldots, \ell_N] = \frac{1}{N} \sum_{i=1}^N \delta\left(\frac{\ell_i}{N}\right).$$

Let us obtain a lower bound for the partition function Z_N in (92). For that let x_i, $i = 1, \ldots, N$ be quantiles of μ defined through

$$\int_0^{x_i} \mu(x)dx = \frac{i - 1/2}{N}, \quad 1 \leq i \leq N.$$

Since $\mu(x) \leq \theta^{-1}$, $\theta(x_{i+1} - x_i) \geq 1/N$. Therefore, using the asymptotics of $a_i(N)$ and $b_i(N)$ of Assumption 1 we conclude that there exists an absolute constant U (independent of N) and a configuration $(\hat{\ell}_1, \ldots, \hat{\ell}_N) \in \mathbf{W}_N^\theta$ such that

(**93**)
$$\left|Nx_i - \hat{\ell}_i\right| \leq U$$

for all $1 \leq i \leq N$ except for $O(\ln(N))$ ones. Arguing as in Proposition 2.16, we write

$$\frac{Z_N}{N^{2\theta N(N-1)}} \geq \exp\left(N^2 I_V\left[\mathrm{mes}\left[\hat{\ell}_1, \ldots, \hat{\ell}_N\right]\right] + O(N\ln N)\right)$$
$$= \exp\left(N^2 I_V[\mu] + O\left(N\ln^2(N)\right)\right).$$

Note that the remainder in the last formula is $O(N\ln^2(N))$ instead of $O(N\ln(N))$ in the proof of Proposition 2.16. This is due to an additional error produced by the ℓ_i for which (93) does not hold: there are $O(\ln(N))$ of such ℓ_i and each of them produces (at most) $O(N\ln(N))$ of error.

From here on the proof literally repeats that of Proposition 2.16. $\qquad\square$

Corollary **5.7.** — *Recall the notations $\|g\|_{1/2}$ and $\|g\|_{\mathrm{Lip}}$ of Corollary 2.17. Fix any $p > 2$ and let $\mu(x)dx$ be the maximizer of Lemma 5.1. Then there exists $C \in \mathbf{R}$ such that for all $\gamma > 0$, all N and all f we have*

$$(94) \qquad \mathbf{P}_N\left(\left|\int_{\mathbf{R}} g(x)\mu_N(dx) - \int_{\mathbf{R}} g(x)\mu(x)dx\right| \geq \gamma\|g\|_{1/2} + \frac{\|g\|_{\mathrm{Lip}}}{N^p}\right)$$

$$\leq \exp\left(CN\ln^2(N) - \frac{\gamma^2 N^2}{2}\right).$$

Proof. — The proof is the same as for Corollary 2.17. $\qquad\square$

Proof of Theorem 5.3. — The desired statement follows from Corollary 5.7 with $\gamma = q \cdot N^{\varepsilon-1/2}$, $q > 0$. $\qquad\square$

Recall that the set of measures Θ and the equilibrium measure μ of Lemma 5.1 implicitly depend on many parameters. For the next statement we reconstruct a part of this dependence in the notations. For $\hat{n} = (\hat{n}_1, \ldots, \hat{n}_k)$ let $\mu^{\hat{n}}$ be the maximizer of I_V over $\Theta^{\hat{n}}$, the set of probability measures μ absolutely continuous with respect to Lebesgue measure, with density bounded above by $1/\theta$ and with mass $\mu([\hat{a}_i, \hat{b}_i]) = \hat{n}_i$, $\sum_{i=1}^k \hat{n}_i = 1$.

Proposition **5.8.** — *Assume that $0 < \hat{n}_i < \theta^{-1}(\hat{b}_i - \hat{a}_i)$ for all $i = 1, \ldots, k$ and that the measure $\mu^{\hat{n}_i}$ has at least one band in each of the intervals $[\hat{a}_i, \hat{b}_i]$, $i = 1, \ldots, k$. Then there exists a finite constant $C_{\hat{n}}$ such that for $\hat{n}'$ close enough to $\hat{n}$*

$$(95) \qquad \mathcal{D}(\mu^{\hat{n}}, \mu^{\hat{n}'}) \leq C_{\hat{n}}\|\hat{n} - \hat{n}'\|_\infty.$$

Remark **5.9.** — Arguing as in the proof of Corollary 2.17, one deduces from (95) that for any smooth function $g(x)$, the averages $\int g(x)\mu^{\hat{n}}(x)dx$ and $\int g(x)\mu^{\hat{n}'}(x)dx$ are close. Their difference goes to zero as $\hat{n}' \to \hat{n}$.

Remark **5.10.** — In all the applications that we present, the assumption that the measure $\mu^{\hat{n}_i}$ has at least one band in each of the intervals $[\hat{a}_i, \hat{b}_i]$, $i = 1, \ldots, k$, is satisfied automatically, see the end of Section 9.1 for a general argument in this direction.

Proof of Proposition 5.8. — By definition, for any probability measure ν with density $\nu(x)$

$$I_V(\nu) = I_V(\mu^{\hat{n}}) - \theta\mathcal{D}^2(\nu, \mu^{\hat{n}})$$
$$- \int\left(V(x) - 2\theta\int\log|x-y|\mu^{\hat{n}}(y)dy\right)(\nu(x) - \mu^{\hat{n}}(x))dx$$

$$= I_V(\mu^{\hat{n}}) - \theta \mathcal{D}^2(\nu, \mu^{\hat{n}}) + \int \tilde{F}_V(x)(\nu(x) - \mu^{\hat{n}}(x))dx$$

$$+ \sum_{i=1}^{k} f_i(\nu - \mu^{\hat{n}})([\hat{a}_i, \hat{b}_i])$$

where $\tilde{F}_V(x) = F_V(x) - f_i$ on $[\hat{a}_i, \hat{b}_i]$ is nonpositive on voids, nonnegative on saturated regions and vanishes on bands for $\mu^{\hat{n}}$, cf. (88)–(90). We next choose $\hat{n}' \neq \hat{n}$ and let ν be a probability measure with filling fractions $\hat{n}'$ so that

$$I_V(\nu) \leq I_V(\mu^{\hat{n}'}).$$

The above decomposition and $(\nu - \mu^{\hat{n}'})([\hat{a}_i, \hat{b}_i]) = 0$ implies

$$\theta \mathcal{D}^2(\mu^{\hat{n}'}, \mu^{\hat{n}}) - \int \tilde{F}_V(x)(\mu^{\hat{n}'}(x) - \mu^{\hat{n}}(x))dx$$

$$\leq \theta \mathcal{D}^2(\nu, \mu^{\hat{n}}) - \int \tilde{F}_V(x)(\nu(x) - \mu^{\hat{n}}(x))dx.$$

Observe that $\int \tilde{F}_V(x)(\nu(x) - \mu^{\hat{n}}(x))dx$ is nonpositive for any probability measure ν whose density is at most θ^{-1}, and vanishes if $\nu - \mu^{\hat{n}}$ is supported on the bands of $\mu^{\hat{n}}$. We assume that this is the case to deduce that

$$(\textbf{96}) \qquad \mathcal{D}^2(\mu^{\hat{n}'}, \mu^{\hat{n}}) \leq \mathcal{D}^2(\nu, \mu^{\hat{n}}).$$

Finally we choose $\nu = \mu^{\hat{n}} + \sum_i (\hat{n}'_i - \hat{n}_i) \frac{1_{B_i}}{|B_i|} dx$ where B_i denotes a subset in the interior of bands in $[\hat{a}_i, \hat{b}_i]$ so that on B_i we have $\delta \leq d\mu^{\hat{n}}/dx \leq 1/\theta - \delta$ for some small $\delta > 0$. Note that our assumption on the existence of bands in all the intervals $[\hat{a}_i, \hat{b}_i]$ implies that B_i is not empty. For $n'_i - n_i$ small enough (smaller than $\delta|B_i|$), $\nu \in \Theta^{\hat{n}'}$. Hence, we have

$$(\textbf{97}) \qquad \mathcal{D}^2(\nu, \mu^{\hat{n}}) = -\sum_{i,j=1}^{k} \frac{(\hat{n}'_i - \hat{n}_i)(\hat{n}'_j - \hat{n}_j)}{|B_i||B_j|} \int_{t \in B_i} \int_{s \in B_j} \ln|t - s| dt ds \leq C_{\hat{n}} \|\hat{n} - \hat{n}'\|_{\infty}^2$$

for a constant $C_{\hat{n}} > 0$. Combining (96) and (97) we obtain

$$\mathcal{D}^2(\mu^{\hat{n}'}, \mu^{\hat{n}}) \leq C_{\hat{n}} \|\hat{n} - \hat{n}'\|_{\infty}^2. \qquad \square$$

5.2. *Functions* R *and* Q. — We work with the equilibrium measure μ of density $\mu(x)$ in terms of its Stieltjes transform $G(z)$ defined through

$$(\textbf{98}) \qquad G_\mu(z) = \int_{-\infty}^{\infty} \frac{\mu(x)dx}{z - x}.$$

Observe that (98) makes sense for all z outside the support of $\mu(x)$ and $G_\mu(z)$ is holomorphic there.

We are going to extensively use the following two notations for any function $F(z)$ of a complex variable z:

$$F(z + \mathbf{i}0) := \lim_{\delta \to 0} F(z + \mathbf{i}\delta), \qquad F(z - \mathbf{i}0) := \lim_{\delta \to 0} F(z - \mathbf{i}\delta).$$

At every point x where $\mu(x)$ is continuous, it can be reconstructed by its Stieltjes transform via

(**99**) $$\mu(x) = \frac{1}{2\pi\mathbf{i}}\big(G_\mu(x + \mathbf{i}0) - G_\mu(x - \mathbf{i}0)\big), \quad x \in \mathbf{R}.$$

On the other hand, the equilibrium measure characterization (90) implies (cf. [DS], [ST]) that for x in a band of the equilibrium measure

(**100**) $$\theta\big(G_\mu(x + \mathbf{i}0) + G_\mu(x - \mathbf{i}0)\big) = \frac{\partial}{\partial x}V(x).$$

We also define the Stieltjes transform $G_N(z)$ of the prelimit empirical measure (86) through

(**101**) $$G_N(z) = \int_{-\infty}^{\infty} \frac{1}{z - x}\mu_N(dx) = \frac{1}{N}\sum_{i=1}^{N} \frac{1}{z - \ell_i/N},$$

$$(\ell_1, \ldots, \ell_N) \in \mathbf{W}_N^\theta \text{ is } \mathbf{P}_N\text{-distributed.}$$

The functions $R_\mu(z)$, $Q_\mu(z)$ defined in (75), (76) are important for our asymptotic study.

Proposition **5.11.** — *Under Assumptions 1, 2, 3, 5, $R_\mu(z)$ is holomorphic in $\mathcal{M}_N$.*

Proof of Proposition 5.11. — Since $G_\mu(z)$ is holomorphic everywhere outside the (real) support of the equilibrium measure $\mu(x)$, so is $R_\mu(z)$.

Further take any point x in the support of $\mu(x)$. Choose a simple contour γ in $\mathcal{M}_N \setminus [\hat{a}_i, \hat{b}_i]$ enclosing x. Observe that as $N \to \infty$ under the change of variables $\xi = Nz$, we have

$$\left(1 - \frac{\theta}{\xi - \ell_i}\right) = \left(1 - \frac{1}{N} \cdot \frac{\theta}{z - \ell_i/N}\right)$$

(**102**) $$= \exp\left(-\frac{1}{N} \cdot \frac{\theta}{z - \ell_i/N} + O\left(\frac{1}{N}\right)\right),$$

$$\left(1 + \frac{\theta}{\xi - \ell_i - 1}\right) = \left(1 + \frac{1}{N} \cdot \frac{\theta}{z - \ell_i/N - 1/N}\right)$$
$$= \exp\left(\frac{1}{N} \cdot \frac{\theta}{z - \ell_i/N} + O\left(\frac{1}{N}\right)\right).$$

Now fix $\hat{n}_i$, $i = 1, \ldots, k$ and choose the filling fractions $n_i(N)$ in such a way that $|n_i(N) - N\hat{n}_i| \leq 1$ for all N and i. We claim that $R_N(Nz)$ of Theorem 4.1 converges to $R_\mu(z)$ uniformly on γ. Indeed, the functions $\phi_N^\pm$ converge to $\phi^\pm$ by Assumption 3. By (102) and Theorem 5.3 the two expectations in the definition of $R_N(Nz)$ approximate $\exp(-\theta G_\nu(z))$ and $\exp(-\theta G_\nu(z))$, respectively, where ν is the maximizer of Lemma 5.1 with filling fractions $\frac{n_1(N)}{N}, \ldots, \frac{n_k(N)}{N}$. Since $\frac{n_i(N)}{N} \to \hat{n}_i$ as $N \to \infty$, Proposition 5.8 implies that $G_\nu(z)$ converges to $G_\mu(z)$.

Since uniform convergence on γ of holomorphic (everywhere inside the contour) functions $R_N(Nz)$ implies the same convergence inside the contour and the holomorphicity of the limit, we are done. $\qquad\square$

Remark **5.12.** — Alternatively, one can prove Proposition 5.11 by showing that $R_\mu(z)$ is continuous near real x via combining (100) with (73).

Despite their similar form, the analytic properties of the function $Q_\mu(z)$ are very different. Observe that

(**103**) $$\left(Q_\mu(z)\right)^2 = \left(R_\mu(z)\right)^2 - 4\phi^+(z)\phi^-(z).$$

Thus, $Q_\mu(z)$ is a branch of a two-valued analytic function which is the square-root of a function holomorphic in $\mathcal{M}_N$. Note that $Q_\mu(z)$ is analytic outside $\bigcup_{i=1}^k [\hat{a}_i, \hat{b}_i]$. On the other hand, combining (100) with (73) one observes that $Q_\mu(z)$ must have discontinuities in the bands of $\mu(x)$ and the endpoints of the bands should be its branching points. Therefore, $Q_\mu(z)$ has at least two branching points inside each interval $[\hat{a}_i, \hat{b}_i]$, which also have to be zeros of $Q_\mu(z)$. For our method it is important to assume that there are precisely two zeros in each interval, as is summarized in Assumption 4.

6. Second order expansion

The goal of this section is to prove a generalization of Theorem 2.5 in the setup of Section 3.

Theorem 5.3 implies that $G_N(z) - G_\mu(z)$ vanishes as $N \to \infty$ uniformly over z in compact subsets of $\mathbf{C} \setminus \bigcup_{i=1}^k [\hat{a}_i, \hat{b}_i]$. Moreover, since all the involved functions are analytic in z, we can infer also the uniform convergence of the derivatives.

Similarly to Section 2.2, we introduce a deformed version of $G_N(z)$. Take $2m$ parameters $\mathbf{t} = (t_1, \ldots, t_m)$, $\mathbf{v} = (v_1, \ldots, v_m)$ such that $v_a + t_a - y \neq 0$ for all $a = 1, \ldots, m$

and all $y \in \bigcup_{i=1}^{k} [\hat{a}_i, \hat{b}_i]$, and let the deformed distribution $\mathbf{P}_N^{t,v}$ be defined through

$$(\mathbf{104}) \qquad \mathbf{P}_N^{t,v}(\ell_1, \ldots, \ell_N) = \frac{Z_N}{Z_N(\mathbf{t}, \mathbf{v})} \cdot \mathbf{P}_N(\ell_1, \ldots, \ell_N) \cdot \prod_{i=1}^{N} \prod_{a=1}^{m} \left(1 + \frac{t_a}{v_a - \ell_i/N} \right),$$

where $\mathbf{P}_N$ was defined in (70). In general, $\mathbf{P}_N^{t,v}$ may be a *complex-valued* measure, but the normalizing constant $Z_N(\mathbf{t}, \mathbf{v})$ in (104) is chosen so that the total mass of $\mathbf{P}_N^{t,v}$ is 1, i.e. $\sum_{\ell \in \mathbf{W}_N^\theta} \mathbf{P}_N^{t,v}(\ell) = 1$. The numbers t_a, $a = 1, \ldots, k$ are always assumed to be in a small neighborhood of 0, which guarantees, in particular, that the measure is normalizable, i.e. $Z_N(\mathbf{t}, \mathbf{v}) \neq 0$.

Observe that $\mathbf{P}_N^{t,v}$ satisfies the assumptions of Theorem 4.1 with (in the notations of Assumption 3)

$$\phi_N^+(x) = \left(\phi^+ \left(\frac{x}{N} \right) + \frac{1}{N} \varphi_N^+ \left(\frac{x}{N} \right) + O \left(\frac{1}{N^2} \right) \right)$$

$$\times \prod_{a=1}^{m} \left(v_a - \frac{x}{N} + \frac{1}{N} \right) \left(t_a + v_a - \frac{x}{N} \right),$$

$$\phi_N^-(x) = \left(\phi^- \left(\frac{x}{N} \right) + \frac{1}{N} \varphi_N^- \left(\frac{x}{N} \right) + O \left(\frac{1}{N^2} \right) \right)$$

$$\times \prod_{a=1}^{m} \left(v_a - \frac{x}{N} \right) \left(t_a + v_a - \frac{x}{N} + \frac{1}{N} \right).$$

We also define

$$\psi_N^-(z) = \varphi_N^-(z) + \sum_{a=1}^{m} \frac{\phi^-(z)}{t_a + v_a - z}, \qquad \psi_N^+(z) = \varphi_N^+(z) + \sum_{a=1}^{m} \frac{\phi^+(z)}{v_a - z}.$$

Clearly,

$$\phi_N^\pm(Nz) = \left(\phi^\pm(z) + \frac{\psi_N^\pm(z)}{N} + O \left(\frac{1}{N^2} \right) \right) \cdot \prod_{a=1}^{m} (v_a - z)(t_a + v_a - z).$$

We further define $\mu_N^{t,v}$ as the empirical distribution of $\mathbf{P}_N^{t,v}$ and set, cf. (18)

$$(\mathbf{105}) \qquad G_N(z) = \int_{-\infty}^{\infty} \frac{1}{z - x} \mu_N^{t,v}(dx) = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{z - \ell_i/N},$$

$(\ell_1, \ldots, \ell_N)$ is $\mathbf{P}_N^{t,v}$-distributed.

Define

$$(\mathbf{106}) \qquad \Delta G_N(z) = N \big(G_N(z) - G_\mu(z) \big).$$

We will now formulate a generalization of Theorem 2.5. For that we need to introduce certain notations from the theory of hyperelliptic integrals.

Fix k simple positively-oriented complex contours $\gamma_1, \gamma_2, \ldots, \gamma_k$, such that each γ_i $(1 \leq i \leq k)$ encloses the segment $[\hat{a}_i, \hat{b}_i]$ (and thus, also $[\alpha_i, \beta_i]$), but not the other segments.

Let $P(z) = p_0 + p_1 z + \cdots + p_{k-2} z^{k-2}$ be a polynomial of degree $k-2$, and consider the map

$$\Omega : P(z) \mapsto \left(\frac{1}{2\pi \mathbf{i}} \oint_{\gamma_1} \frac{P(z)dz}{\prod_{i=1}^{k} \sqrt{(z-\alpha_i)(z-\beta_i)}}, \ldots, \right.$$
$$\left. \frac{1}{2\pi \mathbf{i}} \oint_{\gamma_k} \frac{P(z)dz}{\prod_{i=1}^{k} \sqrt{(z-\alpha_i)(z-\beta_i)}} \right)$$

Note that the sum of the integrals in the definition of Ω equals $(-2\pi \mathbf{i})$ times the residue of $\frac{P(z)}{\prod_{i=1}^{k} \sqrt{(z-\alpha_i)(z-\beta_i)}}$ at infinity, which is zero. Therefore, Ω is a linear map between $(k-1)$-dimensional vector spaces. This map is very non-trivial, but it is known to be an isomorphism for any $k \geq 2$ (cf. [Du, Section 2.1]).

Using Ω we can now define a more complicated map Υ. Given a (continuous) function $f(z)$ defined on the contours γ_i and such that the sum of its integrals over these contours is zero, we define a function $\Upsilon_z[f]$ through

$$\Upsilon_z[f] = f(z) + \frac{P(z)}{\prod_{i=1}^{k} \sqrt{(z-\alpha_i)(z-\beta_i)}},$$

where $P(z)$ is a unique polynomial of degree at most $k-2$, such that for each $i = 1, \ldots, k$

$$\frac{1}{2\pi \mathbf{i}} \oint_{\gamma_i} \Upsilon_z[f]dz = 0.$$

The polynomial $P(z)$ can be evaluated in terms of the map Ω via

$$P = \Omega^{-1} \left(-\frac{1}{2\pi \mathbf{i}} \oint_{\gamma_1} f(z)dz, \ -\frac{1}{2\pi \mathbf{i}} \oint_{\gamma_2} f(z)dz, \ \ldots, \ -\frac{1}{2\pi \mathbf{i}} \oint_{\gamma_k} f(z)dz \right).$$

Note that the map $f \mapsto \Upsilon_z[f]$ is linear and does not depend on $\mathbf{t}, \mathbf{v}$.

Theorem **6.1.** — *Fix $m = 0, 1, \ldots$ and choose complex numbers $\mathbf{v} = (v_1, \ldots, v_m) \subset (\mathbf{C} \setminus \bigcup_{i=1}^{k} [\hat{a}_i, \hat{b}_i])^m$. Under Assumptions 1–5 we have as $N \to \infty$*

(**107**)
$$\frac{\partial^m}{\partial t_1 \cdots \partial t_m} \mathbf{E}_{\mathbf{P}_N^{\mathbf{t},\mathbf{v}}} \big(\Delta G_N(u) \big) \bigg|_{t_a=0, 1 \leq a \leq m}$$
$$= o(1) + \frac{\partial^m}{\partial t_1 \cdots \partial t_m} \left(\Upsilon_u \left[\frac{\theta^{-1}}{2\pi \mathbf{i} \prod_{i=1}^{k} \sqrt{(u-\alpha_i)(u-\beta_i)}} \right. \right.$$

$$\times \oint_{\bigcup_{i=1}^{k} \gamma_i} \frac{dz}{(u-z)\mathrm{H}(z)} \bigg(\psi_{\mathrm{N}}^{-}(z)e^{-\theta \mathrm{G}_\mu(z)} + \psi_{\mathrm{N}}^{+}(z)e^{\theta \mathrm{G}_\mu(z)}$$

$$+ \phi^{-}(z)e^{-\theta \mathrm{G}_\mu(z)} \frac{\theta^2}{2} \frac{\partial}{\partial z} \mathrm{G}_\mu(z)$$

$$+ \phi^{+}(z)e^{\theta \mathrm{G}_\mu(z)} \bigg(\frac{\theta^2}{2} - \theta \bigg) \frac{\partial}{\partial z} \mathrm{G}_\mu(z) \bigg) \bigg] \bigg)_{l_a=0,\, 1 \le a \le m},$$

where γ_i are simple positively-oriented contours enclosing the segment $[\hat{a}_i, \hat{b}_i]$ (the points u and $v_1, \ldots, v_m$ are outside the contours). The remainder $o(1)$ is uniform over $u, v_1, \ldots, v_m$ in compact subsets of the unbounded component of $\mathbf{C} \setminus \bigcup_{i=1}^{k} \gamma_i$.

Remark **6.2.** — Theorem 6.1 generalizes Theorem 2.5 in several directions. First, before we had $k = 1$ and now we allow any $k = 1, 2, \ldots$. Second, θ was equal to 1 and now $\theta > 0$ is arbitrary. Finally, in Theorem 2.5 the functions $\mathrm{H}(z)$, $\phi^{\pm}(z)$, $\mathrm{G}_\mu(z)$ had a specific form.

Proof. — Similarly to Theorem 2.5, the proof has two parts: algebraic manipulations and self-improving estimates for the remainders. The latter literally repeats Section 2.5 and we are not going to present it.

For the former we start from the statement of Theorem 4.1. Making the change of variables $\xi = \mathrm{N}z$, we can write for z outside $\bigcup_{i=1}^{k}[\hat{a}_i, \hat{b}_i]$

$$(108) \qquad \prod_{i=1}^{\mathrm{N}} \bigg(1 - \frac{\theta}{\xi - \ell_i} \bigg) = \exp\bigg(\sum_{i=1}^{\mathrm{N}} \ln\bigg(1 - \frac{1}{\mathrm{N}} \cdot \frac{\theta}{z - \ell_i/\mathrm{N}} \bigg) \bigg)$$

$$= \exp\bigg(-\theta \mathrm{G}_{\mathrm{N}}(z) + \frac{\theta^2}{2\mathrm{N}} \cdot \frac{\partial}{\partial z} \mathrm{G}_{\mathrm{N}}(z) + \mathrm{O}\bigg(\frac{1}{\mathrm{N}^2} \bigg) \bigg)$$

$$= \exp\bigg(-\theta \mathrm{G}_\mu(z) - \frac{\theta}{\mathrm{N}} \Delta \mathrm{G}_{\mathrm{N}}(z) + \frac{\theta^2}{2\mathrm{N}} \cdot \frac{\partial}{\partial z} \mathrm{G}_{\mathrm{N}}(z)$$

$$+ \mathrm{O}\bigg(\frac{1}{\mathrm{N}^2} \bigg) \bigg).$$

Similarly, also

$$(109) \qquad \prod_{i=1}^{\mathrm{N}} \bigg(1 + \frac{\theta}{\xi - \ell_i - 1} \bigg) = \exp\bigg(\sum_{i=1}^{\mathrm{N}} \ln\bigg(1 + \frac{1}{\mathrm{N}} \cdot \frac{\theta}{z - \ell_i/\mathrm{N} - 1/\mathrm{N}} \bigg) \bigg)$$

$$= \exp\bigg(\theta \mathrm{G}_\mu(z) + \frac{\theta}{\mathrm{N}} \Delta \mathrm{G}_{\mathrm{N}}(z) + \frac{\theta^2/2 - \theta}{\mathrm{N}} \cdot \frac{\partial}{\partial z} \mathrm{G}_{\mathrm{N}}(z) + \mathrm{O}\bigg(\frac{1}{\mathrm{N}^2} \bigg) \bigg).$$

Recalling the definition of $R_\mu(z)$, we conclude that the function $R_N(Nz)$ from Theorem 4.1 for $\mathbf{P}_N^{\mathbf{t},\mathbf{v}}$ can be written in the following form

$$(\mathbf{110}) \qquad R_N(Nz) = \prod_{a=1}^{m} (v_a - z)(t_a + v_a - z)$$

$$\times \left[R_\mu(z) + \phi^-(z) e^{\theta G_\mu(z)} \mathbf{E}_{\mathbf{P}_N^{\mathbf{t},\mathbf{v}}} \left(\exp\left(\frac{\theta}{N} \Delta G_N(z) \right) - 1 \right) \right.$$

$$+ \phi^+(z) e^{-\theta G_\mu(z)} \mathbf{E}_{\mathbf{P}_N^{\mathbf{t},\mathbf{v}}} \left(\exp\left(-\frac{\theta}{N} \Delta G_N(z) \right) - 1 \right)$$

$$+ \frac{\psi_N^-(z)}{N} e^{-\theta G_\mu(z)} + \frac{\psi_N^+(z)}{N} e^{\theta G_\mu(z)} + \frac{\phi^-(z) e^{-\theta G_\mu(z)}}{N} \frac{\theta^2}{2} \frac{\partial}{\partial z} G_\mu(z)$$

$$\left. + \frac{\phi^+(z) e^{\theta G_\mu(z)}}{N} \left(\frac{\theta^2}{2} - \theta \right) \frac{\partial}{\partial z} G_\mu(z) \right] + o\left(\frac{1}{N} \right).$$

We further want to simplify the expression in the second line of (110), by replacing $e^h - 1$ by h under expectations. As in Theorem 2.5, for that we use the inequality

$$\left| e^h - h - 1 \right| \leq |h|^2 e^{|h|}, \quad h \in \mathbf{C},$$

and the fact that

$$\mathbf{E}_{\mathbf{P}_N^{\mathbf{t},\mathbf{v}}} \left(\left| \frac{\theta}{N} \Delta G_N(z) \right|^2 \exp\left(\left| \frac{\theta}{N} \Delta G_N(z) \right| \right) \right) = o\left(\frac{1}{N} \right), \quad N \to \infty,$$

which is established via self-improving estimates.

We therefore can rewrite (110) as

$$(\mathbf{111}) \qquad R_N(Nz) = \prod_{a=1}^{m} (v_a - z)(t_a + v_a - z) \cdot \left[R_\mu(z) - \frac{\theta Q_\mu(z)}{N} \mathbf{E}_{\mathbf{P}_N^{\mathbf{t},\mathbf{v}}} \left(\Delta G_N(z) \right) \right.$$

$$+ \frac{\psi_N^-(z)}{N} e^{-\theta G_\mu(z)} + \frac{\psi_N^+(z)}{N} e^{\theta G_\mu(z)} + \frac{\phi^-(z) e^{-\theta G_\mu(z)}}{N} \frac{\theta^2}{2} \frac{\partial}{\partial z} G_\mu(z)$$

$$\left. + \frac{\phi^+(z) e^{\theta G_\mu(z)}}{N} \left(\frac{\theta^2}{2} - \theta \right) \frac{\partial}{\partial z} G_\mu(z) \right] + o\left(\frac{1}{N} \right),$$

where the remainder is uniform over z in compact subsets of $\mathbf{C} \setminus \bigcup_{i=1}^{k} [\hat{a}_i, \hat{b}_i]$.

Let us now fix u outside the contours $\gamma_1, \ldots, \gamma_k$, divide (111) by $2\pi \mathbf{i} \prod_{a=1}^{m} (v_a - z)$ $(t_a + v_a - z) \cdot H(z) \cdot (u - z)$ (the definition of $H(z)$ is given in Assumption 4) and integrate over the union of contours γ_i. We get

$$(112) \qquad \frac{\theta}{2\pi \mathbf{i}} \oint_{\bigcup_{i=1}^{k} \gamma_i} \frac{\prod_{i=1}^{k} \sqrt{(z-\alpha_i)(z-\beta_i)}}{u-z} \cdot \mathbf{E}_{\mathbf{P}_N^{\mathbf{t},\mathbf{v}}}\big(\Delta G_N(z)\big) dz$$

$$= \frac{1}{2\pi \mathbf{i}} \oint_{\bigcup_{i=1}^{k} \gamma_i} \frac{dz}{(u-z)H(z)} \cdot \left(\psi_N^-(z) e^{-\theta G_\mu(z)} + \psi_N^+(z) e^{\theta G_\mu(z)} \right.$$

$$+ \phi^-(z) e^{-\theta G_\mu(z)} \frac{\theta^2}{2} \frac{\partial}{\partial z} G_\mu(z) + \phi^+(z) e^{\theta G_\mu(z)} \left(\frac{\theta^2}{2} - \theta \right) \frac{\partial}{\partial z} G_\mu(z) \right)$$

$$+ \frac{1}{2\pi \mathbf{i}} \oint_{\bigcup_{i=1}^{k} \gamma_i} \left(\frac{-N R_N(Nz)}{(u-z)H(z)} + \frac{N R_\mu(z)}{(u-z)H(z)} + o(1) \right) dz.$$

Since $H(z)$ is non-zero on all the segments $[\hat{a}_i, \hat{b}_i]$, by a proper choice of small contours γ_i we can guarantee that $H(z)$ is non-zero inside the integration contours. Then the terms in the last line of (112) vanish as they have no singularities inside the integration contours.

Turning to the first line of (112), note that $\mathbf{E}_{\mathbf{P}_N^{\mathbf{t},\mathbf{v}}}(\Delta G_N(z))$ is analytic outside the contours of integration and decays as $1/z^2$ when $z \to \infty$. Therefore, we can compute the integral as (minus) the sum of the residues at $z = u$ and at $z = \infty$. The former is

$$\theta \prod_{i=1}^{k} \sqrt{(u-\alpha_i)(u-\beta_i)} \cdot \mathbf{E}_{\mathbf{P}_N^{\mathbf{t},\mathbf{v}}}\big(\Delta G_N(u)\big),$$

while the latter is a polynomial $P_N(u)$ of degree at most $k-2$ (to see that one uses $(z-u)^{-1} = z^{-1} \sum_{\mathfrak{m} \geq 0} (uz)^{\mathfrak{m}}$). We conclude that

$$(113) \qquad \theta \cdot \mathbf{E}_{\mathbf{P}_N^{\mathbf{t},\mathbf{v}}}\big(\Delta G_N(u)\big) + \frac{P_N(u)}{\prod_{i=1}^{k} \sqrt{(u-\alpha_i)(u-\beta_i)}}$$

$$= \frac{1}{2\pi \mathbf{i} \cdot \prod_{i=1}^{k} \sqrt{(u-\alpha_i)(u-\beta_i)}} \oint_{\bigcup_{i=1}^{k} \gamma_i} \frac{dz}{(u-z)H(z)}$$

$$\times \left(\psi_N^-(z) e^{-\theta G_\mu(z)} + \psi_N^+(z) e^{\theta G_\mu(z)} + \phi^-(z) e^{-\theta G_\mu(z)} \frac{\theta^2}{2} \frac{\partial}{\partial z} G_\mu(z) \right.$$

$$\left. + \phi^+(z) e^{\theta G_\mu(z)} \left(\frac{\theta^2}{2} - \theta \right) \frac{\partial}{\partial z} G_\mu(z) \right) + o(1).$$

Now we are in a position to apply the map Υ_u. Indeed, the integral of $G_N(z)$ around γ_i is deterministic and equals $n_i(N)/N$. On the other hand, the integral of $G_\mu(z)$ around γ_i equals the total mass of $\mu(x)$ inside γ_i, which is $\hat{n}_i$. Since, $\hat{n}_i = n_i(N)/N$, the integral of $\mathbf{E}_{\mathbf{P}_N^{\mathbf{t},\mathbf{v}}}(\Delta G_N(u))$ around each loop γ_i vanishes.

Therefore,

(**114**)
$$\mathbf{E}_{\mathbf{P}_N^{t,v}}\big(\Delta G_N(u)\big)$$

$$= \Upsilon_u\Bigg[\frac{\theta^{-1}}{2\pi\mathbf{i}\cdot\prod_{i=1}^k\sqrt{(u-\alpha_i)(u-\beta_i)}}\oint_{\bigcup_{i=1}^k\gamma_i}\frac{dz}{(u-z)H(z)}$$

$$\times\left(\psi_N^-(z)e^{-\theta G_\mu(z)}+\psi_N^+(z)e^{\theta G_\mu(z)}+\phi^-(z)e^{-\theta G_\mu(z)}\left(\frac{\theta^2}{2}\frac{\partial}{\partial z}G_N(z)(z)\right)\right.$$

$$\left.+\phi^+(z)e^{\theta G_\mu(z)}\left((\theta^2/2-\theta)\frac{\partial}{\partial z}G_\mu(z)(z)\right)\right)\Bigg]+o(1).$$

It remains to show that (114) can be differentiated with respect to $t_1,\dots,t_m$ and $t_1=\cdots=t_m=0$, which is done in the same way as in Theorem 2.5, see Section 2.5 for the details. It is important to note here that due to its explicit definition, the map Υ_u is continuous and even Lipshitz (in uniform norm on the contours γ_i, $i=1,\dots,k$). Moreover, it is linear and does not depend on $\mathbf{t}$. Therefore, its appearance in the formulas does not affect the argument. $\qquad\square$

7. Central Limit Theorem

The following two theorems are corollaries of Theorem 6.1, cf. Section 2.2.

Theorem **7.1.** — *Under Assumptions 1–5 the joint moments of the random variables* $N(G_N(z)-\mathbf{E}_{\mathbf{P}_N}G_N(z))$ *approximate (uniformly in z in compact subsets of $\mathbf{C}\setminus\bigcup_{i=1}^k[\hat{a}_i,\hat{b}_i]$) those of centered Gaussian random variables with covariance*

$$N^2\big(\mathbf{E}_{\mathbf{P}_N}\big(G_N(z)G_N(w)\big)-\mathbf{E}_{\mathbf{P}_N}G_N(z)\mathbf{E}_{\mathbf{P}_N}G_N(w)\big)=\theta^{-1}\mathfrak{C}(z,w)+o(1),$$
$$N\to\infty,$$

where

(**115**)
$$\mathfrak{C}(z,w)=\Upsilon_w\Bigg[\frac{1}{(w-z)^2}-\frac{\sqrt{\prod_i(z-\alpha_i)(z-\beta_i)}}{\sqrt{\prod_i(w-\alpha_i)(w-\beta_i)}}\left(\frac{1}{(z-w)^2}\right.$$

$$\left.-\frac{1}{2(z-w)}\sum_{i=1}^k\left(\frac{1}{z-\alpha_i}+\frac{1}{z-\beta_i}\right)\right)\Bigg].$$

Remark **7.2.** — The map Υ was defined in Section 6. In cases $k=1$ and $k=2$ the resulting covariance function $\mathfrak{C}(z,w)$ can be brought to a more explicit form. When $k=1$,

$$\mathfrak{C}(z,w)=\frac{1}{(z-w)^2}\left(1-\frac{zw-\frac{1}{2}(\alpha_1+\beta_1)(z+w)+\alpha_1\beta_1}{\sqrt{(z-\alpha_1)(z-\beta_1)}\sqrt{(w-\alpha_1)(w-\beta_1)}}\right).$$

The expression for $k = 2$ involves the values of complete elliptic integrals and we do not pursue it here, cf. [BDE, Section 3].

Remark **7.3.** — Note that the numbers α_i, β_i may depend on $\mu(x)$, which, in turn, may depend on N through filling fractions $\hat{n}_i = n_i(N)/N$. However, this dependence is continuous (see Proposition 5.8) and thus, if $\hat{n}_i$ converges as $N \to \infty$, then so is $\mathfrak{C}(z, w)$, and Theorem 7.1 turns into a conventional central limit theorem for $NG_N(z)$ as $N \to \infty$. In particular, if $k = 1$, then this is always the case.

Remark **7.4.** — In the general beta random matrix models the covariance has precisely the same form, cf. [J3], [Shch], [BoGu2].

Corollary **7.5.** — *Take $m \geq 1$ real functions $f_1(z), \ldots, f_m(z)$ on $\bigcup_{i=1}^{k} [\hat{a}_i, \hat{b}_i]$ that can be extended to holomorphic functions in a complex neighborhood $\mathcal{B}$ of $\bigcup_{i=1}^{k} [\hat{a}_i, \hat{b}_i]$. Under Assumptions 1–5, as $N \to \infty$ the joint moments of the m random variables*

$$\mathcal{L}_{f_j} = \sum_{i=1}^{N} \big(f_j(\ell_i) - \mathbf{E}_{\mathbf{P}_N} f_j(\ell_i)\big), \quad (\ell_1, \ldots, \ell_N) \text{ is } \mathbf{P}_N\text{-distributed,}$$

approximate those of centered Gaussian random variables with asymptotic covariance

$$(\mathbf{116}) \qquad \mathbf{E}_{\mathbf{P}_N} \mathcal{L}_{f_i} \mathcal{L}_{f_j} = \frac{1}{(2\pi \mathbf{i})^2} \oint_{\bigcup_i \gamma_i} \oint_{\bigcup_i \gamma_i} f_i(u) f_j(v) \mathfrak{C}(u, v) du dv + o(1),$$

where γ_i, $i = 1, \ldots, k$, are positively oriented contours in $\mathcal{B}$ that enclose $[\hat{a}_i, \hat{b}_i]$, respectively, and $\mathfrak{C}(u, v)$ is given by (115).

Remark **7.6.** — Similarly to Corollary 7.5, Theorem 6.1 can be used to obtain the first two terms in the asymptotic expansion of $\sum_{i=1}^{N} \mathbf{E}_{\mathbf{P}_N} f(\ell_i)$ for functions f holomorphic in a neighborhood of $\bigcup_{i=1}^{k} [\hat{a}_i, \hat{b}_i]$.

8. Non-vanishing weights

Assumption 5 of the general setup of Section 3 was vanishing of the weight at the end-points of the supporting intervals. For future applications it is convenient to relax this condition and replace it by the following exponential bound on probabilities of having particles at $a_i(N) + 1$ or $b_i(N) - 1$.

Assumption **6.** — *We require the existence of constants C_1, C_2, $C_3 > 0$ such that for all $N = 1, 2, \ldots$, the $\mathbf{P}_N$-probability of the event*

$$\ell_j = a_i(N) + 1 \text{ or } \ell_j = b_i(N) - 1 \quad \text{for at least one pair } 1 \leq j \leq N, \ 1 \leq i \leq k$$

is bounded from above by $C_1 \exp(-C_2 N^{C_3})$.

Theorem **8.1.** — *If we replace Assumption 5 by Assumption 6, then the results of Theorem 6.1, Theorem 7.1, Corollary 7.5 are still valid for measures* $\mathbf{P}_N$.

The proof of Theorem 8.1 is based on the following observation.

Proposition **8.2.** — *In the notations of Theorem 4.1, suppose that the condition* $\phi_N^-(a_i(N) + 1) = \phi_N^+(b_i(N)) = 0$, $i = 1, \ldots, k$, *does not hold. If* $\phi_N^\pm(\xi)$ *are holomorphic in a domain* $\mathcal{M}_N \subset \mathbf{C}$, *then so is* $R_N(\xi)$ *except for at most* $2k$ *simple poles. These poles are at points* $\{a_i(N) + 1, b_i(N) - 1\}_{i=1}^k$. *Under Assumption 6 the corresponding residues decay exponentially in* N *as* $N \to \infty$ *in the same sense as in the bound of Assumption 6 (perhaps, with different constants* C_1, C_2, C_3).

Proof. — We repeat the proof of Theorem 4.1 and observe the same cancellation of the poles. The only poles for which the cancellations do not occur, are endpoints of the interval: in Theorem 4.1 the functions $\phi^\pm$ were vanishing at these endpoints, but this is no longer the case. The residue at an end-point can be bounded for the first term in (78) by the probability to have a particle at such an end-point (denote it by m) multiplied by

$$\max_{\ell_i} \left| \prod_{i:\ell_i \neq m} \left(1 - \frac{\theta}{m - \ell_i} \right) \right| \leq \left| \prod_{i=1}^{N-1} \left(1 + \frac{\theta}{\theta i} \right) \right| = N.$$

Thus, the exponential decay of the probability in Assumption 6 implies the exponential decay of the residue. For the second term in (78) the argument is the same. $\square$

Proof of Theorem 8.1. — Note that an analogue of Proposition 8.2 is readily established also for the deformed measures $\mathbf{P}_N^{\mathbf{t},\mathbf{v}}$ as in Section 6. Indeed, we need such measures only for small (i.e. tending to 0) values of t_i, but then the exponential bounds on probability remain valid. Therefore, all the arguments of Section 6 go through for the measures $\mathbf{P}_N$. Indeed, the only difference between Theorem 4.1 and Proposition 8.2 is in the appearance of finitely many simple poles. However, since the residues are exponentially small, these poles will only add exponentially small terms to all the remainders and thus will not contribute to the expansions in powers of $1/N$ that we study. $\square$

9. Examples

The aim of this section is to demonstrate how the Assumptions 1–4 are checked in applications, which yields the validity of Theorem 6.1, Theorem 7.1, and Corollary 7.5 for certain stochastic systems.

9.1. *Multi-cut general θ extension of the Krawtchouk ensemble.* — The first example is an extension of that of Section 2 to general values of θ and k.

We fix $k = 1, 2, \ldots$ and take $3k$ numbers $\hat{a}_i$, $\hat{b}_i$, $\hat{n}_i$, such that

$$\hat{a}_1 < \hat{b}_1 < \hat{a}_2 < \hat{b}_2 < \cdots < \hat{a}_k < \hat{b}_k,$$

$0 < \hat{n}_i < \theta^{-1}(\hat{b}_i - \hat{a}_i)$ for all $i = 1, \ldots, k$, and $\sum_{i=1}^{k} \hat{n}_i = 1$. Then we further choose numbers $a_i(N)$, $b_i(N)$, $n_i(N)$ such that the model fits into the setup of Section 3. The weight $w(x; N)$ is then defined through the identity

$$(\mathbf{117}) \qquad \frac{w(x; N)}{w(x-1; N)} = -\prod_{i=1}^{k} \frac{x - b_i(N)}{x + 1 - a_i(N)}.$$

Note that (117) agrees with the conditions $w(x; N) > 0$, $x \in \bigcup_{i=1}^{k}[a_i(N) + 1, b_i(N) - 1]$ and with Assumption 5. Since the weight $w(x; N)$ is supported on $\bigcup_{i=1}^{k}(a_i(N), b_i(N))$, we need to supplement (117) by the choice of k constants

$$(\mathbf{118}) \qquad \mathfrak{c}_i(N) = w\big(a_i(N) + 1; N\big), \quad i = 1, \ldots, k.$$

Observe that the multiplication of all $\mathfrak{c}_i$ by a same constant leaves the probability distribution unchanged. Therefore, for each N the system depends on the choice of $4k - 2$ constants. In particular, if $k = 1$ and $\theta = 1$, then (up to a shift of the lattice) we arrive at the example of Section 2.

If we now assume that all the parameters are chosen so that as $N \to \infty$

$$a_i(N) = N\hat{a}_i + O(1), \quad b_i(N) = N\hat{b}_i + O(1), \quad n_i(N) = N\hat{n}_i + O(1),$$

$$\mathfrak{c}_i(N) = \exp(N\hat{\mathfrak{c}}_i), \quad i = 1, \ldots, k,$$

then using Stirling's formula for the factorials appearing in the explicit expressions for $w(x; N)$, it is easy to see that the model satisfies Assumptions 1, 2, 3. Further,

$$\phi^+(z) = -\prod_{i=1}^{k}(z - \hat{b}_i), \quad \phi^-(z) = \prod_{i=1}^{k}(z - \hat{a}_i).$$

The function $R_\mu(z)$ is an analytic function in z, which is $O(z^{k-1})$ as $z \to \infty$. Therefore, by Liouville's theorem $R_\mu(z)$ is a polynomial of degree at most $k - 1$. Hence, the quadratic equation (103) implies that $Q_\mu(z)$ is the square root of a degree $2k$ polynomial. In other words, decomposing into linear factors we get

$$(\mathbf{119}) \qquad Q_\mu(z) = 2\sqrt{\prod_{i=1}^{k}(z - \alpha_i)(z - \beta_i)}.$$

On the other hand, our choice of $\hat{n}_i$ guarantees, that the equilibrium measure $\mu(x)dx$ in each interval $[\hat{a}_i, \hat{b}_i]$ has some points where $\mu(x) > 0$ and some points where $\mu(x) < \theta^{-1}$.

Combining this with the observation (that is immediate from (76)) that $Q_\mu(\hat{a}_i)$ and $Q_\mu(\hat{b}_i)$ have different signs for each $i = 1, \ldots, k$, we conclude that $\alpha_i, \beta_i \in [\hat{a}_i, \hat{b}_i]$. For each i there are two options: either $\mu(x)$ has a band inside $[\hat{a}_i, \hat{b}_i]$, and then α_i and β_i are endpoints inside this band, or $[\hat{a}_i, \hat{b}_i]$ is a union of a void and a saturated region. Let us explain that the latter is impossible. Indeed, then there must be a point $x \in \mathbf{R}$ and $\varepsilon > 0$ such that $\mu(x) = 0$ on $(x - \varepsilon, x)$ and $\mu(x) = \theta^{-1}$ on $(x, x + \varepsilon)$ (or vice versa which is considered in the same way). Therefore, as z approaches x along the real axis from the left, the Stieltjes transform $G_\mu(z)$ explodes:

$$\lim_{z \to x-} G_\mu(z) = \lim_{z \to x-} \int \frac{\mu(t)}{z - t} dt = -\infty.$$

This implies $\lim_{z \to x-} \exp(G_\mu(z)) = 0$ and $\lim_{z \to x-} \exp(-G_\mu(z)) = +\infty$. But then the definition (75) of $R_\mu(z)$ implies that $R_\mu(z)$ has a singularity at $z = x$, which contradicts the fact that $R_\mu(z)$ is a polynomial. Note that it is crucial in this argument that $\phi^\pm$ does not have zeros inside $\bigcup(\hat{a}_i, \hat{b}_i)$, and indeed the end-points $\hat{a}_i, \hat{b}_i$ might separate voids and saturated regions.

The conclusion is that this class of probability models satisfies Assumptions 1–5 and Theorem 6.1, Theorem 7.1, Corollary 7.5 are valid for them. As far as we know, these results are new with the exception of the case $\theta = k = 1$.

9.2. *Lozenge tilings.* — The second example demonstrates that at least two instances of probability measures arising in the study of uniformly random lozenge tilings fit into our framework.

Consider an $A \times B \times C$ hexagon drawn on the regular triangular lattice. We tile this hexagon with three types of elementary lozenges (which are unions of adjacent triangular faces of the lattice), cf. Figure 3. There are finitely many such tilings, and we are interested in the asymptotic behavior of uniformly random tiling as $A, B, C \to \infty$. This is a well-studied model, with many results available, cf. [CLP], [BKMM], [JN], [G], [BoGo], [P1], [P2].

Let us dissect the hexagon by a vertical line at distance t from the left boundary. There will be a fixed (depending on A, B, C, t) number N of horizontal lozenges on this line; let $\mathbf{P}_N$ denote the distribution of these lozenges. This distribution can be computed by noticing that the tiling can be viewed as two *Gelfand-Tsetlin patterns* glued together, as shown in the right panel of Figure 3. The enumeration of Gelfand-Tsetlin patterns is well-known and can be used to compute the distribution $\mathbf{P}_N$ (see [CLP], [G], [BP], [BuGo] for more details). Assuming $t > \max(B, C)$ as in Figure 3, which yields $N = B + C - t$, and introducing the coordinate system such that the lowest possible position for horizontal lozenges on the tth vertical line is 1 and the highest one is $A + B + C - t$, we obtain the
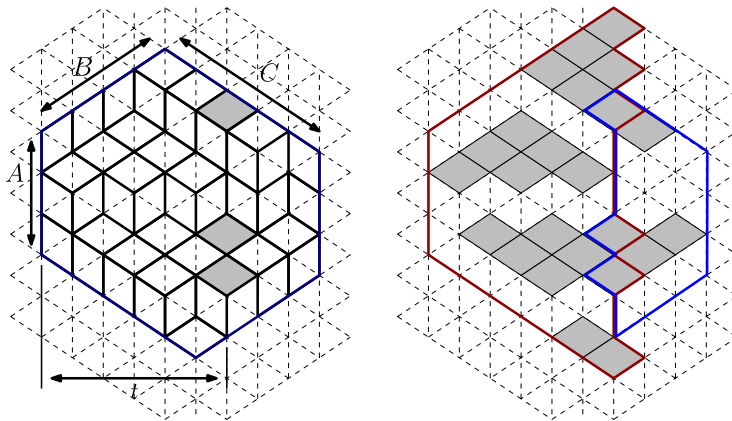
Fɪɢ. 3. — *Left panel*: Lozenge tiling of the $3 \times 4 \times 5$ hexagon and 3 horizontal lozenges on the sixth from the *left vertical line*. *Right panel*: two Gelfand-Tsetlin patterns (*inside blue and red contours*) corresponding to each tiling (Color figure online)

formula

$$(\mathbf{120}) \qquad \mathbf{P}_{\mathrm{N}}(\ell_1, \ldots, \ell_{\mathrm{N}}) = \frac{1}{\mathrm{Z}_{\mathrm{N}}} \prod_{i<j} (\ell_i - \ell_j)^2$$

$$\times \prod_{i=1}^{\mathrm{N}} \big[(\mathrm{A} + \mathrm{B} + \mathrm{C} + 1 - t - \ell_i)_{t-\mathrm{B}} (\ell_i)_{t-\mathrm{C}} \big],$$

where $(a)_n$ is the Pochhammer symbol, $(a)_n = a(a+1) \cdots (a+n-1)$, and Z_{N} is a normalizing constant (which can be computed explicitly in this case). The distribution of the form (120) is known as *Hahn orthogonal polynomial ensemble.*

Take a large parameter L and suppose that

$$(\mathbf{121}) \qquad \begin{aligned} \mathrm{A} &= \hat{\mathrm{A}}\mathrm{L} + \mathrm{O}(1), & \mathrm{B} &= \hat{\mathrm{B}}\mathrm{L} + \mathrm{O}(1), \\ \mathrm{C} &= \hat{\mathrm{C}}\mathrm{L} + \mathrm{O}(1), & t &= \hat{t}\mathrm{L} + \mathrm{O}(1), \end{aligned}$$

for positive constants $\hat{\mathrm{A}}$, $\hat{\mathrm{B}}$, $\hat{\mathrm{C}}$, $\hat{t}$. We assume that $\hat{t} > \max(\hat{\mathrm{B}}, \hat{\mathrm{C}})$—other possibilities for t are considered similarly. Let us check that under such choice of parameters the ensemble (120) satisfies Assumptions 1–5.

- Assumption 1 is satisfied due to Stirling's formula applied to Pochhammer symbols. The potential $\mathrm{V}(u)$ has the form:

$$\begin{aligned} \mathrm{V}(u) = &-(\hat{\mathrm{A}} + \hat{\mathrm{C}} - u) \ln(\hat{\mathrm{A}} + \hat{\mathrm{C}} - u) \\ &+ (\hat{\mathrm{A}} + \hat{\mathrm{B}} + \hat{\mathrm{C}} - \hat{t} - u) \ln(\hat{\mathrm{A}} + \hat{\mathrm{B}} + \hat{\mathrm{C}} - \hat{t} - u) \\ &- (\hat{t} - \hat{\mathrm{C}} + u) \ln(\hat{t} - \hat{\mathrm{C}} + u) + u \ln(u). \end{aligned}$$
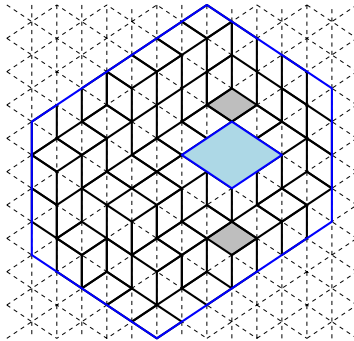
FIG. 4. — Lozenge tiling of the $4 \times 7 \times 5$ hexagon with a rhombic 2×2 *hole* (shown in *blue*). The center of the hole is at distance t from the left side of the hexagon. The remaining *horizontal lozenges* on tth *vertical line* are shown in *gray* (Color figure online)

- Assumption 2 is empty, as there is only one filling fraction, $n_1(N) = N$.
- Assumption 3 is immediate from the definitions, and we have

$$\phi^+(z) = (\hat{t} - \hat{C} + z)(\hat{A} + \hat{B} + \hat{C} - \hat{t} - z), \quad \phi^-(z) = z(\hat{A} + \hat{C} - z).$$

- For Assumption 4 note that $R_\mu(z)$ is an analytic function which grows as $O(z^2)$ as $z \to \infty$ and therefore it is a polynomial of degree at most two. Hence, by (103), the function $Q_\mu(z)$ is a square root of a polynomial. The definition implies that $Q_\mu(z)$ is $O(z)$ as $z \to \infty$. Thus $Q_\mu(z)$ is a square root of degree two polynomial and has the form

$$Q_\mu(z) = \text{const} \cdot \sqrt{(z-a)(z-b)}.$$

  The points a and b are necessarily endpoints of the band of the equilibrium measure $\mu(x)dx$.
- Assumption 5 is immediate from the formula (120).

The conclusion is that Theorem 6.1, Theorem 7.1, Corollary 7.5 are valid. This implies a form of the Central Limit Theorem for fluctuations of lozenge tilings of a hexagon. We remark that the same CLT (and even stronger statement concerning joint asymptotic Gaussianity for several values of $\hat{t}$) can be also established by other methods, cf. [P2], [BD], [BuGo2].

Remark **9.1.** — One can probably use the fact that $R_\mu(z)$ is a degree two polynomial to find an explicit formula for $R_\mu(z)$ and thus, also for $G_\mu(z)$, and for the equilibrium measure $\mu(x)dx$ describing the *limit shape* for lozenge tilings, cf. Section 9.4 for a similar argument. Explicit formulas for $\mu(x)$ were previously found by other methods in [CLP], [BKMM], [KO], [G], [P1].

One can also analyze tilings of more complicated domains. Let us cut a rhombic $D \times D$ hole in the hexagon, as shown in Figure 4. Assume that the bottom point of the

hole is at distance t from the left side of the hexagon and at height H (counted from the bottom of the hexagon along tth vertical line). Let $\mathbf{P}_N$ be the probability distribution of the horizontal lozenges (outside the hole) on tth vertical line induced by the uniform measure on all tilings of the hexagon with the hole. We can repeat the same argument as the one used for the complete hexagon for computing $\mathbf{P}_N$. Assuming $t > \max(B, C)$ as in Figure 3, which yields $N = B + C - D - t$, and introducing the coordinate system such that the lowest possible position for horizontal lozenge on the tth vertical line is 1 and the highest one is $A + B + C - t$, we obtain

$$(\mathbf{122}) \qquad \mathbf{P}_N(\ell_1, \dots, \ell_N)$$

$$= \frac{1}{Z_N} \prod_{1 \le i < j \le N} (\ell_i - \ell_j)^2$$

$$\times \prod_{i=1}^{N} \big[(A + B + C + 1 - t - \ell_i)_{t-B} (\ell_i)_{t-C} (H - \ell_i)_D (H - \ell_i)_D \big].$$

We also need two filling fractions n_1 and n_2: we consider only such tilings that there are n_1 horizontal lozenges (on tth vertical line) below the hole and n_2 lozenges above. Take a large parameter L and suppose that in addition to (121) we have

$$H = \hat{H}L + O(1), \qquad D = \hat{D}L + O(1),$$

$$\hat{H} > 0, \ \hat{D} > 0, \ \hat{H} + \hat{D} < \hat{A} + \hat{B} + \hat{C} - \hat{t},$$

$$n_1 = \hat{n}_1 L + O(1), \qquad n_2 = \hat{n}_2 L + O(1),$$

$$0 < \hat{n}_1 < \hat{H}, \ 0 < n_2 < \hat{A} + \hat{B} + \hat{C} - \hat{t} - \hat{H} - \hat{D}.$$

We again assume that $\hat{t} > \max(\hat{B}, \hat{C})$ and remark that other possibilities for t are considered similarly. Let us check that under such choice of parameters the ensemble (122) satisfies Assumptions 1–5.

- Assumption 1 is satisfied due to Stirling's formula applied to Pochhammer symbols.
- Assumption 2 is satisfied due to restrictions on $\hat{n}_1$, $\hat{n}_2$.
- Assumption 3 is immediate from the definitions, and we have

$$\phi^+(z) = (\hat{t} - \hat{C} + z)(\hat{A} + \hat{B} + \hat{C} - \hat{t} - z)(\hat{H} - z)^2,$$

$$\phi^-(z) = z(\hat{A} + \hat{C} - z)(\hat{H} + \hat{D} - z)^2.$$

- For Assumption 4 note that $R_\mu(z)$ is an analytic function which grows as $O(z^4)$ as $z \to \infty$ and therefore it is a polynomial of degree at most four. Hence, by (103), the function $Q_\mu(z)$ is a square root of a polynomial. The definition

implies that $Q_\mu(z)$ is $O(z^3)$ as $z \to \infty$. Thus, $Q_\mu(z)$ is a square root of a degree six polynomial. Further observe that according to definitions $Q_\mu(0) < 0$, $Q_\mu(\hat{H}) > 0$, $Q_\mu(\hat{H} + \hat{D}) < 0$, $Q_\mu(\hat{A} + \hat{B} + \hat{C} - \hat{\imath}) > 0$. Together with analyticity of $Q_\mu(z)$ outside the support of the equilibrium measure this yields

$$Q_\mu(z) = \mathrm{const}(z - c)\sqrt{(z - a_1)(z - b_1)(z - a_2)(z - b_2)},$$

where

$$0 < a_1 \le b_1 < \hat{H} < c < \hat{H} + \hat{D} < a_2 < b_2 < \hat{A} + \hat{B} + \hat{C} - \hat{\imath}.$$

As before, the points a_1, b_1, a_2, b_2 can be identified with the endpoints of the bands of the equilibrium measure.

- Assumption 5 is immediate from the formula (122).

The conclusion is that Theorem 6.1, Theorem 7.1, Corollary 7.5 are valid. This implies a form of the Central Limit Theorem for fluctuations of lozenge tilings of a hexagon with a hole (with fixed filling fractions above and below the hole). As far as we know, these results are new; the same applies to the examples of the next sections.

9.3. *Arbitrary convex potential on $\mathbf{R}$ with no saturation.* — For our third example take a real *convex* analytic function $V(x)$, i.e. such that $V''(x) > 0$ for all $x \in \mathbf{R}$. Fix a constant $\kappa > 0$ such that

$$(\mathbf{123}) \qquad \liminf_{x \to \infty} \frac{\kappa V(x)}{2\theta \ln |x|} > 1$$

and consider a probability distribution

$$(\mathbf{124}) \qquad \mathbf{P}_N(\ell_1, \ldots, \ell_N) = \frac{1}{Z_N} \prod_{1 \le i < j \le N} \frac{\Gamma(\ell_j - \ell_i + 1)\Gamma(\ell_j - \ell_i + \theta)}{\Gamma(\ell_j - \ell_i)\Gamma(\ell_j - \ell_i + 1 - \theta)}$$

$$\times \prod_{i=1}^{N} \exp\left(-\kappa N \cdot V\left(\frac{\ell_i}{N}\right)\right)$$

on N-tuples $\ell_1 < \ell_2 < \cdots < \ell_N$ such that

$$\ell_i = \lambda_i + \theta i, \quad \lambda_1 \le \lambda_2 \le \cdots \le \lambda_N, \quad \lambda_i \in \mathbf{Z}.$$

In other words, we are in the framework of Section 3 except that now the distribution is supported on an *infinite and unbounded* subset of $\mathbf{R}^N$. In particular, when $V(x) = x^2$, we obtain a discretization of the celebrated general β Gaussian ensemble from random matrix theory.

Proposition **9.2.** *— For any real analytic* $V(x)$ *satisfying* $V''(x) > 0$ *and* (123)*, the functional* $I_{\kappa V}$ *of* (87) *has a unique maximizer (equilibrium measure) on the set of all absolutely continuous probability measures on* **R** *(without any restrictions on the density). The equilibrium measure is compactly supported, has a continuous density* $\mu(x)$ *and has a single band. At the end-points of the band the density of the equilibrium measure behaves like* $c\sqrt{x}$.

Proof. — This is well-known, see [DS], [J1] and references therein.   □

Note that the equilibrium measure $\mu(x)dx$ of Proposition 9.2 is unchanged when we multiply both θ and κ on the same positive constant (as the functional is then multiplied by the same constant) and, therefore, depends only on their ratio κ/θ. In particular, if we fix ratio and then choose small κ (equivalently, small θ), then the density $\mu(x)$ will be smaller than θ^{-1} for all θ. But then the solution to constrained maximization problem with $0 \le \mu(x) \le \theta^{-1}$ as in Theorem 5.3 will be the same as the solution of the unconstrained minimization problem as in Proposition 9.2. The conclusion is that for small κ the equilibrium measure has no saturated regions; this is crucial for our considerations.

We now want to show that Theorem 6.1, Theorem 7.1, and Corollary 7.5 apply in this situation. This is done by *localizing* the probability measure $\mathbf{P}_N$ onto a finite interval with help of the following proposition.

Proposition **9.3.** *— Assume that* $V(x)$ *is convex and satisfies* (123)*. Then with exponentially high probability the measure* (124) *is supported on configurations in a linearly growing interval, i.e. there exist constants* $C, D > 0$ *such that*

$$\mathbf{P}_N\left(-D \le \frac{\ell_1}{N} \le \frac{\ell_N}{N} \le D\right) > 1 - \frac{1}{C} \cdot \exp(-NC), \quad N = 1, 2, \dots.$$

Proof. — This a particular case of Theorem 10.1. We remark that statements of this flavor are well-known as large deviations principles for the largest/smallest particles in both discrete and continuous log-gases, cf. [J3, Theorem 2.2], [Fe, Theorem 4.2], [AGZ, Section 2.6].   □

Proposition 9.3 motivates the definition of measure $\widehat{\mathbf{P}_N}$ as $\mathbf{P}_N$ conditioned on the event that $|\ell_i/N| \le D + 1$ for all $i = 1, \dots, N$. An advantage of the measure $\widehat{\mathbf{P}_N}$ is that ℓ_i/N are $\widehat{\mathbf{P}_N}$-almost surely bounded, and therefore, there exists a finite complex contour which enclose them all and we can apply the developments of previous sections.

Lemma **9.4.** *— Assume* (123) *and that analytic* $V(x)$ *satisfies* $V''(x) > 0$ *for all* $x \in \mathbf{R}$. *Then the results of Theorem 6.1, Theorem 7.1, Corollary 7.5 are valid for measures* $\widehat{\mathbf{P}_N}$. *Here* $k = 1$, *the*

points a_1 and b_1 are two endpoints of the band of the equilibrium measure,

$$\phi^-(z) = 1, \qquad \varphi_N^-(z) = 0, \qquad \phi^+(z) = \exp\bigl(-\kappa V'(z)\bigr),$$

$$\varphi_N^+(z) = \frac{\kappa V''(z)}{2} \exp\bigl(-\kappa V'(z)\bigr).$$

Proof. — We will apply Theorem 8.1, and we need to check all its assumptions. With $w(x; N) = \exp(-\kappa N V(x/N))$, we have

(**125**) $$\frac{w(x; N)}{w(x-1; N)} = \exp\left[\kappa N\left(V\left(\frac{\xi-1}{N}\right) - V\left(\frac{\xi}{N}\right)\right)\right].$$

Therefore, the formulas for ϕ^+ and φ^+ are obtained through large N expansion of the right-hand side in (125).

We only check that Assumption 4 is satisfied, as the rest is automatic. We have

$$Q_\mu(z) = \exp\bigl(-\theta G_\mu(z)\bigr) - \exp\bigl(-\kappa V'(z)\bigr)\exp\bigl(\theta G_\mu(z)\bigr).$$

As before, $Q_\mu(z)$ is a square root of an analytic function due to (103). Further, Q_μ is analytic outside the single band of the equilibrium measure $\mu(x)dx$. Near the end-points of the band, the density $\mu(x)$ behaves like a square root and thus similar behavior for $Q_\mu(z)$. Therefore,

$$Q_\mu(z) = H(z)\sqrt{(z-a)(z-b)},$$

where $a < b$ are endpoints of the band. We will now show that $H(z)$ has no zeros on **R**, which would also imply that $H(z)$ is holomorphic, since $H^2(z)$ is.

We already ruled out $H(a) = 0$ and $H(b) = 0$, and there are 3 more cases to consider:

- $H(x) = 0$, $a < x < b$. Then also $Q_\mu(x + \mathbf{i}0) = Q_\mu(x - \mathbf{i}0) = 0$. Therefore,

$$\exp\bigl(\kappa V'(x)\bigr) = \exp\bigl(2\theta G_\mu(x + \mathbf{i}0)\bigr) = \exp\bigl(2\theta G_\mu(x - \mathbf{i}0)\bigr).$$

  Comparing with (100) ($V'(x)$ should be replaced by $\kappa V'(x)$ to match the notations) we conclude that

$$\exp\bigl(\theta G_\mu(x + \mathbf{i}0) - \theta G_\mu(x - \mathbf{i}0)\bigr) = 1.$$

  But this contradicts $0 < \mu(x) < \theta^{-1}$ and (99).
- For $x > b$, if $H(x) = 0$, then also $Q(x) = 0$ and

(**126**) $$\exp\bigl(\kappa V'(x)\bigr) = \exp\bigl(2\theta G_\mu(x)\bigr).$$

But note that (126) holds at $x = b$ as a corollary of (100). And as $x > b$ grows, the left-hand side of (126) also grows due to $V''(x) > 0$, but the right-hand side decays due to the definition of $G_\mu(x)$. Thus, (126) can not hold at $x > b$.

- For $x < a$ the argument is similar. □

As a corollary we obtain a central limit theorem for the original measure $\mathbf{P}_N$.

Corollary **9.5.** — *Assume* (123), $V''(x) > 0$ *for all* $x \in \mathbf{R}$, *and consider the probability measure* $\mathbf{P}_N$ *given by* (124). *Suppose that κ is so small that the equilibrium measure has one band. Take any m functions $f_1(z), \ldots, f_m(z)$ on $\mathbf{R}$ which are bounded and extend to holomorphic functions in a complex neighborhood of* $(-D, D)$. *Then the m random variables*

$$\mathcal{L}_{f_j} = \sum_{i=1}^{N} \big(f_j(\ell_i) - \mathbf{E}_{\mathbf{P}_N} f_j(\ell_i) \big), \quad (\ell_1, \ldots, \ell_N) \text{ is } \mathbf{P}_N\text{-distributed,}$$

converge (in distribution and in the sense of moments) to centered Gaussian random variables with covariance

$$(\mathbf{127}) \qquad \lim_{N \to \infty} \mathbf{E}_{\mathbf{P}_N} \mathcal{L}_{f_i} \mathcal{L}_{f_j} = \frac{\theta^{-1}}{(2\pi\mathbf{i})^2} \oint_\gamma \oint_\gamma f_i(u) f_j(v) \mathcal{C}(u, v) \, du \, dv,$$

where $\mathcal{C}(u, v)$ is given by (22) *with $a_\pm$ being endpoints of the band of the equilibrium measure $\mu(x)dx$, and γ is a positively oriented contour which encloses* $[a_-, a_+]$.

Proof. — Due to Proposition 9.3 and boundedness of functions f_j, the joint moments of $\sum_{i=1}^{N} f_j(\ell_i)$ with respect to $\mathbf{P}_N$ and with respect to $\widehat{\mathbf{P}_N}$ differ by exponentially small (in N) error. It remains to use Corollary 7.5 for $\widehat{\mathbf{P}_N}$. □

Remark **9.6.** — One way to construct functions f_j satisfying conditions of Corollary 9.5 is to take any analytic functions on $[-D, D]$ and extend them to all of $\mathbf{R}$ by setting equal to 0 outside $[-D, D]$. The condition that f_j must be bounded on $\mathbf{R}$ can be weakened. However, we need a growth condition on these functions, as we want to be able to replace $\mathbf{E}_{\mathbf{P}_N}(\sum_{i=1}^{N} f_j(\ell_i))^r$ by $\mathbf{E}_{\widehat{\mathbf{P}_N}}(\sum_{i=1}^{N} f_j(\ell_i))^r$ with negligible error. The analyticity assumption can also likely be weakened, but we do not address this in the present paper.

Remark **9.7.** — The covariance (127) has the same form as for random matrices and log-gases in the one cut regime. It depends only on the restrictions of functions f_j onto the interval $[a_-, a_+]$ and can be rewritten in several other equivalent forms, cf. [J1, Theorem 4.2], [PS, Chapter 3].

9.4. (z, w)-*Measures.* — Our last example originates in the asymptotic representation theory of unitary groups U(N), cf. [O], [BO1], [O2].

Fix two sequences of non-real parameters $\mathbf{z}(N)$ and $\mathbf{w}(N)$ and define

(**128**) $\qquad \mathbf{P}_N(\ell_1, \ldots, \ell_N)$

$$= \frac{1}{Z_N} \prod_{1 \leq i < j \leq N} \frac{\Gamma(\ell_j - \ell_i + 1)\Gamma(\ell_j - \ell_i + \theta)}{\Gamma(\ell_j - \ell_i)\Gamma(\ell_j - \ell_i + 1 - \theta)}$$

$$\times \prod_{i=1}^{N} \frac{1}{\Gamma(\mathbf{z}(N) - \ell_i)\Gamma(\bar{\mathbf{z}}(N) - \ell_i)\Gamma(\mathbf{w}(N) + \ell_i)\Gamma(\bar{\mathbf{w}}(N) + \ell_i)}$$

on N-tuples $\ell_1 < \ell_2 < \cdots < \ell_N$ such that

$$\ell_i = \lambda_i + \theta i, \quad \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_N, \quad \lambda_i \in \mathbf{Z}.$$

Here $\bar{\mathbf{z}}(N)$ and $\bar{\mathbf{w}}(N)$ are complex conjugates of $\mathbf{z}(N)$ and $\mathbf{w}(N)$, respectively.

If $\theta = 1$, $\mathbf{z}(N) = z + N + 1$, $\mathbf{w}(N) = w$ with $\mathbf{Re}(z + w) > -1/2$, and $\{\lambda_i\}_{i=1}^{N}$ are identified with highest weights of irreducible representations of U(N), then (128) describes the decomposition of the character of the "generalized bi-regular" representation of the infinite-dimensional unitary group U(∞), see [O].

With the notation

$$w(x; N) = \frac{1}{\Gamma(\mathbf{z}(N) - x)\Gamma(\bar{\mathbf{z}}(N) - x)\Gamma(\mathbf{w}(N) + x)\Gamma(\bar{\mathbf{w}}(N) + x)},$$

we have

$$\frac{w(x; N)}{w(x - 1; N)} = \frac{(x - \mathbf{z}(N))(x - \bar{\mathbf{z}}(N))}{(x + \mathbf{w}(N) - 1)(x + \bar{\mathbf{w}}(N) - 1)}$$

$$= 1 - \frac{\mathbf{w}(N) + \bar{\mathbf{w}}(N) + \mathbf{z}(N) + \bar{\mathbf{z}}(N) - 2}{x} + O\left(\frac{1}{x^2}\right).$$

The last formula implies that as $|x| \to \infty$, the weight decays as

(**129**) $\qquad w(x; N) = O\left(|x|^{\mathbf{w}(N) + \bar{\mathbf{w}}(N) + \mathbf{z}(N) + \bar{\mathbf{z}}(N) - 2}\right).$

Therefore, the real part of $\mathbf{w}(N) + \bar{\mathbf{w}}(N) + \mathbf{z}(N) + \bar{\mathbf{z}}(N)$ needs to be large in order to guarantee that the measure $\mathbf{P}_N$ is finite. Let us assume that as $N \to \infty$

(**130**) $\qquad \begin{aligned} &\mathbf{w}(N) = \mathbf{w}_\infty \cdot N + O(1), \qquad \mathbf{z}(N) = \mathbf{z}_\infty \cdot N + O(1), \\ &\mathbf{Re}(\mathbf{w}_\infty + \mathbf{z}_\infty) > 1. \end{aligned}$

Then $w(x; N)$ decays fast enough so that a condition of the form (123) is satisfied. Note that the original representation-theoretic case $\mathbf{z}(N) = z + N + 1$, $\mathbf{w}(N) = w$ does not

satisfy this assumption, and indeed the equilibrium measure in this case is known to be somewhat degenerate, and one generally does not expect to see the Gaussian behavior. This case with $\theta = 1$ was studied in [BO1].

Proposition **9.8.** — *Assume* (130). *Then with exponentially high probability the measure* (128) *is supported on configurations in a linearly growing interval, i.e. there exist constants* $C, D > 0$ *such that*

$$\mathbf{P}_N\left(-D \le \frac{\ell_1}{N} \le \frac{\ell_N}{N} \le D\right) > 1 - \frac{1}{C} \cdot \exp(-NC), \quad N = 1, 2, \ldots.$$

Proof. — This is a particular case of Theorem 10.1.   □

Proposition 9.8 implies that we can use for $\mathbf{P}_N$ the techniques developed in Sections 3–6. Let us present the functions R_μ and Q_μ, as they can be found explicitly in this case. Indeed, we have

$$R_\mu(\xi) = (\xi - \mathbf{z}_\infty)(\xi - \bar{z}_\infty) \exp\big(\theta G_\mu(\xi)\big)$$
$$+ (\xi + \mathbf{w}_\infty)(\xi + \bar{w}_\infty) \exp\big(-\theta G_\mu(\xi)\big).$$

Since $R_\mu(\xi)$ is analytic and grows as $2\xi^2$ as $|\xi| \to \infty$, $R_\mu(\xi)$ is a degree two polynomial, i.e.

$$R_\mu(\xi) = 2\xi^2 + A\xi + B.$$

Let us find the coefficients A and B. Expand G_μ in power series near $\xi = \infty$ as

$$G_\mu(\xi) = \frac{1}{\xi} + \frac{p_1}{\xi^2} + O\big(\xi^{-3}\big),$$

where p_1 is unknown. Plugging this into the definition of $R_\mu(\xi)$ and expanding up to $O(\xi^{-1})$ we get

$$(\mathbf{131}) \qquad R_\mu(\xi) = (\xi - \mathbf{z}_\infty)(\xi - \bar{\mathbf{z}}_\infty)\left(1 + \frac{\theta}{\xi} + \frac{\theta p_1 + \theta^2}{\xi^2}\right)$$

$$+ (\xi + \mathbf{w}_\infty)(\xi + \bar{\mathbf{w}}_\infty)\left(1 - \frac{\theta}{\xi} + \frac{-\theta p_1 + \theta^2}{\xi^2}\right) + O\big(\xi^{-1}\big)$$

$$= 2\xi^2 + (\mathbf{w}_\infty + \bar{\mathbf{w}}_\infty - \mathbf{z}_\infty - \bar{\mathbf{z}}_\infty)\xi + \mathbf{z}_\infty \bar{\mathbf{z}}_\infty + \mathbf{w}_\infty \bar{\mathbf{w}}_\infty$$

$$- \theta(\mathbf{z}_\infty + \bar{\mathbf{z}}_\infty + \mathbf{w}_\infty + \bar{\mathbf{w}}_\infty) + 2\theta^2.$$

Therefore, we can also find Q_μ through (103):

$$(\mathbf{132}) \qquad Q_\mu(\xi) = \sqrt{\big(R_\mu(\xi)\big)^2 - 4(\xi - \mathbf{z}_\infty)(\xi - \bar{\mathbf{z}}_\infty)(\xi + \mathbf{w}_\infty)(\xi + \bar{\mathbf{w}}_\infty)}$$

$$= c\sqrt{(\xi - a_-)(\xi - a_+)}$$

with explicit $a_\pm$ and c which are found by plugging (131) in (103) (the resulting formulas are somewhat complicated and we omit them). This is precisely the form we need for Assumption 4.

Now repeating the argument of Lemma 9.4, we see that the results of Theorem 6.1, Theorem 7.1, Corollary 7.5 are valid for measures $\mathbf{P}_N$ conditioned on the event of Proposition 9.8. Thus, as in Corollary 9.5, we arrive at the following result.

Corollary **9.9.** *— Assume* (130) *and consider the probability measure* $\mathbf{P}_N$ *given by* (128). *Take* $m \geq 1$ *functions* $f_1(z), \ldots, f_m(z)$ *on* $\mathbf{R}$ *that are bounded and extend to holomorphic functions in a complex neighborhood of* $(-D, D)$, *where* D *is given by Proposition* 9.8. *Then m random variables*

$$\mathcal{L}_{f_j} = \sum_{i=1}^{N} \big(f_j(\ell_i) - \mathbf{E}_{\mathbf{P}_N} f_j(\ell_i)\big), \quad (\ell_1, \ldots, \ell_N) \text{ is } \mathbf{P}_N\text{-distributed,}$$

converge (in distribution and in the sense of moments) to centered Gaussian random variables with covariance

$$(\mathbf{133}) \qquad \lim_{N \to \infty} \mathbf{E}_{\mathbf{P}_N} \mathcal{L}_{f_i} \mathcal{L}_{f_j} = \frac{\theta^{-1}}{(2\pi\mathbf{i})^2} \oint_\gamma \oint_\gamma f_i(u) f_j(v) \mathcal{C}(u, v) \, du \, dv,$$

where $\mathcal{C}(u, v)$ *is given by* (22) *with* $a_\pm$ *found from* (131), (132), *and* γ *is a positively oriented contour which encloses* $[a_-, a_+]$.

10. Exponential bound on the support

Take a continuous function $V : \mathbf{R} \to \mathbf{R}$ and numbers $\varepsilon > 0$, $H > 1 + \theta^{-1}$ such that

$$(\mathbf{134}) \qquad \frac{V(x)}{2\theta \ln |x|} > 1 + \varepsilon, \quad \text{when } |x| > H.$$

Set $T = \lfloor \theta \rfloor + 1$ and assume that $V(x)$ is increasing for $x > H$, decreasing for $x < -H$ and is Lipshitz with a constant $\mathfrak{s}$ for $|x| < H + T$.

Consider a probability distribution

$$(\mathbf{135}) \qquad \mathbf{P}_N(\ell_1, \ldots, \ell_N) = \frac{1}{Z_N} \prod_{1 \leq i < j \leq N} \frac{\Gamma(\ell_j - \ell_i + 1)\Gamma(\ell_j - \ell_i + \theta)}{\Gamma(\ell_j - \ell_i)\Gamma(\ell_j - \ell_i + 1 - \theta)}$$

$$\times \prod_{i=1}^{N} \exp\left(-N \cdot V\left(\frac{\ell_i}{N}\right)\right)$$

on N-tuples $\ell_1 < \ell_2 < \cdots < \ell_N$ such that

$$(\mathbf{136}) \qquad \ell_i = \lambda_i + \theta i, \qquad \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_N, \qquad \lambda_i \in \mathbf{Z}.$$

The aim of this section is to prove the following statement describing the tails of $\mathbf{P}_N$.

Theorem **10.1.** — *There exist two constants* $C = C(\theta, \varepsilon, H, \mathfrak{s})$ *and* $D = D(\theta, \varepsilon, H, \mathfrak{s})$ *that depend only on* $\theta, \varepsilon, H, \mathfrak{s}$, *and such that*

$$\mathbf{P}_{N,N}\left(-D \leq \frac{\ell_1}{N} \leq \frac{\ell_N}{N} \leq D\right) > 1 - \frac{1}{C} \cdot \exp(-NC), \quad N = 1, 2, \ldots.$$

The proof of Theorem 10.1 borrows ideas from similar proofs in [J3], [AGZ, Section 2.7], [Fe], but additional care is required because of the shifts by θ in the definition of ℓ_i. We present the proof as a series of lemmas.

It is useful to consider several modifications of the measure $\mathbf{P}_N$, which we now introduce. The probability measure $\mathbf{P}_{N,+}$ is defined on the same space of ℓs by the formula

$$(\mathbf{137}) \qquad \mathbf{P}_{N,+}(\ell_1, \ldots, \ell_N) = \frac{1}{Z_{N,+}} \prod_{1 \leq i < j \leq N} \frac{\Gamma(\ell_j - \ell_i + 1)\Gamma(\ell_j - \ell_i + \theta)}{\Gamma(\ell_j - \ell_i)\Gamma(\ell_j - \ell_i + 1 - \theta)}$$

$$\times \prod_{i=1}^{N} \exp\left(-N \cdot V\left(\frac{\ell_i}{N+1}\right)\right).$$

The probability measure $\mathbf{P}_{N,++}$ is defined on the same space of ℓs by the formula

$$(\mathbf{138}) \qquad \mathbf{P}_{N,++}(\ell_1, \ldots, \ell_N)$$

$$= \frac{1}{Z_{N,++}} \prod_{1 \leq i < j \leq N} \frac{\Gamma(\ell_j - \ell_i + 1)\Gamma(\ell_j - \ell_i + \theta)}{\Gamma(\ell_j - \ell_i)\Gamma(\ell_j - \ell_i + 1 - \theta)}$$

$$\times \prod_{i=1}^{N} \exp\left(-(N+1) \cdot V\left(\frac{\ell_i}{N+1}\right)\right).$$

We also define shifted measures $\mathbf{P}_N^{(k)}$, $\mathbf{P}_{N,+}^{(k)}$, $\mathbf{P}_{N,++}^{(k)}$, $k = 1, \ldots, N+1$, which are given by the same formulas (135), (137), (138) as $\mathbf{P}_N^{(k)}$, $\mathbf{P}_{N,+}^{(k)}$, $\mathbf{P}_{N,++}^{(k)}$, respectively, but with ℓ_i confined to a different lattice, namely

$$(\mathbf{139}) \qquad \ell_i = \begin{cases} \lambda_i + \theta i, & i < k, \\ \lambda_i + \theta(i+1), & i \geq k, \end{cases} \quad \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_N, \ \lambda_i \in \mathbf{Z}.$$

We also let $Z_N^{(k)}$, $Z_{N,+}^{(k)}$, $Z_{N,++}^{(k)}$ to be the normalizing constant for the corresponding measures. Further denote

$$\mathbf{M}_N(\ell_1, \ldots, \ell_N) = Z_N \mathbf{P}_N(\ell_1, \ldots, \ell_N),$$

and similarly for $\mathbf{P}_{N,+}$, $\mathbf{P}_{N,++}$, $\mathbf{P}_N^{(k)}$, $\mathbf{P}_{N,+}^{(k)}$, $\mathbf{P}_{N,++}^{(k)}$.

All the constants $c_1, c_2, \ldots$ in the following statements depend only on $\theta, \varepsilon, H, \mathfrak{s}$, the exact values of the constants might change from statement to statement.

Lemma **10.2.** — *Then there exists $c_1 > 0$ such that for any $1 \leq k \leq N + 1$, and any ℓ satisfying* (136) *there exists ℓ' satisfying* (139), *for which*

$$(\mathbf{140}) \qquad \left| \ell_i - \ell_i' \right| \leq T, \quad \text{for all } 1 \leq i \leq N$$

and

$$(\mathbf{141}) \qquad \frac{\mathbf{M}_N^{(k)}(\ell')}{\mathbf{M}_N(\ell)} \leq c_1 \exp(N c_1).$$

There also exists a (possibly different) ℓ'' such that

$$(\mathbf{142}) \qquad \left| \ell_i - \ell_i'' \right| \leq T, \quad \text{for all } 1 \leq i \leq N$$

and

$$(\mathbf{143}) \qquad \frac{\mathbf{M}_N^{(k)}(\ell'')}{\mathbf{M}_N(\ell)} \geq \frac{1}{c_1} \exp(-N c_1).$$

Similarly, for any ℓ' satisfying (139), *there exists ℓ satisfying* (136) *such that* (140) *and* (141) *hold. Further, for any ℓ'' satisfying* (139), *there exists ℓ satisfying* (136) *such that* (142) *and* (143) *hold. Finally, the same statements hold for the measures $\mathbf{P}_{N,+}^{(k)}$ and $\mathbf{P}_{N,++}^{(k)}$.*

Proof. — We will only prove the first two statements of the Lemma, as the rest can be proven similarly.

Let $\mathbf{x} = (x_1, \ldots, x_N) \in \mathbf{R}^N$ and $\mathbf{y} = (y_1, \ldots, y_N) \in \mathbf{R}^N$ be such that $x_j - x_i \geq \theta(j - i)$, for $1 \leq i < j \leq N$ and there exists $m = 1, \ldots, N$ and $M \in \mathbf{R}$ such that

$$y_i = \begin{cases} x_i, & 1 \leq i \leq m, \\ x_i + M, & m < i \leq N. \end{cases}$$

We claim that there exists $c_2 = c_2(M)$ such that

$$(\mathbf{144}) \qquad \frac{1}{c_2} \exp(-N c_2)$$

$$\leq \prod_{1 \leq i < j \leq N} \left(\frac{\Gamma(x_j - x_i + 1)\Gamma(x_j - x_i + \theta)}{\Gamma(x_j - x_i)\Gamma(x_j - x_i + 1 - \theta)} \cdot \frac{\Gamma(y_j - y_i)\Gamma(y_j - y_i + 1 - \theta)}{\Gamma(y_j - y_i + 1)\Gamma(y_j - y_i + \theta)} \right)$$

$$\leq c_2 \exp(N c_2).$$

Indeed, (91) implies that

$$\prod_{i < j} \left(\frac{\Gamma(x_j - x_i + 1)\Gamma(x_j - x_i + \theta)}{\Gamma(x_j - x_i)\Gamma(x_j - x_i + 1 - \theta)} \cdot \frac{\Gamma(y_j - y_i)\Gamma(y_j - y_i + 1 - \theta)}{\Gamma(y_j - y_i + 1)\Gamma(y_j - y_i + \theta)} \right)$$

$$= \prod_{i \leq m < j} \left(1 + O\left(\frac{1}{x_j - x_i} \right) \right)$$

Since $x_j - x_i \geq \theta(j - i)$, the last product is bounded below by $\frac{1}{c_2} \exp(-Nc_2)$ and above by $c_2 \exp(-Nc_2)$ for some $c_2 > 0$.

Take ℓ satisfying (136). Suppose that $\ell_i < -H$ for $i = 1, \ldots, m_1$, $|\ell_i| \leq H$ for $i = m_1 + 1, \ldots, m_2$, and $\ell_i > H$ for $i = m_2 + 1, \ldots, N$.

Define $\tilde{\ell} = (\tilde{\ell}_1, \ldots, \tilde{\ell}_N)$ through

$$\tilde{\ell}_i = \begin{cases} \ell_i, & i < k, \\ \ell_i + \theta, & i \geq k. \end{cases}$$

And further define $\ell' = (\ell'_1, \ldots, \ell'_N)$ through

$$\ell'_i = \begin{cases} \tilde{\ell}_i - T, & i \leq m_1, \\ \tilde{\ell}_i, & i > m_1 \end{cases}$$

We claim that (141) holds. Indeed, the ratio of the factors in double product $\prod_{i<j}$ is bounded by two applications of (144), and it remains to bound

$$\prod_{i=1}^{N} \exp\left(NV\left(\frac{\ell_i}{N}\right) - NV\left(\frac{\ell'_i}{N}\right) \right).$$

If $i \leq m_1$, then $\ell'_i \leq \ell_i$ and the monotonicity of $V(x)$ implies that corresponding factors are less than 1. If $i > m_2$, then $\ell'_i \geq \ell_i$ and again the monotonicity of $V(x)$ implies that corresponding factors are less than 1. Finally, if $m_1 < i \leq m_2$, then the Lipshitz property of $V(x)$ gives the desired bound.

Next, we construct ℓ''. For $i \leq m_1$ we set $\ell''_i := \tilde{\ell}_i$. For $i > m_2$ we set $\ell''_i = \tilde{\ell}_i - T$. Since $H > 1 + \theta^{-1}$, we can always choose the remaining coordinates ℓ''_i, $m_1 < i \leq m_2$ in such a way that ℓ'' satisfies (139) and (140). We claim that (143) holds. Indeed, the ratio of the factors in double product $\prod_{i<j}$ is bounded by (144), and it remains to bound

$$\prod_{i=1}^{N} \exp\left(NV\left(\frac{\ell_i}{N}\right) - NV\left(\frac{\ell'_i}{N}\right) \right).$$

If $i \leq m_1$, then $\ell'_i \geq \ell_i$ and the monotonicity of $V(x)$ implies that corresponding factors are greater than 1. If $i > m_2$, then $\ell'_i \leq \ell_i$ and again the monotonicity of $V(x)$ implies that corresponding factors are greater than 1. Finally, if $m_1 < i \leq m_2$, then the Lipshitz property of $V(x)$ gives the desired bound.   □

Lemma **10.3.** — *There exists $c_2 > 0$ such that*

$$(\textbf{145}) \qquad \frac{Z_{N,+}^{(k)}}{Z_{N,+}} \leq c_2 \exp(c_2 N), \quad 1 \leq k \leq N + 1.$$

Proof. — For ℓ' satisfying (139) let $s(\ell)$ denote the corresponding ℓ of Lemma 10.2. Then

$$Z_{N,+}^{(k)} = \sum_{\ell'} \mathbf{M}_{N,+}^{(k)}(\ell') \le c_1 \exp(Nc_1) \sum_{\ell'} \mathbf{M}_{N,+}\big(s(\ell')\big)$$

$$\le c_1 \exp(Nc_1)(2T+1)^N \sum_{\ell} \mathbf{M}_{N,+}(\ell) = c_1 \exp(Nc_1)(2T+1)^N Z_{N,+}.$$
$\square$

Lemma **10.4.** *— There exists $c_3 > 0$ such that*

$$\frac{Z_{N-1,++}}{Z_N} \le c_3 \exp(Nc_3) \cdot N^{-2\theta N}, \quad N = 1, 2, \ldots.$$

Proof. — We have

(**146**)
$$N\frac{Z_N}{Z_{N-1,++}} = N \sum_{\ell} \frac{\mathbf{M}_N(\ell)}{Z_{N-1,++}}$$

$$= \sum_{k=1}^{N} \sum_{\ell^{(k)}} \frac{\mathbf{M}_{N-1,++}^{(k)}(\ell^{(k)})}{Z_{N-1,++}}$$

$$\times \sum_{m=\ell_{k-1}^{(k)}+\theta}^{\ell_k^{(k)}-\theta} \prod_{i=1}^{N-1} \frac{\Gamma(|m-\ell_i^{(k)}|+1)\Gamma(|m-\ell_i^{(k)}|+\theta)}{\Gamma(|m-\ell_i^{(k)}|)\Gamma(|m-\ell_i^{(k)}|+1-\theta)}$$

$$\times \exp\left(-NV\left(\frac{m}{N}\right)\right),$$

where $\ell^{(k)}$ varies over (139) with N replaced by $N-1$, and we use the notation $\ell_0^{(1)} = -\infty$, $\ell_N^{(N)} = +\infty$. Note that by the definitions, the sum $\sum_{m=\ell_{k-1}^{(k)}+\theta}^{\ell_k^{(k)}-\theta}$ is always non-empty. Take ℓ satisfying (136) with N replaced by $N-1$ and let $s^{(k)}(\ell)$ denote the corresponding ℓ' of Lemma 10.2. Then (146) implies

(**147**)
$$N\frac{Z_N}{Z_{N-1,++}} \ge (2T+1)^{-N} \sum_{k=1}^{N} \sum_{\ell} \frac{\mathbf{M}_{N-1,++}^{(k)}(s^{(k)}(\ell))}{Z_{N-1,++}}$$

$$\times \sum_{m=s^{(k)}(\ell)_{k-1}+\theta}^{s^{(k)}(\ell)_k-\theta} \prod_{i=1}^{N-1} \frac{\Gamma(|m-s^{(k)}(\ell)_i|+1)\Gamma(|m-s^{(k)}(\ell)_i|+\theta)}{\Gamma(|m-s^{(k)}(\ell)_i|)\Gamma(|m-s^{(k)}(\ell)_i|+1-\theta)}$$

$$\times \exp\left(-NV\left(\frac{m}{N}\right)\right),$$

where ℓ varies over (136).

Using Stirling's formula, its corollary (91), and definition of the state space (139), we see that the product of Gamma functions in the above formula is bounded from below by

$$\frac{\exp(-\mathrm{N}c_4)}{c_4} \cdot \mathrm{N}^{2\theta\mathrm{N}}$$

with a constant $c_4 > 0$. Let v be the maximum of $\mathrm{V}(x)$ over $[-2\mathrm{T}, 2\mathrm{T}]$. Then (147) and Lemma 10.2 imply that for $c_5 > 0$

$$\mathrm{N}\frac{Z_{\mathrm{N}}}{Z_{\mathrm{N}-1,++}} \geq \mathrm{N}^{2\theta\mathrm{N}}\frac{\exp(-c_5\mathrm{N})}{c_5}\sum_{k=1}^{\mathrm{N}}\sum_{\ell}\mathbf{P}_{\mathrm{N}-1,++}(\ell)$$

$$\times \sum_{m=s^{(k)}(\ell)_{k-1}+\theta}^{s^{(k)}(\ell)_k-\theta} \mathrm{I}_{-2\mathrm{T}\leq m\leq 2\mathrm{T}}\exp(-\mathrm{N}v)$$

Note that for every ℓ satisfying (136) there exists at least one k, such that for at least one $m \in s^{(k)}(\ell)_{k-1} + \theta, s^{(k)}(\ell)_{k-1} + \theta + 1, \ldots, s^{(k)}(\ell)_k - \theta$, we have $|m| \leq 2\mathrm{T}$ (here again we use the notation $s^{(1)}(\ell)_0 = -\infty$, $s^{(\mathrm{N})}(\ell)_{\mathrm{N}} = +\infty$). Therefore,

$$\mathrm{N}\frac{Z_{\mathrm{N}}}{Z_{\mathrm{N}-1,++}} \geq \mathrm{N}^{2\theta\mathrm{N}}\frac{\exp(-c_5\mathrm{N})}{c_5}\sum_{\ell}\mathbf{P}_{\mathrm{N}-1,++}(\ell)\exp(-\mathrm{N}v)$$

$$= \mathrm{N}^{2\theta\mathrm{N}}\frac{\exp(-c_5\mathrm{N})}{c_5}\exp(-\mathrm{N}v). \qquad \square$$

Lemma **10.5.** — *There exists $c_4 > 0$ such that*

(**148**) $$\frac{Z_{\mathrm{N},+}}{Z_{\mathrm{N},++}} \leq c_4 \exp(c_4\mathrm{N}), \quad \mathrm{N} = 1, 2, \ldots.$$

Proof. — Define the random probability measure ν_{N} through

$$\nu_{\mathrm{N}} = \frac{1}{\mathrm{N}}\sum_{i=1}^{\mathrm{N}}\delta_{\ell_i/(\mathrm{N}+1)}, \quad (\ell_1, \ldots, \ell_{\mathrm{N}}) \text{ is } \mathbf{P}_{\mathrm{N},++}\text{-distributed}$$

Then

(**149**) $$\frac{Z_{\mathrm{N},+}}{Z_{\mathrm{N},++}} = \sum_{\ell}\frac{\mathbf{M}_{\mathrm{N},+}(\ell)}{Z_{\mathrm{N},++}} = \mathbf{E}_{\mathbf{P}_{\mathrm{N},++}}\left[\exp\left(\mathrm{N}\int \mathrm{V}(x)\nu_{\mathrm{N}}(dx)\right)\right].$$

In order to bound (149) we start with a lower bound

$$Z_{\mathrm{N},++} \geq \mathbf{M}_{\mathrm{N},++}(\theta, 2\theta, \ldots, \mathrm{N}\theta) \geq \exp\left(-c_5\mathrm{N}^2\right) \cdot \mathrm{N}^{2\theta\mathrm{N}^2}$$

for some $c_5 > 0$. On the other hand, using (91) we obtain

$$\ln \mathbf{M}_{N,++}(\ell_1, \ldots, \ell_N)$$

$$\leq \exp\left(2\theta N^2 \ln N - N^2 \iint_{x \neq y} \left(\frac{V(x) + V(y)}{2} - \theta \ln|x - y| \right) \right.$$

$$\left. \times \nu_N(dx)\nu_N(dy) + c_6 N^2 \right),$$

for some $c_6 > 0$. Assumption (134) and inequality $\ln|x - y| \leq \ln(|x| + 1) + \ln(|y| + 1)$ imply that that there exist $c_7 > 0$ such that for all $x \neq y$,

$$V(x) + V(y) - 2\theta \ln|x - y|$$

$$\geq \frac{\varepsilon}{1 + \varepsilon}\left(V(x) + V(y)\right) + \left(\frac{V(x)}{1 + \varepsilon} - 2\theta \ln(|x| + 1) \right)$$

$$+ \left(\frac{V(y)}{1 + \varepsilon} - 2\theta \ln(|y| + 1) \right)$$

$$\geq \frac{\varepsilon}{1 + \varepsilon}\left(V(x) + V(y)\right) - c_7.$$

Therefore, for each $L > 0$ we have

$$\mathbf{E}_{\mathbf{P}_{N,++}}\left[\int V(x)\mu_N(dx) > L \right]$$

$$\leq \exp\left(c_5 N^2 + c_6 N^2 + \frac{c_7}{2}N^2 - \frac{\varepsilon}{2(1 + \varepsilon)}N^2 L \right)$$

$$\times \prod_{i=1}^{N}\left(\sum_{m \in \mathbf{Z} + i\theta} \exp\left(-\frac{\varepsilon}{2(1 + \varepsilon)}NV(m/N) \right) \right).$$

Hence, there exists $c_8 > 0$, and L_0 such that for all $L > L_0$,

$$\mathbf{E}_{\mathbf{P}_{N,++}}\left[\int V(x)\mu_N(dx) > L \right] \leq \exp\left(-c_8 L N^2 \right).$$

Together with (149) this implies (148). □

Proof of Theorem 10.1. — We have

$$(\mathbf{150}) \qquad \mathbf{P}_N(\ell_N > DN) = \frac{Z_{N-1,+}}{Z_N} \sum_{\tilde{\ell}} \mathbf{P}_{N-1,+}(\tilde{\ell}) \sum_{m > \max(\tilde{\ell}_{N-1}, DN)} \exp\left(-NV\left(\frac{m}{N} \right) \right)$$

$$\times \prod_{i=1}^{N-1}\left[\frac{\Gamma(m - \tilde{\ell}_i + 1)\Gamma(m - \tilde{\ell}_i + \theta)}{\Gamma(m - \tilde{\ell}_i)\Gamma(m - \tilde{\ell}_i + 1 - \theta)} \exp\left(-V\left(\frac{\tilde{\ell}_i}{N} \right) \right) \right],$$

where $\tilde{\ell}$ ranges over (136) with N replaced by N $-$ 1. The combination of Lemmas 10.3 and 10.5 implies that for $c_5 > 0$ we have a bound

$$\frac{Z_{N-1,+}}{Z_N} \le c_5 \exp(Nc_5) \cdot N^{-2\theta N}.$$

For the product of gamma-functions in (150) we use Stirling's formula and its corollary (91), which yields

$$\prod_{i=1}^{N-1}\left[\frac{\Gamma(m-\tilde{\ell}_i+1)\Gamma(m-\tilde{\ell}_i+\theta)}{\Gamma(m-\tilde{\ell}_i)\Gamma(m-\tilde{\ell}_i+1-\theta)}\right]$$

$$\le c_6 \exp(c_6 N)N^{2\theta N}\prod_{i=1}^{N-1}\left|\frac{m}{N}-\frac{\tilde{\ell}_i}{N}\right|^{2\theta}$$

$$\le c_6 \exp(c_6 N)N^{2\theta N}2^{(2\theta+1)(N-1)}\prod_{i=1}^{N-1}\left[\left|\frac{m}{N}\right|^{2\theta}+\left|\frac{\tilde{\ell}_i}{N}\right|^{2\theta}\right].$$

It follows that for $c_7 > 0$,

(**151**) $\quad \mathbf{P}_N(\ell_N > DN) \le c_7 \exp(c_7 N)\sum_{\tilde{\ell}}\mathbf{P}_{N-1,+}(\tilde{\ell})\sum_{m>\max(\tilde{\ell}_{N-1},DN)}\exp\left(-NV\left(\frac{m}{N}\right)\right)$

$$\times \prod_{i=1}^{N-1}\left[\left(|\frac{m}{N}|^{2\theta}+|\frac{\tilde{\ell}_i}{N}|^{2\theta}\right)\exp\left(-V\left(\frac{\tilde{\ell}_i}{N}\right)\right)\right].$$

Further, take $\delta > (2\theta)^{-1}$, and observe that when $D > H + 1$,

$$\sum_{m>\max(\tilde{\ell}_{N-1},DN)}\exp\left(-\delta V\left(\frac{m}{N}\right)\right) \le \sum_{m=\lfloor DN\rfloor}^{\infty}\exp\left(-2\delta\theta\ln(m/N)\right)$$

$$= N^{2\delta\theta}\sum_{m=\lfloor DN\rfloor}^{\infty}\frac{1}{m^{2\delta\theta}}$$

$$\le 2ND^{1-2\delta\theta}.$$

Also due to (134), when $m > DN > (H+1)N$, we have

$$\left(\left|\frac{m}{N}\right|^{2\theta}+\left|\frac{\tilde{\ell}_i}{N}\right|^{2\theta}\right)\exp\left(-V\left(\frac{\tilde{\ell}_i}{N}\right)-\frac{1+\varepsilon/2}{1+\varepsilon}V\left(\frac{m}{N}\right)\right)$$

$$\le \text{const}\cdot\left|\frac{m}{N}\right|^{2\theta}\exp\left(-\frac{1+\varepsilon/2}{1+\varepsilon}V\left(\frac{m}{N}\right)\right) \le \text{const}\left|\frac{N}{m}\right|^{\theta\varepsilon} \le \text{const}\cdot D^{-\theta\varepsilon}$$

Therefore, writing $N(V(\frac{m}{N}))$ in (151) as

$$N\left(V\left(\frac{m}{N}\right)\right) = \frac{1+\varepsilon/2}{1+\varepsilon}(N-1)V\left(\frac{m}{N}\right) + \delta V\left(\frac{m}{N}\right)$$
$$+ \left(N - \frac{1+\varepsilon/2}{1+\varepsilon}(N-1) - \delta\right)V\left(\frac{m}{N}\right),$$

and noticing that the last term is positive when $m > DN$ and N is large, we conclude that for some $c_8 > 0$, all $D > H + 1$ and all $N > N_0$

$$\mathbf{P}_N(\ell_N > DN) \le c_8 \exp(c_8 N) \cdot N^2 \cdot D^{-\theta\varepsilon N}.$$

Choosing D large enough, we obtain the desired exponential estimate for $\mathbf{P}_N(\ell_N > DN)$. The estimate for $\mathbf{P}_N(\ell_1 < -DN)$ is obtained in the same way—the only difference is that we now need to bound $\frac{Z_{N-1,+}^{(1)}}{Z_N}$ instead of $\frac{Z_{N-1,+}}{Z_N}$, but for that we use Lemma 10.4. $\qquad\square$

Remark **10.6.** — It is very plausible that one can similarly establish an analogue of Theorem 10.1 for more general models in the framework of Section 3 with $a_1(N) = -\infty$ and $b_k(N) = +\infty$. The only necessary modification in the above proofs is in Lemma 10.2, where we should take into account that the Lipshitz property of $V(x)$ might fail near the endpoints $a_i(N)$, $b_i(N)$, cf. (72). We will not address here the exact conditions on $V(x)$ under which an analogue of Lemma 10.2 holds for the models in framework of Section 3.

Acknowledgements

REFERENCES

[AM]   J. AMBJØRN and Yu. MAKEENKO, Properties of loop equations for the Hermitian matrix model and for two-dimensional gravity, *Mod. Phys. Lett. A*, **5** (1990), 1753–1763.

[AGZ]   G. ANDERSON, A. GUIONNET and O. ZEITOUNI, *Introduction to Random Matrices*, Cambridge Studies in Advanced Mathematics, 2009.

[BBDT] J. BAIK, A. BORODIN, P. DEIFT and T. SUIDAN, A model for the bus system in Cuernevaca (Mexico), *J. Phys. A, Math. Gen.*, **39** (2006), 8965, arXiv:math/0510414.

[BKMM] J. Baik, T. Kriecherbauer, K. T.-R. McLaughlin and P. D. Miller, Uniform asymptotics for polynomials orthogonal with respect to a general class of discrete weights and universality results for associated ensembles, arXiv:math/0310278.

[BFG] F. Bekerman, A. Figalli and A. Guionnet, Transport maps for Beta-matrix models and universality, *Commun. Math. Phys.*, **338** (2015), 589–619, arXiv:1311.2315.

[BeGu] G. Ben Arous and A. Guionnet, Large deviations for Wigner's law and Voiculescu's non-commutative entropy, *Probab. Theory Relat. Fields*, **108** (1997), 517–542.

[BDE] G. Bonnet, F. David and B. Eynard, Breakdown of universality in multi-cut matrix models, *J. Phys. A, Math. Gen.*, **33** (2000), 6739, arXiv:cond-mat/0003324.

[B1] A. Borodin, Schur dynamics of the Schur processes, *Adv. Math.*, **228** (2011), 2268–2291, arXiv:1001.3442.

[B2] A. Borodin, CLT for spectra of submatrices of Wigner random matrices, *Mosc. Math. J.*, **14** (2014), 29–38, arXiv:1010.0898.

[BF] A. Borodin and P. Ferrari, Anisotropic growth of random surfaces in $2 + 1$ dimensions, *Commun. Math. Phys.*, **325** (2014), 603–684, arXiv:0804.3035.

[BoGo] A. Borodin and V. Gorin, Shuffling algorithm for boxed plane partitions, *Adv. Math.*, **220** (2009), 1739–1770, arXiv:0804.3071.

[BoGo2] A. Borodin and V. Gorin, General beta Jacobi corners process and the Gaussian Free Field, *Commun. Pure Appl. Math.*, **68** (2015), 1683–1884, arXiv:1305.3627.

[BO1] A. Borodin and G. Olshanski, Harmonic analysis on the infinite-dimensional unitary group and determinantal point processes, *Ann. Math.*, **161** (2005), 1319–1422, arXiv:math/0109194.

[BO2] A. Borodin and G. Olshanski, Asymptotics of Plancherel-type random partitions, *J. Algebra*, **313** (2007), 40–60, arXiv:math/0610240.

[BP] A. Borodin and L. Petrov, Integrable probability: from representation theory to Macdonald processes, *Probab. Surv.*, **11** (2014), 1–58, arXiv:1310.8007.

[BoGu1] G. Borot and A. Guionnet, Asymptotic expansion of beta matrix models in the one-cut regime, *Commun. Math. Phys.*, **317** (2013), 447–483, arXiv:1107.1167.

[BoGu2] G. Borot and A. Guionnet, Asymptotic expansion of beta matrix models in the multi-cut regime, *Commun. Math. Phys.*, **317** (2013), 447–483, arXiv:1303.1045.

[BEY] P. Bourgade, L. Erdos and H.-T. Yau, Edge universality of beta ensembles, *Commun. Math. Phys.*, **332** (2014), 261–353, arXiv:1306.5728.

[BPS] A. Boutet de Monvel, L. Pastur and M. Shcherbina, On the statistical mechanics approach in the random matrix theory: integrated density of states, *J. Stat. Phys.*, **79** (1995), 585–611.

[BD] J. Breuer and M. Duits, Central limit theorems for biorthogonal ensembles and asymptotics of recurrence coefficients, J. Amer. Math. Soc. arXiv:1309.6224.

[BIPZ] E. Brézin, C. Itzykson, G. Parisi and J. B. Zuber, Planar diagrams, *Commun. Math. Phys.*, **59** (1978), 35–51.

[BuGo] A. Bufetov and V. Gorin, Representations of classical Lie groups and quantized free convolution, *Geom. Funct. Anal.*, **25** (2015), 763–814, arXiv:1311.5780.

[BuGo2] A. Bufetov and V. Gorin, Fluctuations of particle systems determined by Schur generating functions, arXiv:1604.01110.

[Cha] S. Chatterjee, Rigorous solution of strongly coupled SO(N) lattice gauge theory in the large N limit, arXiv:1502.07719.

[CE] L. O. Chekhov and B. Eynard, Matrix eigenvalue model: Feynman graph technique for all genera, *J. High Energy Phys.*, **0612**, 026 (2006), arXiv:math-ph/0604014.

[CJY] S. Chhita, K. Johansson and B. Young, Asymptotic domino statistics in the Aztec diamond, *Ann. Appl. Probab.*, **25** (2015), 1232–1278, arXiv:1212.5414.

[CLP] H. Cohn Larsen and J. Propp, The shape of a typical boxed plane partition, *N.Y. J. Math.*, **4** (1998), 137–165, arXiv:math/9801059.

[CGM] B. Collins, A. Guionnet and E. Maurel-Segala, Asymptotics of unitary and orthogonal matrix integrals, *Adv. Math.*, **222** (2009), 172–215, arXiv:math/0608193.

[DF] M. Dolega and V. Feray, Gaussian fluctuations of Young diagrams and structure constants of Jack characters, *Duke Math. J.*, **7** (2016), 1193–1282, arXiv:1402.4615.

[DS] P. D. Dragnev and E. B. Saff, Constrained energy problems with applications to orthogonal polynomials of a discrete variable, *J. Anal. Math.*, **72** (1997), 223–259.

[Du] B. A. Dubrovin, Theta functions and non-linear equations, *Russ. Math. Surv.*, **36** (1981), 11–92.

[Ey1]   B. EYNARD, All genus correlation functions for the hermitian 1-matrix model, *J. High Energy Phys.*, **0411**, 031 (2004), arXiv:hep-th/0407261.

[Ey2]   B. EYNARD, All order asymptotic expansion of large partitions, *J. Stat. Mech. Theory Exp.*, **2008**, P07023 (2008).

[Ey3]   B. EYNARD, A matrix model for plane partitions, *J. Stat. Mech. Theory Exp.*, **0910**, P10011, (2009).

[EO]   B. EYNARD and N. ORANTIN, Topological recursion in enumerative geometry and random matrices, *J. Phys. A*, **42** (2009), 293001.

[Fe]   D. FERAL, On large deviations for the spectral measure of discrete Coulomb gas, in *Seminaire de Probabilites XLI Lecture Notes in Mathematics*, vol. 1934, pp. 19–49, 2008.

[Fo]   P. J. FORRESTER, *Log-Gases and Random Matrices*, Princeton University Press, Princeton, 2010.

[G]   V. GORIN, Non-intersecting paths and Hahn orthogonal ensemble, *Funct. Anal. Appl.*, **42** (2008), 180–197, arXiv:0708.2349.

[GS]   V. GORIN and M. SHKOLNIKOV, Multilevel Dyson Brownian motions via Jack polynomials, *Probab. Theory Relat. Fields*, **163** (2015), 413, arXiv:1401.5595.

[GN]   A. GUIONNET and J. NOVAK, Asymptotics of unitary multimatrix models: the Schwinger–Dyson lattice and topological recursion, *J. Funct. Anal.*, **268** (2015), 2851–2905.

[HO]   A. HORA and N. OBATA, *Quantum Probability and Spectral Analysis of Graphs*, Theoretical and Mathematical Physics, Springer, Berlin, 2007.

[J1]   K. JOHANSSON, On fluctuations of eigenvalues of random Hermitian matrices, *Duke Math. J.*, **91** (1998), 151–204.

[J2]   K. JOHANSSON, Discrete orthogonal polynomial ensembles and the Plancherel measure, *Ann. Math. (2)*, **153** (2001), 259–296, math.CO/9906120.

[J3]   K. JOHANSSON, Shape fluctuations and random matrices, *Commun. Math. Phys.*, **209** (2000), 437–476, arXiv:math/9903134.

[J4]   K. JOHANSSON, Non-intersecting paths, random tilings and random matrices, *Probab. Theory Relat. Fields*, **123** (2002), 225–280, arXiv:math/0011250.

[JN]   K. JOHANSSON and E. NORDENSTAM, Eigenvalues of GUE minors, *Electron. J. Probab.*, **11** (2006), 50, arXiv:math/0606760.

[K]   R. KENYON, Height fluctuations in the honeycomb dimer model, *Commun. Math. Phys.*, **281** (2008), 675–709, arXiv:math-ph/0405052.

[KO]   R. KENYON and A. OKOUNKOV, Limit shapes and the complex burgers equation, *Acta Math.*, **199** (2007), 263–302, arXiv:math-ph/0507007.

[KOR]   W. KONIG, N. O'CONNEL and S. ROCH, Non-colliding random walks, tandem queues, and discrete orthogonal polynomial ensembles, *Electron. J. Probab.*, **7** (2002), 1–24.

[KS]   T. KRIECHERBAUER and M. SHCHERBINA, Fluctuations of eigenvalues of matrix models and their applications, arXiv:1003.6121.

[Ma]   I. G. MACDONALD, *Symmetric Functions and Hall Polynomials*, 2nd ed., Oxford University Press, London, 1999.

[1]   M. MAIDA and E. MAUREL-SEGALA, Free transport-entropy inequalities for non-convex potentials and application to concentration for random matrices, *Probab. Theory Relat. Fields*, **159** (2014), 329–356, arXiv:1204.3208.

[Me]   M. L. MEHTA, *Random Matrices*, 3rd ed., Elsevier/Academic Press, Amsterdam, 2004.

[Mo]   A. MOLL, in preparation.

[Mi]   A. A. MIGDAL, Loop equations and 1/N expansion, *Phys. Rep.*, **102** (1983), 199–290.

[NS]   N. NEKRASOV, V. PESTUN and S. SHATASHVILI, Quantum geometry and quiver gauge theories, *High Energy Physics - Theory*, (2013), 1–83, arXiv:1312.6689.

[NP]   N. NEKRASOV and V. PESTUN, Seiberg-Witten geometry of four dimensional N = 2 quiver gauge theories, arXiv:1211.2240.

[N]   N. NEKRASOV, Non-perturbative Dyson–Schwinger equations and BPS/CFT correspondence, in preparation.

[O]   G. OLSHANSKI, The problem of harmonic analysis on the infinite-dimensional unitary group, *J. Funct. Anal.*, **205** (2003), 464–524, arXiv:math/0109193.

[O2]   G. OLSHANKSI, Probability measures on dual objects to compact symmetric spaces and hypergeometric identities, *Funct. Anal. Appl.*, **37** (2001), 281–301.

[Ox]   G. Akemann, J. Baik and P. Di Francesco (eds.), *The Oxford Handbook of Random Matrix Theory*, Oxford University Press, London, 2011.

[PS]   L. PASTUR and M. SHCHERBINA, *Eigenvalue Distribution of Large Random Matrices*, AMS, Providence, 2011.

[P1]   L. PETROV, Asymptotics of random lozenge tilings via Gelfand-Tsetlin schemes, *Probab. Theory Relat. Fields*, **160** (2014), 429–487, arXiv:1202.3901.

[P2]   L. Petrov, Asymptotics of uniformly random lozenge tilings of polygons. Gaussian free field, *Ann. Probab.*, **43** (2015), 1–43, arXiv:1206.5123.

[ST]   E. B. Saff and V. Totik, *Logarithmic Potentials with External Fields*, Springer, Berlin, 1997.

[Shch]   M. Shcherbina, Fluctuations of linear eigenvalue statistics of β matrix models in the multi-cut regime, *J. Stat. Phys.*, **151** (2013), 1004–1034, arXiv:1205.7062.

[Wi]   E. P. Wigner, On the distribution of the roots of certain symmetric matrices, *Ann. Math.*, **67** (1958), 325–327.

A. B.
Department of Mathematics,
Massachusetts Institute of Technology,
Cambridge, MA, USA
and
Institute for Information Transmission
Problems of Russian Academy of Sciences,
Moscow, Russia

V. G.
Department of Mathematics,
Massachusetts Institute of Technology,
Cambridge, MA, USA
vadicgor@gmail.com
and
Institute for Information Transmission
Problems of Russian Academy of Sciences,
Moscow, Russia

A. G.
Department of Mathematics,
Massachusetts Institute of Technology,
Cambridge, MA, USA