Journal of Functional Analysis 276 (2019) 3067–3169

Contents lists available at ScienceDirect

Journal of Functional Analysis

www.elsevier.com/locate/jfa

Log-gases on quadratic lattices via discrete loop equations and q-boxed plane partitions

Evgeni Dimitrov *, Alisa Knizel

Department of Mathematics, Columbia University, New York, NY, USA

ARTICLE INFO

Article history:
Received 29 April 2018
Accepted 17 December 2018
Available online 25 January 2019
Communicated by D. Stroock

Keywords:
Log-gases
Loop equations

ABSTRACT

We study a general class of log-gas ensembles on (shifted) quadratic lattices. We prove that the corresponding empirical measures satisfy a law of large numbers and that their global fluctuations are Gaussian with a universal covariance. We apply our general results to analyze the asymptotic behavior of a q-boxed plane partition model introduced by Borodin, Gorin and Rains. In particular, we show that the global fluctuations of the height function on a fixed slice are described by a one-dimensional section of a pullback of the two-dimensional Gaussian free field.

Our approach is based on a q-analogue of the Schwinger–Dyson (or loop) equations, which originate in the work of Nekrasov and his collaborators, and extends the methods developed by Borodin, Gorin and Guionnet to quadratic lattices.

Contents

* Corresponding author.
E-mail addresses: esd2138@columbia.edu (E. Dimitrov), knizel@math.columbia.edu (A. Knizel).

https://doi.org/10.1016/j.jfa.2018.12.008
0022-1236/

1. Introduction

1.1. Preface

A β-ensemble (or continuous log-gas) is a probability distribution $\mathbb{P}_N^c$ on N-tuples of ordered real numbers $x_1 < x_2 < \cdots < x_N$ with density proportional to

$$\prod_{1\le i<j\le N} (x_j - x_i)^\beta \cdot \prod_{i=1}^N \exp(-NV(x_i)), \tag{1.1}$$

where $V(x)$ is a continuous function called *potential*. The study of continuous log-gases for general potentials is a rich subject that is of special interest to random matrix theory, see e.g. [1,32,53,60]. For example, when $V(x) = x^2$ and $\beta = 1, 2, 4$ distribution (1.1) is the joint density of the eigenvalues of random matrices from the Gaussian Orthogonal/Unitary/Symplectic ensembles [1].

Recently, [15] initiated a detailed study of a particular discrete version of (1.1) called *discrete β-ensembles* or *discrete log-gases*. These are probability distributions depending on a parameter $\theta = \beta/2 > 0$ and a positive real-valued function $w(x; N)$ of the form

$$\begin{aligned} &\mathbb{P}_N^d(\lambda_1, \dots, \lambda_N) \propto \prod_{1\le i<j\le N} H_\theta(\lambda_i, \lambda_j) \prod_{i=1}^N w(\lambda_i; N), \text{ with} \\ &H_\theta(\lambda_i, \lambda_j) = \frac{\Gamma(\lambda_j - \lambda_i + 1)\Gamma(\lambda_j - \lambda_i + \theta)}{\Gamma(\lambda_j - \lambda_i)\Gamma(\lambda_j - \lambda_i + 1 - \theta)}, \end{aligned} \tag{1.2}$$

where $\lambda_i = x_i + \theta(i-1)$ and $x_1 \le x_2 \le \cdots \le x_N$ are integers. The interest in these discrete models comes from integrable probability; specifically, due to their connection to uniform random tilings, (z, w)-measures, Jack measures, etc.

In the present paper we consider the following two-parameter generalization of discrete β-ensembles

$$\begin{aligned} &\mathbb{P}_N(\ell_1, \dots, \ell_N) \propto \prod_{1\le i<j\le N} H_\theta^{q,v}(\ell_i, \ell_j) \cdot \prod_{i=1}^N w(\ell_i; N), \text{ with} \\ &H_\theta^{q,v}(\ell_i, \ell_j) \\ &\quad = q^{-2\theta\lambda_j} \frac{\Gamma_q(\lambda_j - \lambda_i + 1)\Gamma_q(\lambda_j - \lambda_i + \theta)}{\Gamma_q(\lambda_j - \lambda_i)\Gamma_q(\lambda_j - \lambda_i + 1 - \theta)} \frac{\Gamma_q(\lambda_j + \lambda_i + v + 1)\Gamma_q(\lambda_j + \lambda_i + v + \theta)}{\Gamma_q(\lambda_j + \lambda_i + v)\Gamma_q(\lambda_j + \lambda_i + v + 1 - \theta)}, \end{aligned} \tag{1.3}$$

where $\ell_i = q^{-\lambda_i} + u \cdot q^{\lambda_i}$, $u = q^v$ and λ_i are as in the definition of the discrete β-ensembles, while $q \in (0, 1)$ and $v \in (1, \infty)$. The measures (1.2) are recovered from (1.3) by setting $u \to 0$ and $q \to 1$. We interpret the random vector $(\ell_1, \dots, \ell_N)$ as locations of N

particles. If $\theta = 1$ then all particles ℓ_i live on the same space $\mathfrak{Z} := \{q^{-x} + u \cdot q^x : x \in \mathbb{Z}\}$, and we refer to the set $\mathfrak{Z}$ as a *quadratic lattice* in the spirit of [58] (note that in this case $H_\theta^{q,v}(\ell_i, \ell_j) = (\ell_i - \ell_j)^2$). For general θ we call the class of measures (1.3) *discrete β-ensembles on shifted quadratic lattices*.

Our study is motivated by random matrix theory on one side, and by integrable models on the other. We first investigate $\mathbb{P}_N$ for a general choice of weights w in the *multu-cut* and *fixed filling fractions* regime. We prove that these systems obey a law of large numbers under a certain scaling as N goes to infinity and also show that their global fluctuations are asymptotically Gaussian with a *universal covariance*. The same phenomenon is present in the case of discrete and continuous log-gases. Subsequently, we apply our general results to a class of tiling models that was introduced in [16] and obtain explicit formulas for their limit shape and global fluctuations. The tiling model we investigate corresponds to a special case of (1.3) when $\theta = 1$, and we remark that for general $\theta > 0$ the interaction term $H_\theta^{q,v}(\ell_i, \ell_j)$ can be linked to *Macdonald–Koorwinder polynomials* [49] similarly to how $H_\theta(\lambda_i, \lambda_j)$ in (1.2) is linked to Jack symmetric polynomials, see also Remark 2.1.2.

1.1.1. Log-gases

The probability measures from (1.1) and (1.2) have been extensively studied in the past, see [1,32,53,60] for $\mathbb{P}_N^c$ and [13,25,31,39,40] for $\mathbb{P}_N^d$ among many others.

Under weak assumptions on the potential $V(x)$ or weight function $w(x; N)$, continuous and discrete log-gases exhibit a *law of large numbers* as $N \to \infty$. Specifically, if one forms the (random) empirical measures

$$\mu_N = \frac{1}{N} \sum_{i=1}^{N} \delta\left(x_i / N\right), \quad \text{where } (x_1, \ldots, x_N) \text{ is } \mathbb{P}_N^{c,d}\text{-distributed},$$

then the measures μ_N converge weakly in probability to a deterministic measure μ, called the *equilibrium measure*. In the continuous case with $V(x) = x^2$ this statement goes back to the work of Wigner [70], and is called *Wigner's semicircle law*. The analogous statements for generic $V(x)$ were proved much later, see [6,20,38]. In discrete settings similar law of large numbers type results were obtained in [31,39,40]. In both cases the equilibrium measure μ is the solution to a suitable variational problem and one establishes the convergence of μ_N to μ by proving large deviation estimates. In essence, μ maximizes the density (1.1) or (1.2) and the large deviation estimates show that μ_N concentrate around that maximum.

The next order asymptotics asks about the fluctuations of $\mu - \mu_N$ as $N \to \infty$. One natural way to analyze this difference is to form the pairings with smooth test functions f and consider the asymptotic behavior of the random variables

$$N \left(\int_{\mathbb{R}} f(x) \mu_N(x) - \int_{\mathbb{R}} f(x) \mu(x) \right), \quad \text{as } N \to \infty. \tag{1.4}$$

There is an efficient method, which establishes that the limits of (1.4) are Gaussian in a very general setup and its key ingredient is the so-called *loop equations* (also known as Schwinger–Dyson equations), see [17,18,38,50,64] and references therein. These are functional equations for certain observables of the log-gases (1.1) that are related to the Stieltjes transforms of the empirical measure μ_N and their cumulants. Since their introduction loop equations have become a powerful tool for studying not only global fluctuations but also *local universality* for random matrices [5,19].

In [15] the authors presented an analogue of the above method for discrete β-ensembles. They introduced *discrete loop equations* and used them to establish that the limits of (1.4) for the measures in 1.2) are Gaussian with a covariance that is the same as in the continuous case for a large class potentials. These discrete loop equations originate in the work of Nekrasov [57] and are also called *Nekrasov's equations*. The central limit theorem for (1.2) had been previously known for various very specific *integrable* choices of the potential, see e.g. [13,21,62]. The main contribution of [15] is that it establishes *general conditions* on the potential $V(x)$ that lead to the asymptotic Gaussianity of (1.4). Similarly to the continuous case, discrete loop equations have become a valuable tool to study not only global fluctuations [15] but also edge universality for discrete β-ensembles [33].

In the present paper we establish the universality type results for the global fluctuations of discrete β-ensembles on shifted quadratic lattices (1.3). To obtain the law of large numbers we use a similar combination of large deviation estimates and variational problems that proved to be successful for $\mathbb{P}_N^{c/d}$. In order to study the next order fluctuations we introduce a new version of discrete loop equations for a quadratic lattice, which we also call Nekrasov's equations, and view the latter as one of the main contributions of this paper. We remark that it is hard to guess that there even exists an analogue of the Nekrasov's equation in this setting, since it is a very subtle equation which reflects some specific algebraic structure of the system. Equipped with these new equations, we establish global central limit theorems for log-gases on a quadratic lattice for a multi-cut general potential by adapting the arguments in [15].

Our main motivation for considering the class of measures $\mathbb{P}_N$ comes from an interesting tiling model introduced in [16] which we describe next.

1.1.2. The q-Racah tiling model

Consider a hexagon drawn on a regular triangular lattice, whose side lengths are given by integers $a, b, c \geq 1$, see Fig. 1. We are interested in random tilings of such a hexagon by rhombi, also called lozenges (these are obtained by gluing two neighboring triangles together). There are three types of rhombi that arise in such a way, and they are all colored differently in Fig. 1. This model also has a natural $3D$ interpretation as a boxed plane partition or, equivalently, a random stepped surface formed by a stack of boxes. One can assign to every lattice vertex (i, j) inside the hexagon an integer $h(i, j)$, which reflects the height of the $3D$ stack at that point, see an example in Fig. 1. One typically calls h the *height function* and formulates results in terms of it.

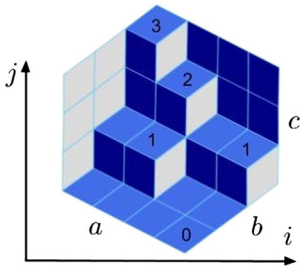

Fig. 1. Tiling of a hexagon and the corresponding height function.

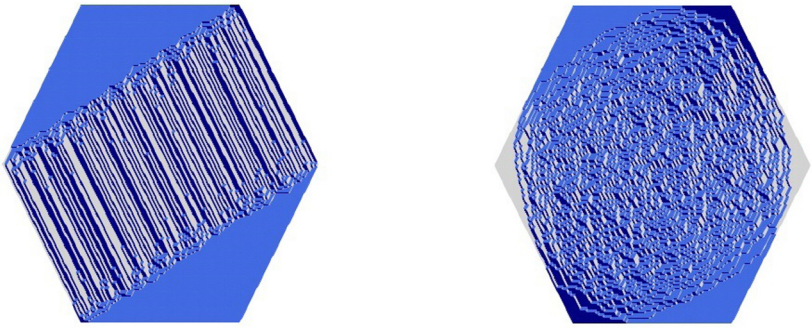

Fig. 2. A simulation for $a = 80$, $b = 80$, $c = 80$. On the left picture the parameters are $\kappa^2 = -1$, $q = 0.8$, and on the right picture the parameters are $\kappa^2 = -1$, $q = 0.98$.

The probability measures on the set of tilings that we consider were introduced in [16] and form a 2-parameter generalization of the uniform distribution. Denoting the two parameters by q and κ, one defines the weight of a tiling as the product of simple factors $(\kappa q^{j-(c+1)/2} - q^{-j+(c+1)/2}/\kappa)$ over all horizontal rhombi $\diamond$, where j is the coordinate of the topmost point of the rhombus. The dependence of the factors on the location of the lozenge makes the model *inhomogeneous*. Note that the uniform measure on tilings is recovered if one sends $\kappa \to 0$ and $q \to 1$. Other interesting cases include $\kappa \to 0$, then the weight becomes proportional to q^{-V} (here V refers to the number of boxes in the $3D$ interpretation). In addition, the same way the Hahn orthogonal polynomial ensemble arises in the analysis of uniform lozenge tilings, our measures are related to the q-Racah orthogonal polynomials. In this sense, the model goes all the way up to the top of the Askey scheme [46], and we call it the q-Racah tiling model.

We believe that the q-Racah tiling model is a source of rich and interesting structures that are worth investigating. The presence of two parameters allows one to consider various limit regimes that lead to quite different behavior of the system as can be seen in Fig. 2. One of the central goals of this paper is to understand the asymptotic behavior of the height function of the q-Racah tiling model when the sides of the hexagon become large, and simultaneously $q \to 1, \kappa \to \kappa_0$, where $\kappa_0 \in (0, 1)$ is fixed, see Fig. 3 for a sample tiling in this case.

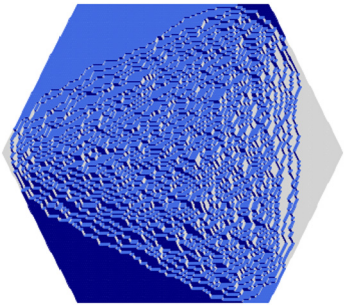

Fig. 3. A random tiling of a hexagon of side lengths $a = 50$, $b = 60$, $c = 40$ for $\kappa^2 = 0.05$ and $q = 0.99$.

It turns out that one can relate one-dimensional sections of the q-Racah tiling model to measures from (1.3) with $\theta = 1$. We will elaborate on this point later in Sections 7 and 9.2, but the identification goes as follows. One places a particle in the center of each horizontal lozenge $\diamond$ and takes a vertical section of the model; the resulting "holes" (positions, where there are no particles) form an N-point process. Under a suitable change of variables this point process has the same distribution as (1.3) for a set of parameters and weight w that depend on the location of the vertical slice. Using this identification, our general results for log-gases on (shifted) quadratic lattices imply a law of large numbers and central limit theorem for the height function h of the tiling model.

Informally, our law of large numbers states that there exists deterministic limit shape $\hat{h}$ and the random height functions h concentrate near it with high probability as the parameters of the model scale to their critical values. An important feature of our model is that the limit shape develops *frozen facets* where the height function is linear. In addition, the frozen facets are interpolated by a connected disordered *liquid region*. In terms of the tiling a frozen facet corresponds to a region where asymptotically only one type of lozenge is present and in the liquid region one sees lozenges of all three types, see Fig. 3.

Similar concentration phenomena for the random height function in the case of the uniform measure and the measure proportional to q^{-V} are well-understood. In particular, in these cases convergence of the random height function to a deterministic function for a large class of domains was established in [24,26,27,37,44,61]. The limit shape is given by the unique solution of a suitable variational problem. For the q-Racah tiling model we compute the limit shape explicitly introducing a method, which we believe to be novel. This method uses discrete Riemann–Hilbert problems.

The next order asymptotics we obtain show that the one-dimensional fluctuations of the height function around the limit shape are Gaussian with an explicit covariance kernel. An important additional contribution of our work is the introduction of a (rather nontrivial) *complex structure* Ω on the liquid region. The significance of this map is that the fluctuations of h on fixed vertical slices are asymptotically described by the one-dimensional sections of the pullback of the Gaussian free field (GFF for short) on the upper half-plane $\mathbb{H}$ under the map Ω – see Theorem 7.2.4 for the precise statement. This

result admits a natural two-dimensional generalization, which we formulate as Conjecture 8.4.1 in the main text. At this time our methods only provide access to the global fluctuations at fixed vertical sections of the model, and so we cannot establish the full $2D$ result.

The GFF is believed to be a universal scaling limit for various models of random surfaces in $\mathbb{R}^3$. The appearance of the GFF in tiling models with no frozen zones dates back to [41,42] and the fluctuations of the liquid region for a random tiling model containing both frozen facets and a liquid region were first studied in [13]. In case of the uniform measure on domino and lozenge tilings the convergence to the GFF has been established for a wide class of domains in [22,23,62], but there are no results in this direction for more general measures. One possible reason that explains why the GFF was not recognized in the q-Racah tiling model is the rather non-trivial change of coordinates that makes the correct covariance structure appear (see Section 8), and already in the q^{-V} (or $\kappa = 0$) case our result is new. We remark that there is a natural complex coordinate on the liquid region defined by the so-called *complex slope*, which in the uniform tiling case is known to be intimately related to the complex structure that gives rise to the GFF. For the q-Racah tiling model an expression for the complex slope was obtained in [16] and we connect it to our complex structure Ω through an explicit functional dependence, see Remark 8.2.2.

1.2. Main results

We present here our main results for the log-gas on a quadratic lattice and forego stating our results on the q-Racah tiling model until the main text – Section 7.2 – since it requires the introduction of more notation. Moreover, to simplify the discussion in the introduction we will formulate our results for the one-cut case and $\theta = 1$. The general statement of the law of large numbers is given in Theorem 3.1.1 and the general statement of the central limit theorem is given in Theorem 5.2.7.

Let us first explain our regularity assumptions on the parameters and the weight function. We assume we are given parameters $\mathtt{q} \in (0,1)$, $\mathtt{M} \geq 1$ and $\mathtt{u} \in [0,1)$. In addition, let $q_N \in (0,1)$, $M_N \in \mathbb{N}$ and $u_N \in [0,1)$ be sequences of parameters such that

$$M_N \geq N-1 \text{ and } \max\left(N^2\left|q_N - \mathtt{q}^{1/N}\right|, |M_N - N\mathtt{M}|, N|u_N - \mathtt{u}|\right) \leq A_1, \quad \text{for some } A_1 > 0. \tag{1.5}$$

We assume that $w(x;N)$ has the form

$$w(x;N) = \exp\left(-NV_N(x)\right),$$

for a function V_N that is continuous in the intervals $[1+u_N, q_N^{-M_N} + u_N q_N^{M_N}]$ and such that

$$|V_N(s) - V(s)| \leq A_2 \cdot N^{-1} \log(N), \text{ where } V \text{ is continuous and } |V(s)| \leq A_3, \tag{1.6}$$

for some positive constants $A_2, A_3 > 0$. We also require that $V(s)$ is differentiable and for some $A_4 > 0$ there is a bound

$$|V'(s)| \leq A_4 \cdot \left[1 + |\log|s - 1 - \mathtt{u}|| + |\log|s - \mathtt{q}^{-\mathtt{M}} - \mathtt{u}\mathtt{q}^{\mathtt{M}}||\right], \text{ for } s \in \left[1 + \mathtt{u}, \mathtt{q}^{-\mathtt{M}} + \mathtt{u}\mathtt{q}^{\mathtt{M}}\right]. \tag{1.7}$$

We let $\mathbb{P}_N$ be as in (1.3) for $q = q_N, q^v = u = u_N, M = M_N, N$ and weight function $w(\cdot) = w(\cdot; N)$.

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

$$\mu_N = \frac{1}{N} \sum_{i=1}^{N} \delta(\ell_i), \quad \text{where } (\ell_1, \dots, \ell_N) \text{ is } \mathbb{P}_N\text{-distributed.}$$

Theorem 1.2.1. *There is a deterministic, compactly supported and absolutely continuous probability measure $\mu(x)dx$[1] such that μ_N concentrate (in probability) near μ. More precisely, for each Lipschitz function $f(x)$ defined in a real neighborhood of the interval $[1 + \mathtt{u}, \mathtt{q}^{-\mathtt{M}} + \mathtt{u}\mathtt{q}^{\mathtt{M}}]$ and each $\varepsilon > 0$ the random variables*

$$N^{1/2-\varepsilon} \left| \int_{\mathbb{R}} f(x)\mu_N(dx) - \int_{\mathbb{R}} f(x)\mu(x)dx \right| \tag{1.8}$$

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

Remark 1.2.2. Theorem 1.2.1 is a special case of Theorem 3.1.1, where we extend the statement to the multi-cut regime with fixed filling fractions and for general $\theta > 0$.

To obtain our central limit theorem we need to impose certain analyticity conditions on the weight $w(x; N)$ that we now detail. We assume that we have an open set $\mathcal{M} \subset \mathbb{C} \setminus \{0, \pm\sqrt{\mathtt{u}}\}$, such that for large N

$$\left(\left[q_N^1, q_N^{-M_N-1}\right] \cup \left[u_N q_N^{M_N+1}, u_N q_N^{-1}\right]\right) \subset \mathcal{M}.$$

In addition, we require that for all sufficiently large N there exist analytic functions Φ_N^+, Φ_N^- on $\mathcal{M}$ such that for $z \in \mathcal{M}$ and $\sigma_N(z) = z + u_N z^{-1}$ the following holds

$$\frac{w(\sigma_N(z); N)}{w(\sigma_N(q_N z); N)} = \frac{q_N(z^2 - u_N)\Phi_N^+(z)}{(q_N^2 z^2 - u_N)\Phi_N^-(z)}. \tag{1.9}$$

Moreover, the functions $\Phi_N^{\pm}$ satisfy the following vanishing conditions

[1] Throughout the paper we denote the density of a measure μ by $\mu(x)$.

$$\Phi_N^+\left(q_N^{-M_N-1}\right) = \Phi_N^-(1) = \Phi_N^+\left(u_N q_N^{-1}\right) = \Phi_N^-\left(u_N q_N^{M_N}\right) = 0,$$

and asymptotic expansion

$$\Phi_N^-(z) = \Phi^-(z) + \varphi_N^-(z) + O\left(N^{-2}\right) \text{ and } \Phi_N^+(z) = \Phi^+(z) + \varphi_N^+(z) + O\left(N^{-2}\right),$$

where $\varphi_N^{\pm}(z) = O(N^{-1})$ and the constants in the big O notation are uniform over z in compact subsets of $\mathcal{M}$. All the aforementioned functions are holomorphic in $\mathcal{M}$.

The assumptions in (1.9) are the analogues of Assumptions 3 and 5 in [15], and similarly to that paper their importance to the analysis comes from the following observation, which is the starting point for our results. We discuss the general β setup and the corresponding Nekrasov's equation in Section 4.

Theorem 1.2.3 *(Nekrasov's equation). Suppose that (1.9) hold and define*

$$R_N(z) = \Phi_N^-(z) \cdot \mathbb{E}_{\mathbb{P}_N}\left[\prod_{i=1}^N \frac{\sigma_N(q_N z) - \ell_i}{\sigma_N(z) - \ell_i}\right] + \Phi_N^+(z) \cdot \mathbb{E}_{\mathbb{P}_N}\left[\prod_{i=1}^N \frac{\sigma_N(z) - \ell_i}{\sigma_N(q_N z) - \ell_i}\right]. \quad (1.10)$$

Then $R_N(z)$ is analytic in $\mathcal{M}$. If $\Phi_N^{\pm}(z)$ are polynomials of degree at most d, then so is $R_N(z)$.

Remark 1.2.4. If μ denotes the equilibrium measure from Theorem 1.2.1, and $G_\mu(z) = \int_{\mathbb{R}} \frac{\mu(x)dx}{z-x}$ is its Stieltjes transform then as explained in Section 4 one has

$$\lim_{N\to\infty} \mathbb{E}_{\mathbb{P}_N}\left[\prod_{i=1}^N \frac{\sigma_N(q_N z) - \ell_i}{\sigma_N(z) - \ell_i}\right] = \exp\left(\mathfrak{G}(z)\right) \text{ with } \mathfrak{G}(z) = \log(\mathtt{q}) \cdot (z - \mathtt{u} z^{-1}) \cdot G_\mu(z + \mathtt{u} z^{-1}).$$

In this sense, the Nekrasov's equation lead to a functional equation for $\mathfrak{G}(z)$, and our central limit theorem is a consequence of a careful analysis of the lower order terms of the above limit. We remark that in [15] the expression that appears in the exponent above is directly the Stieltjes transform $G_\mu(z)$ and not a modified version of it as in our case, which increases the technical difficulty of our arguments. The appearance of $\mathfrak{G}$ is a novel feature that comes from working on a quadratic lattice and we give some explanation of it in Remark 4.2.3.

Our central limit theorem requires that the equilibrium measure μ satisfies Assumption 5 in Section 2.1, which roughly ensures that μ has a single *band* in $[1 + \mathtt{u}, \mathtt{q}^{-\mathtt{M}} + \mathtt{u}\mathtt{q}^{\mathtt{M}}]$. In our context, a band is a maximal interval (a, b) such that $0 < \mu(x) < f_{\mathtt{q}}(\sigma_{\mathtt{q}}^{-1}(x))^{-1}$, where $\sigma_{\mathtt{q}}(x) = \mathtt{q}^{-x} + \mathtt{u}\mathtt{q}^{x}$ and $f_{\mathtt{q}}(x) = \frac{d}{dx}\sigma_{\mathtt{q}}\left(\mathtt{q}^{-x}\right)$ (see also Section 4.2). The parameters α_1, β_1 that appear in the next Theorem 1.2.5 are then precisely the endpoints of this band.

Theorem 1.2.5. *Suppose that (1.5), (1.6), (1.7), (1.9) and that (technical) Assumption 5 from Section 2.2 hold. For $m \geq 1$ let $f_1, \dots, f_m$ be real analytic functions in a neighborhood of $[1+\mathtt{u}, \mathtt{q}^{-\mathtt{M}} + \mathtt{u}\mathtt{q}^{\mathtt{M}}]$ and define*

$$\mathcal{L}_{f_i} = N \int_{\mathbb{R}} f_j(x)\mu_N(dx) - N\mathbb{E}_{\mathbb{P}_N}\left[\int_{\mathbb{R}} f_j(x)\mu_N(dx)\right] \quad \text{for } i = 1, \dots, m.$$

Then the random variables $\mathcal{L}_{f_i}$ converge jointly in the sense of moments to an m-dimensional centered Gaussian vector $X = (X_1, \dots, X_m)$ with covariance[2]

$$Cov(X_i, X_j) = \frac{1}{(2\pi i)^2} \oint_{\Gamma} \oint_{\Gamma} f_i(s) f_j(t) \mathcal{C}(s,t) ds dt, \quad \text{where}$$

$$\mathcal{C}(s,t) = -\frac{1}{2(s-t)^2}\left(1 - \frac{(s-\alpha_1)(t-\beta_1) + (t-\alpha_1)(s-\beta_1)}{2\sqrt{(s-\alpha_1)(s-\beta_1)}\sqrt{(t-\alpha_1)(t-\beta_1)}}\right), \tag{1.11}$$

where $\alpha_1, \beta_1 \in [1+\mathtt{u}, \mathtt{q}^{-\mathtt{M}} + \mathtt{u}\mathtt{q}^{\mathtt{M}}]$ are given in Assumption 5 and Γ is a positively oriented contour that encloses the interval $[1+\mathtt{u}, \mathtt{q}^{-\mathtt{M}} + \mathtt{u}\mathtt{q}^{\mathtt{M}}]$.

We emphasize that the covariance in (1.11) depends only on α_1, β_1, and is not sensitive to other features of the equilibrium measure μ. Furthermore, the covariance is the same as for the continuous log-gases, cf. [38, Theorem 2.4] and [60, Chapter 3]. Thus, the discreteness of the model is invisible on the level of the central limit theorem, which is consistent with what was observed for the discrete β-ensembles in [15].

Remark 1.2.6. Theorem 1.2.5 is a special case of Theorem 5.2.7, where we extend the statement to the multi-cut regime with fixed filling fractions and general $\theta > 0$.

Remark 1.2.7. Observe that the covariance $\mathcal{C}(s,t)$ has no singularity when $s = t$, since the RHS of (1.11) has a finite limit when s tends to t.

Outline

In Section 2 we describe the general framework of our study, the scaling regime we consider and the assumptions on the weight w. In Section 3 we establish a general law of large numbers as Theorem 3.1.1. Nekrasov's equation is discussed in Section 4. Sections 5

[2] Throughout the paper, given $a, b \in \mathbb{R}$ with $a < b$, we write $f(z) = \sqrt{(z-a)(z-b)}$ to mean

$$f(z) = \begin{cases} \sqrt{z-a}\sqrt{z-b} & \text{when } z \in \mathbb{C} \setminus (-\infty, b], \\ -\sqrt{a-z}\sqrt{b-z} & \text{when } z \in (-\infty, a). \end{cases}$$

Observe that in this way f is holomorphic on $\mathbb{C} \setminus [a,b]$, cf. Theorem 2.5.5 in [68].

and 6 contain the proof of Theorem 5.2.7 (our general central limit theorem). A detailed description of the q-Racah tiling model is given in Section 7 and we give the proof of our results about its random height function in Section 8. Section 9 provides the verification that the tiling model fits into the general framework of Section 2. Finally, Section 10 contains the asymptotic analysis of the Nekrasov's equation for the tiling model using discrete Riemann–Hilbert problems.

Acknowledgments

The authors are deeply grateful to Alexei Borodin, Vadim Gorin and Alice Guionnet for very helpful discussions. For the second author the financial support was partially available through NSF grant DMS:1704186 and the project started when the second author was still a PhD student at Massachusetts Institute of Technology. We also want to thank the hospitality of PCMI during the summer of 2017 supported by NSF grant DMS:1441467.

2. General setup

In this section we describe the general setting of a multi-cut, fixed filling fractions model that we consider and list the specific assumptions we make about it.

2.1. Definition of the system

We begin with some necessary notation. Let $q \in (0,1)$, $M \in \mathbb{Z}_{\geq 0}$, $u \in [0,1)$, $\theta > 0$ and $N \in \mathbb{N}$ be such that $M \geq N-1$. For such parameters we set

$$\begin{aligned} X_N &= \{(x_1,\dots,x_N) : x_1 \leq x_2 \leq \cdots \leq x_N, x_i \in \mathbb{Z} \text{ and } 0 \leq x_i \leq M-N+1\},\\ \mathbb{W}_N^\theta &= \{(\lambda_1,\dots,\lambda_N) : \lambda_i = x_i + (i-1)\cdot\theta, \text{ with } (x_1,\dots,x_N) \in X_N\},\\ \mathcal{L}_N^\theta &= \{(\ell_1,\dots,\ell_N) : \ell_i = q^{-\lambda_i} + uq^{\lambda_i}, \text{ with } (\lambda_1,\dots,\lambda_N) \in \mathbb{W}_N^\theta\}. \end{aligned} \tag{2.1}$$

We interpret the elements $(\ell_1,\dots,\ell_N)$ in $\mathcal{L}_N^\theta$ as locations of N particles. If $\theta = 1$ then all particles ℓ_i live on the same space $GCLTZ := \{q^{-x} + uq^x : x \in \mathbb{Z}\}$, and we refer to the set $\mathfrak{Z}$ as a *quadratic lattice* in the spirit of [58]. On the other hand, for general $\theta > 0$ the particle ℓ_k lives on an appropriately *shifted* quadratic lattice $\mathfrak{Z}_k^\theta := \{q^{-x-(k-1)\theta} + uq^{x+(k-1)\theta} : x \in \mathbb{Z}\}$. This is similar to the setup in [15]. Throughout the text we will frequently switch from ℓ_i's to λ_i's or x_i's without mention using

$$x_i \in \mathbb{Z},\ \lambda_i = x_i + (i-1)\cdot\theta \in (i-1)\cdot\theta + \mathbb{Z},\ \ell_i = q^{-\lambda_i} + uq^{\lambda_i} \in \mathfrak{Z}_i^\theta. \tag{2.2}$$

We typically choose the coordinate system that leads to the most transparent formulas or arguments.

Our goal is to define probability measures on a subset of $\mathcal{L}_N^\theta$, where particles are split into k groups of prescribed sizes, living on k disjoint prescribed segments. We start by fixing $k \in \mathbb{N}$, which denotes the number of segments (or groups). For each $N \in \mathbb{N}$ we take k integers $n_1(N), \dots, n_k(N)$, set $N_j = \sum_{i=1}^{j} n_i(N)$ with the convention $N_0 = 0$ and assume $N_k = N$. The numbers $n_i(N)$ indicate the number of particles in each segment (or group). In addition, we suppose that we have $2k$ integers $a_i(N), b_i(N)$ such that $0 \le a_i(N) \le b_i(N) - 1 \le M - N + 1$ for $i = 1, \dots, k$ and $b_i(N) \le a_{i+1}(N)$ for $i = 1, \dots, k-1$. With the above data we define the state space of our N-point process as follows.

Definition 2.1.1. The state space $\mathfrak{X}_N^\theta$ consists of N-tuples $\ell = (\ell_1, \dots, \ell_N) \in \mathcal{L}_N^\theta$ such that $a_j(N) \le x_i \le b_j(N) - 1$, see (2.2), whenever $N_{j-1} + 1 \le i \le N_j$ for $i = 1, \dots, N$ and $j = 1, \dots, k$. For future use we also denote $\alpha_i(N) = q^{-a_i(N) - N_{i-1} \cdot \theta} + uq^{a_i(N) + N_{i-1} \cdot \theta}$, $\beta_i(N) = q^{-b_i(N) - (N_i - 1) \cdot \theta} + uq^{b_i(N) + (N_i - 1) \cdot \theta}$ and $\beta_i^-(N)$ the largest element in $\mathfrak{Z}_{N_i}^\theta$ less than $\beta_i(N)$ for $i = 1, \dots, N$.

Utilizing Definition 2.1.1 we define a probability measure $\mathbb{P}_N$ on $\mathfrak{X}_N^\theta$ through

$$\mathbb{P}_N(\ell_1, \dots, \ell_N) = \frac{1}{Z_N} \cdot \prod_{1 \le i < j \le N} H(\ell_i, \ell_j) \cdot \prod_{i=1}^{N} w(\ell_i; N), \text{ where}$$

$$H(\ell_i, \ell_j) = q^{-2\theta\lambda_j} \frac{\Gamma_q(\lambda_j - \lambda_i + 1)\Gamma_q(\lambda_j - \lambda_i + \theta)}{\Gamma_q(\lambda_j - \lambda_i)\Gamma_q(\lambda_j - \lambda_i + 1 - \theta)} \frac{\Gamma_q(\lambda_j + \lambda_i + v + 1)\Gamma_q(\lambda_j + \lambda_i + v + \theta)}{\Gamma_q(\lambda_j + \lambda_i + v)\Gamma_q(\lambda_j + \lambda_i + v + 1 - \theta)}. \tag{2.3}$$

Here Z_N is a normalization constant (called the *partition function*), v is such that $q^v = u$, and $w(x; N)$ is a weight function, which is assumed to be *positive* for $x \in \cup_{i=1}^{k} [\alpha_i(N), \beta_i^-(N)]$. We recall

$$\Gamma_q(x) = (1-q)^{1-x} \frac{(q;q)_\infty}{(q^x;q)_\infty} \text{ where } (a;q)_\infty = \prod_{k=0}^{\infty} (1 - aq^k) \text{ and it satisfies}$$

$$\frac{\Gamma_q(x+1)}{\Gamma_q(x)} = \frac{1 - q^x}{1 - q}. \tag{2.4}$$

Let us remark on a couple of properties of $\mathbb{P}_N$. Firstly, the measure $\mathbb{P}_N$ when $u = 0$ was considered in [15]. Specifically, our measure agrees with equation (82) of that paper with λ_i replaced with ℓ_i and $w(\ell_i; N)$ replaced with $w(\ell_i; N) \cdot q^{-\theta(N-1)\ell_i}$. In addition, from [2, Theorem 10.2.4]

$$\frac{\Gamma_q(x+\alpha)}{\Gamma_q(x)} = (1-q)^{-\alpha} \frac{(q^x;q)_\infty}{(q^{x+\alpha};q)_\infty} \sim (1-q)^{-\alpha}(1-y)^\alpha \text{ as } q \to 1^- \text{ and } q^x \to y \in [0,1)$$

and setting $q^{-\lambda_i} = y_i$ for $i = 1, \dots, N$ we have

$$H(\ell_i, \ell_j) \sim (1-q)^{-4\cdot\theta} \cdot y_j^{-2\theta}(1 - y_i y_j^{-1})^{2\theta}(1 - u y_i^{-1} y_j^{-1})^{2\theta} = (1-q)^{-4\cdot\theta} \cdot (\ell_j - \ell_i)^{2\theta}, \tag{2.5}$$

which is why we view $\mathbb{P}_N$ as a discretization of the general $\beta = 2\theta$ log-gas to a quadratic lattice. The latter is particularly obvious when $\theta = 1$, since then we have

$$\mathbb{P}_N(\ell_1, \dots, \ell_N) = \frac{(1-q)^{-2\cdot N(N-1)}}{Z_N} \cdot \prod_{1 \le i < j \le N} (\ell_j - \ell_i)^2 \prod_{i=1}^{N} w(\ell_i; N).$$

The above connection to log-gases motivates our choice to work with the particles $(\ell_1, \dots, \ell_N)$ and not for example $(\lambda_1, \dots, \lambda_N)$, although most results can be formulated in terms of the latter.

Remark 2.1.2. One way to understand the interaction $H(\ell_i, \ell_j)$ in (2.3) is that it is an *integrable* extension of the interaction $(\ell_j - \ell_i)^2$ to general $\theta > 0$. This should be viewed as an analogue to how (1.2) is a general $\theta > 0$ version of $(\lambda_j - \lambda_i)^2$, and the integrability of that extension can be traced to *discrete Selberg integrals* and *Jack symmetric polynomials*, where analogous expressions appear, see [15, Section 1]. One source of motivation for why $H(\ell_i, \ell_j)$ is the correct generalization of $(\ell_j - \ell_i)^2$ in the setting of a quadratic lattice comes from a connection to Macdonald–Koorwinder polynomials [49] as we detail below.

Following the notation in [63] we let $K_\mu^{(n)}(\cdot; q, t; t_0, t_1, t_2, t_3)$ denote the BC_n-symmetric Koorwinder polynomial in n variables. In addition, if $\mu \subset (m)^n$ we define $\tilde{\mu} \subset (n)^m$ through $\tilde{\mu}_j = n - \mu'_{m-j+1}$. Taking the product of $K_\mu^{(n)}$ and $K_{\tilde{\mu}}^{(m)}$ at the principal and dual principal specializations (such products appear in the dual Cauchy identity for Koorwinder polynomials [55]) gives

$$\begin{aligned} &K_\mu^{(n)}(t^{n-i}t_0; q, t; t_0, t_1, t_2, t_3) \cdot K_{\tilde{\mu}}^{(m)}(q^{m-j}t_0; t, q; t_0, t_1, t_2, t_3) = \\ &C(n, m, \theta) \cdot \prod_{1 \le i < j \le n} H(\ell_i, \ell_j) \cdot \prod_{i=1}^{n} W(\ell_i), \end{aligned} \tag{2.6}$$

where $H(\ell_i, \ell_j)$ is as in (2.3) with $t = q^\theta$, $u = q^v = t_0 t_1 t_2 t_3 / q$ and $C(n, m, \theta)$ is a μ-independent constant. As before we have $\ell_i = q^{-\lambda_i} + u q^{\lambda_i}$ and $\lambda_i = \mu_{n-i+1} + (i-1)\cdot\theta$ (notice that λ's are indexed in increasing order, while μ's are indexed in decreasing order as is typical for partitions). In addition, we have

$$W(\ell_i) = (1 - q^{2\lambda_i} u) \frac{(q^{1+\lambda_i}, q^{m-\lambda_i+1+\theta(n-1)}, q^{\lambda_i+m+1+\theta(n-1)}u; q)_\infty}{(q^{\lambda_i}u; q)_\infty} q^{(n-m-1/2)\lambda_i + \lambda_i^2/2},$$

where $(a_1, \dots, a_r; q)_\infty = \prod_{k=1}^{r} (a_k; q)_\infty$. The obvious parallel between (2.6) and (2.3) is one of the main reasons we view $H(\ell_i, \ell_j)$ as the correct integrable generalization to $\theta > 0$.

2.2. Scaling and regularity assumption

We are interested in obtaining asymptotic statements about $\mathbb{P}_N$ as $N \to \infty$. This requires that we scale our parameters θ, q, u, M in some way and also impose some regularity conditions for the interval endpoints $\alpha_i(N), \beta_i(N)$ and the weight functions $w(x; N)$. We list these assumptions below.

Assumption 1. We assume that we are given parameters $\theta > 0$, $\mathtt{q} \in (0,1)$, $\mathtt{M} \geq 1$, and $\mathtt{u} \in [0,1)$. For future reference we denote the set of parameters $\mathtt{q}, \mathtt{M}, \mathtt{u}$ that satisfy the latter conditions by $\mathtt{P}$ and view it as a subset of $\mathbb{R}^3$ with the subspace topology. In addition, we assume that we have a sequence of parameters $q_N \in (0,1)$, $M_N \in \mathbb{N}$ and $u_N \in [0,1)$ such that

$$M_N \geq N - 1, \ \max\left(N^2 \cdot \left|q_N - \mathtt{q}^{1/N}\right|, |M_N - N\mathtt{M}|, N \cdot |u_N - \mathtt{u}|\right) \leq A_1, \quad \text{for some } A_1 > 0. \tag{2.7}$$

The measures $\mathbb{P}_N$ will then be as in (2.3) for $q = q_N, u = u_N, M = M_N, \theta$ and N.

Assumption 2. We require that for each $i = 1, \ldots, k$ as $N \to \infty$ we have for some $A_2 > 0$

$$|\alpha_i(N) - \hat{a}_i| \leq A_2 \cdot N^{-1} \log(N), \qquad |\beta_i(N) - \hat{b}_i| \leq A_2 \cdot N^{-1} \log(N), \text{ where}$$
$$1 + \mathtt{u} \leq \hat{a}_1 < \hat{b}_1 < \hat{a}_2 < \cdots < \hat{a}_k < \hat{b}_k \leq \mathtt{q}^{-\mathtt{M}-\theta+1} + \mathtt{u}\mathtt{q}^{\mathtt{M}+\theta-1}.$$

In addition, we assume that $w(x; N)$ in the intervals $[\alpha_i(N), \beta_i^-(N)]$, $i = 1, \ldots, k$ has the form

$$w(x; N) = \exp\left(-NV_N(x)\right),$$

for a function V_N that is continuous in the intervals $[\alpha_i(N), \beta_i^-(N)]$ and such that

$$|V_N(s) - V(s)| \leq A_3 \cdot N^{-1} \log(N), \text{ where } V \text{ is continuous and } |V(s)| \leq A_4, \tag{2.8}$$

for some constants $A_3, A_4 > 0$. We also require that $V(s)$ is differentiable and for some $A_5 > 0$

$$|V'(s)| \leq A_5 \cdot \left[1 + \sum_{i=1}^{k} |\log|s - \hat{a}_i|| + |\log|s - \hat{b}_i||\right], \text{ for } s \in \left[1 + \mathtt{u}, \mathtt{q}^{-\mathtt{M}-\theta+1} + \mathtt{u}\mathtt{q}^{\mathtt{M}+\theta-1}\right]. \tag{2.9}$$

Remark 2.2.1. We believe that one can take more general remainders in the above two assumptions, without significantly influencing the arguments in the later parts of the paper. However, we do not pursue this direction due to the lack of natural examples.

Let us denote $\sigma_{\mathtt{q}}(x) = \mathtt{q}^{-x} + \mathtt{u}\mathtt{q}^{x}$ and observe that the latter is a bijective diffeomorphism from $[0, \mathtt{M}+\theta-1]$ to $[1+\mathtt{u}, \mathtt{q}^{-\mathtt{M}-\theta+1} + \mathtt{u}\mathtt{q}^{\mathtt{M}+\theta-1}]$. Let $f_{\mathtt{q}}(x) = \frac{d}{dx}\sigma_{\mathtt{q}}(\mathtt{q}^{-x})$ and note that $f_{\mathtt{q}}$ is positive on the interval $[0, \mathtt{M}]$.

Assumption 3. Set $\hat{n}_i(N) = \frac{n_i(N)}{N}$ for $i = 1, \dots, k$. We will often suppress the dependence of $\hat{n}_i$ on N and we assume that for sufficiently large N these sequences satisfy

$$A_6 < \hat{n}_i < \theta^{-1} \cdot \left[\sigma_{\mathtt{q}}^{-1}(\hat{b}_i) - \sigma_{\mathtt{q}}^{-1}(\hat{a}_i)\right] - A_6,$$

where A_6 is some positive constant. In our future results it will be important that the remainders are uniform over $\hat{n}_i$, satisfying the above conditions.

Remark 2.2.2. The above assumptions will be sufficient to obtain our law of large numbers for $\mathbb{P}_N$. We stated the one-cut $\theta = 1$ case of this law in Theorem 1.2.1. In general, if one assumes that $\hat{n}_i(N) = \nu_i + O(N^{-1}\log(N))$ for some positive constants ν_i for $i = 1, \dots, k$, then the sequence of empirical measures $\mu_N := \frac{1}{N}\sum_{i=1}^{N}\delta(\ell_i)$ converges to a deterministic measure μ, called the *equilibrium measure*. The precise statement detailing this convergence is given in Theorem 3.1.1, and the equilibrium measure turns out to be the maximizer of a certain variational problem – see Lemma 3.1.2. It depends on $\mathtt{q}, \mathtt{u}, \theta$, the limiting potential V, the endpoints $\hat{a}_i, \hat{b}_i$ from Assumption 2 and the limiting filling fractions ν_i for $i = 1, \dots, k$.

We next isolate the assumptions we require for establishing our central limit theorem, starting with the analytic properties of the weight $w(x; N)$.

Assumption 4. We assume that we have an open set $\mathcal{M} \subset \mathbb{C} \setminus \{0, \pm\sqrt{\mathtt{u}}\}$, such that for large N

$$\bigcup_{i=1}^{k}\Big(\Big[q_N^{1-a_i(N)-N_{i-1}\cdot\theta}, q_N^{-b_i(N)-(N_i-1)\cdot\theta}\Big] \cup \Big[u_N q_N^{b_i(N)+(N_i-1)\cdot\theta}, u_N q_N^{a_i(N)-1+N_{i-1}\cdot\theta}\Big]\Big) \subset \mathcal{M}.$$

In addition, we require for all large N the existence of analytic functions Φ_N^+, Φ_N^- on $\mathcal{M}$ such that

$$\frac{w(\sigma_N(z); N)}{w(\sigma_N(q_N z); N)} = \frac{(q_N^2 z^2 - u_N q_N^{\theta})(z^2 - u_N)\Phi_N^+(z)}{(q_N^{\theta} z^2 - u_N)(q_N^2 z^2 - u_N)\Phi_N^-(z)}, \tag{2.10}$$

whenever $\sigma_N(z), \sigma(q_N z) \in \cup_{i=1}^{k}[\alpha_i(N), \beta_i^-(N)]$ where $\sigma_N(z) = z + u_N z^{-1}$. Moreover,

$$\Phi_N^-(z) = \Phi^-(z) + \varphi_N^-(z) + O\left(N^{-2}\right) \text{ and } \Phi_N^+(z) = \Phi^+(z) + \varphi_N^+(z) + O\left(N^{-2}\right),$$

where $\varphi_N^{\pm}(z) = O(N^{-1})$ and the constants in the big O notation are uniform over z in compact subsets of $\mathcal{M}$. All aforementioned functions are holomorphic in $\mathcal{M}$.

The next assumption we require is about the equilibrium measure μ, which was discussed in Remark 2.2.2. A convenient way to encode μ is through its Stieltjes transform

$$G_\mu(z) := \int_{\mathbb{R}} \frac{\mu(x)dx}{z-x}. \tag{2.11}$$

The following two functions $R_\mu(z), Q_\mu(z)$ naturally arise from our discrete loop equations (see Section 4.2) and play an important role in our further analysis

$$\begin{aligned} R_\mu(z) &= \Phi^-(z)\cdot e^{\theta\log(\mathtt{q})(z-\mathtt{u}z^{-1})G_\mu(z+\mathtt{u}z^{-1})} + \Phi^+(z)\cdot e^{-\theta\log(\mathtt{q})(z-\mathtt{u}z^{-1})G_\mu(z+\mathtt{u}z^{-1})} \\ Q_\mu(z) &= \Phi^-(z)\cdot e^{\theta\log(\mathtt{q})(z-\mathtt{u}z^{-1})G_\mu(z+\mathtt{u}z^{-1})} - \Phi^+(z)\cdot e^{-\theta\log(\mathtt{q})(z-\mathtt{u}z^{-1})G_\mu(z+\mathtt{u}z^{-1})} \end{aligned} \tag{2.12}$$

In Section 4.2 we show that $R_\mu(z)$ is analytic, while $Q_\mu(z)$ is a branch of a two-valued analytic function, given by the square root of a holomorphic function on $\mathcal{M}$. Our assumption on μ is expressed through Q_μ as follows.

Assumption 5. For each N let $\hat{\mu}_N$ be the equilibrium measure μ discussed in Remark 2.2.2 for the parameters $\mathtt{q}$, $\mathtt{u}$, endpoints $\hat{a}_i, \hat{b}_i$ as in Assumptions 1–2 and filling fractions $\nu_i = \hat{n}_i = n_i(N)/N$, $i = 1, \dots, k$. Observe that $\hat{\mu}_N$ depends on N only through the filling fractions, in particular in the one-cut case it does not depend on N.

Let $Q_{\hat{\mu}_N}$ be as in (2.12) for the measure $\hat{\mu}_N$. Then we require that for all large N there exist real numbers $r_i(N), s_i(N)$ and functions $H_N(z)$ on $\mathcal{M}$ such that

- $\hat{a}_i \le r_i(N) < s_i(N) \le \hat{b}_i$, and there are constants $\hat{a}_i \le \hat{r}_i < \hat{s}_i \le \hat{b}_i$ such that $r_i(N) - \hat{r}_i = O(N^{-1}\log(N)) = s_i(N) - \hat{s}_i$ for $i = 1, \dots, k$.
- $Q_{\hat{\mu}_N}(z) = H_N(z)\prod_{i=1}^k \sqrt{(z+\mathtt{u}z^{-1} - r_i(N))(z+\mathtt{u}z^{-1} - s_i(N))}$, with $H_N(z) \neq 0$ in $\mathcal{M}$.

Remark 2.2.3. Assumption 5 is the analogue of Assumption 4 in [15, Section 3] for our setting and as discussed there it does not describe a general case. In particular, it implies that $\mu(x)$ has a single interval of support inside each interval $[\hat{a}_i, \hat{b}_i]$. To the authors' knowledge there are no simple conditions on the potential V that ensure that μ has this property.

We will further impose a vanishing condition for the functions $\Phi_N^{\pm}$. We believe that it can be relaxed, but introduce it to simplify our arguments in the text.

Assumption 6. If $a_i(N), b_i(N), N_i$ are as in Section 2.1 then for all $i = 1, \dots, k$ we have

$$\Phi_N^+\left(q_N^{-b_i-(N_i-1)\theta}\right) = \Phi_N^-\left(q_N^{-a_i-N_{i-1}\cdot\theta}\right) = \Phi_N^+\left(u_N q_N^{a_i-1+N_{i-1}\cdot\theta}\right)$$

$$= \Phi_N^- \left(u_N q_N^{b_i - 1 + (N_i - 1)\cdot\theta} \right) = 0.$$

Finally, we state an assumption, under which one can find an explicit formula for the density of μ in Remark 2.2.2 in terms of R_μ and $\Phi^\pm$ as in (2.12) and Assumption 4, see Lemma 4.2.2.

Assumption 7. We assume that $V(s)$ is real analytic in a real open neighborhood of $\cup_{i=1}^k [\hat{a}_i, \hat{b}_i]$ and $\Phi^+(x), \Phi^-(x) \in \mathbb{R}$ with $\Phi^+(x) \cdot \Phi^-(x) > 0$ whenever $x + \mathfrak{u} x^{-1} \in \cup_{i=1}^k (\hat{a}_i, \hat{b}_i)$.

3. Law of large numbers

In this section we establish the law of large numbers for the empirical measures μ_N associated to $\mathbb{P}_N$ from Section 2. In Section 3.1 we provide a variational formulation of the equilibrium measure μ, which describes the limit of μ_N. The convergence of μ_N to μ is detailed in Theorem 3.1.1 and we reduce the proof of the latter to a concentration inequality – see Proposition 3.1.3. This inequality is established in Section 3.2 using arguments similar to [15], which in turn go back to [17] and [52].

3.1. Convergence of empirical measures

We continue with the same notation as in Section 2. With $\mathbb{P}_N$ as in (2.3) we define the empirical measures μ_N as

$$\mu_N = \frac{1}{N} \sum_{i=1}^{N} \delta(\ell_i) \text{ where } (\ell_1, \ldots, \ell_N) \text{ is } \mathbb{P}_N\text{-distributed.}$$

The measures μ_N satisfy the following law of large numbers.

Theorem 3.1.1. *Suppose that Assumptions 1, 2 and 3 from Section 2.2 hold. In addition, suppose that* $|\hat{n}_i - \nu_i| \leq A_7 \cdot N^{-1} \log(N)$ *for some positive constants* A_7 *and* ν_i, $i = 1, \ldots, k$ *such that* $\sum_i \nu_i = 1$. *Then there is a deterministic probability measure* $\mu(x)dx$, *depending on* ν_i *for* $1 \leq i \leq k$, *such that* μ_N *concentrate (in probability) near* μ. *More precisely, for each Lipschitz function* $f(x)$ *defined in a real neighborhood of the interval* $\cup_{i=1}^k [\hat{a}_i, \hat{b}_i]$ *and each* $\varepsilon > 0$ *the random variables*

$$N^{1/2-\varepsilon} \left| \int_{\mathbb{R}} f(x) \mu_N(dx) - \int_{\mathbb{R}} f(x) \mu(x) dx \right| \tag{3.1}$$

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

The limiting measure μ is defined as the maximizer of a certain variational problem, described in the following section.

3.1.1. Variational problem

Define the functional $I_V[\rho]$ of a measure ρ supported in $\cup_{i=1}^{k}[\hat{a}_i, \hat{b}_i]$ via

$$I_V[\rho] = \theta \cdot \iint\limits_{x \neq y} \log|x-y| \, d\rho(x) d\rho(y) - \int\limits_{\mathbb{R}} V(x) d\rho(x). \tag{3.2}$$

Lemma 3.1.2. *Let Θ denote the set of absolutely continuous probability measures $\rho(x)dx$ supported on $\cup_{i=1}^{k}[\hat{a}_i, \hat{b}_i]$, whose denisty $\rho(x)$ is between 0 and $\theta^{-1} \cdot f_{\mathtt{q}}(\sigma_{\mathtt{q}}^{-1}(x))^{-1}$ and such that*

$$\int\limits_{\hat{a}_i}^{\hat{b}_i} \rho(x)dx = \hat{\nu}_i, \text{ with } 1 \leq i \leq k,$$

where $0 < \hat{\nu}_i < \theta^{-1} \cdot \left[\sigma_{\mathtt{q}}^{-1}(\hat{b}_i) - \sigma_{\mathtt{q}}^{-1}(\hat{a}_i)\right]$, $i = 1, \ldots, k$ are such that $\sum_{i=1}^{k} \hat{\nu}_i = 1$ (recall that $\sigma_{\mathtt{q}}$ and $f_{\mathtt{q}}$ were defined in Section 2.2). Then the functional I_V has a unique maximum $\hat{\mu}$ on Θ.

Proof. Observe that by our assumption on $\mathtt{u}$ and $\mathtt{q}$ we know that $\sigma_{\mathtt{q}}(y)$ is a strictly increasing function, whose derivative $f_{\mathtt{q}}(y)$ on $[0, \mathtt{M}]$ lies between $(-\log \mathtt{q})(1-\mathtt{u})$ and $(-\log \mathtt{q})\mathtt{q}^{-\mathtt{M}-\theta+1}$.

Let Θ' be the same as Θ, except that we restrict $0 \leq \rho(x) \leq \frac{1}{\theta \cdot (-\log \mathtt{q}) \cdot (1-\mathtt{u})}$. From the above argument we conclude that Θ is a closed convex subset of Θ'. It follows from the proof of Lemma 5.1 in [15] that $I_V[\rho]$ is a continuous strictly concave functional on Θ' and that the latter is compact. It follows that Θ is convex and compact, and hence $I_V[\rho]$ attains a unique maximum there. □

We call the measure $\hat{\mu}$ from Lemma 3.1.2 the *equilibrium measure*. The significance of $\hat{\mu}$ is that it equals μ from Theorem 3.1.1, when $\hat{\nu}_i = \nu_i$ for $1 \leq i \leq k$. Proving this fact will be the focus of this and the subsequent sections.

For a measure $\rho \in \Theta$ as in Lemma 3.1.2 define the effective potential $F_\rho^V(x)$ through

$$F_\rho^V(x) = 2\theta \cdot \int\limits_{\mathbb{R}} \log|x-t| \rho(t)dt - V(x).$$

Applying Theorem 2.1 in [28] to each interval $[\hat{a}_i, \hat{b}_i]$ we know that there exist real numbers f_i for $1 \leq i \leq k$ such that

$$\begin{cases} F_{\hat{\mu}}^V(x) - f_i \leq 0 & \text{for a.e. } x \in S_i = \{\hat{a}_i \leq x \leq \hat{b}_i | 0 \leq \hat{\mu}(x) < \theta^{-1} \cdot f_{\mathtt{q}}(\sigma_{\mathtt{q}}^{-1}(x))^{-1}\}, \\ F_{\hat{\mu}}^V(x) - f_i \geq 0 & \text{if } x \in [\hat{a}_i, \hat{b}_i] \cap \operatorname{supp}(\hat{\mu}). \end{cases} \tag{3.3}$$

3.1.2. Proof of Theorem 3.1.1

Our approach to proving Theorem 3.1.1 is to reconstruct in our setup the arguments of Section 5 in [15] and we begin by introducing some relevant notation. Take any two compactly supported absolutely continuous probability measures with uniformly bounded densities $\nu(x)dx$ and $\rho(x)dx$ and define $\mathcal{D}(\nu(x),\rho(x))$ through

$$\mathcal{D}^2(\nu(x),\rho(x)) = -\int_{\mathbb{R}}\int_{\mathbb{R}} \log|x-y|(\nu(x)-\rho(x))(\nu(y)-\rho(y))dxdy. \tag{3.4}$$

There is an alternative formula for $\mathcal{D}(\nu(x),\rho(x))$ in terms of Fourier transforms, cf. [6]:

$$\mathcal{D}(\nu(x),\rho(x)) = \sqrt{\int_0^\infty \frac{dt}{t}\left|\int_{\mathbb{R}} e^{-itx}(\nu(x)-\rho(x))dx\right|^2}. \tag{3.5}$$

Fix a parameter $p>2$ and let $\tilde{\mu}_N$ denote the convolution of the empirical measure μ_N with the uniform measure on the interval $[0,N^{-p}]$. With the above notation we can now state the main technical result we require for proving Theorem 3.1.1.

Proposition 3.1.3. *Assume the same notation as in Theorem 3.1.1 and let $\hat{\mu}$ be as in Lemma 3.1.2 for $\hat{\nu}_i=\nu_i$ for $i=1,\dots,k$. There exists a constant $C>0$ such that for all $x>0$ and $N\geq 2$*

$$\mathbb{P}_N\left(\mathcal{D}(\tilde{\mu}_N,\hat{\mu})\geq x\right)\leq \exp\left(CN\log(N)^2-\theta\cdot x^2N^2\right).$$

The constant C depends on the constants $A_1,\dots,A_7$ in Theorem 3.1.1 and Assumptions 1, 2, 3, as well as $\mathtt{q},\mathtt{M},\mathtt{u},\theta$ and is uniform as the latter vary over compact subsets of $\mathtt{P}$.

The proof of Proposition 3.1.3 is the focus of Section 3.2 below. For now we assume its validity and conclude the proof of Theorem 3.1.1. We start by deducing the following corollary.

Corollary 3.1.4. *Assume the same notation as in Proposition 3.1.3. For a compactly supported Lipschitz function $g(x)$ define*

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

Fix any $p>2$. Then for all $a>0$, $N\geq 2$ and g we have

$$\mathbb{P}_N\left(\left|\int_{\mathbb{R}} g(x)\mu_N(dx) - \int_{\mathbb{R}} g(x)\hat{\mu}(dx)\right| \geq a\|g\|_{1/2} + \frac{\|g\|_{\text{Lip}}}{N^p}\right)$$
$$\leq \exp\left(CN\log(N)^2 - 2\theta\pi^2 a^2 N^2\right), \tag{3.6}$$

where the constant C is as in Proposition 3.1.3.

Proof. From the triangle inequality we have

$$\left|\int_{\mathbb{R}} g(x)\mu_N(dx) - \int_{\mathbb{R}} g(x)\mu(dx)\right| \leq \left|\int_{\mathbb{R}} g(x)\mu_N(dx) - \int_{\mathbb{R}} g(x)\tilde{\mu}_N(dx)\right|$$
$$+ \left|\int_{\mathbb{R}} g(x)[\tilde{\mu}_N(x) - \mu(x)]dx\right|.$$

The first term is bounded by $\frac{\|g\|_{\text{Lip}}}{N^p}$ and corresponds to such a term in (3.34). We will thus focus on estimating the second term.

We denote by $\mathcal{F}[\phi](\xi) := \int_{\mathbb{R}} e^{-ix\xi}\phi(x)dx$ the Fourier transform of a function ϕ. Note that $g, \hat{\mu}(x), \tilde{\mu}_N(x)$ all belong to $L^2(\mathbb{R}) \cap L^1(\mathbb{R})$ and so we can use Parseval's formula (see e.g. Theorem 7.1.6, in [35]) and the Cauchy–Schwarz inequality to get

$$\left|\int_{\mathbb{R}} g(x)\tilde{\mu}_N(x)dx - \int_{\mathbb{R}} g(x)\mu(x)dx\right| =$$
$$(2\pi)^{-1}\left|\int_{\mathbb{R}} \left(\sqrt{|\xi|}\mathcal{F}[g](\xi)\right)\cdot\frac{\mathcal{F}[\tilde{\mu}_N](\xi) - \mathcal{F}[\hat{\mu}](\xi)}{\sqrt{|\xi|}}d\xi\right| \leq$$
$$(2\pi)^{-1}\|g\|_{1/2}\cdot\sqrt{\int_{\mathbb{R}}\frac{|\mathcal{F}[\tilde{\mu}_N](\xi) - \mathcal{F}[\hat{\mu}](\xi)|^2}{|\xi|}d\xi} = (2\pi)^{-1}\|g\|_{1/2}\cdot\sqrt{2}\mathcal{D}(\tilde{\mu}_N, \mu).$$

In the last equality above we used (3.5). What remains is to use Proposition 3.1.3. □

Proof of Theorem 3.1.1. Suppose that f and ε are as in the statement of the theorem, $\mu = \hat{\mu}$ from Proposition 3.1.3 and assume without loss of generality that $\varepsilon \in (0, 1/2)$. Fix $\eta > 0$ and let h be a smooth function, whose support is contained in a η-neighborhood of $\cup_{i=1}^k [\hat{a}_i, \hat{b}_i]$ and $h(x) = 1$ on a $\eta/2$-neighborhood of $\cup_{i=1}^k [\hat{a}_i, \hat{b}_i]$. If we set

$$X_N := N^{1/2-\varepsilon}\left|\int_{\mathbb{R}} f(x)\cdot h(x)\mu_N(dx) - \int_{\mathbb{R}} f(x)\cdot h(x)\mu(x)dx\right|,$$

then to prove the theorem we need to show that for each $k \geq 1$ we have $\lim_{N\to\infty} \mathbb{E}\left[X_N^k\right] = 0$.

It follows from Corollary 3.1.4 that there exist positive constants c_1, c_2 and C such that for all $a > 0$ and $N \geq 2$ we have

$$\mathbb{P}\left(X_N \geq c_1 \cdot a \cdot N^{1/2-\varepsilon} + c_2 \cdot N^{-3}\right) \leq \exp(CN\log(N)^2 - 2\theta\pi^2 a^2 N^2).$$

Using the above inequality and setting $a_N = 2c_2 \cdot N^{-\varepsilon}$ we see that for any $k \geq 1$ we have

$$\mathbb{E}\left[X_N^k\right] = \int_0^\infty x^{k-1} \cdot \mathbb{P}(X_N \geq x)dx = \int_0^{a_N} x^{k-1} \cdot \mathbb{P}(X_N \geq x)dx + \int_{a_N}^\infty x^{k-1} \cdot \mathbb{P}(X_N \geq x)dx \leq$$

$$\leq a_N \cdot \max(a_N^{k-1}, 1) + \exp(CN\log(N)^2) \cdot \int_{a_N}^\infty x^{k-1} \cdot \exp\left(-\theta \cdot \frac{\pi^2 x^2 N^{1+2\varepsilon}}{2c_1^2}\right) dx$$

The last inequality implies that $\lim_{N\to\infty} \mathbb{E}\left[X_N^k\right] = 0$. □

3.2. Proof of Proposition 3.1.3

We begin with a technical result about the asymptotics of the ratio of two q-Gamma functions.

Lemma 3.1. *Suppose that $\theta > 0$, $\mathtt{q} \in (0,1)$ and $q_N \in (0,1), \alpha_N > 0$ are sequences such that $|q_N^N - \mathtt{q}| \leq A \cdot N^{-1}$ and $\alpha_N \leq A$ for some $A > 0$. Then for any $x \geq 1$ we have*

$$\frac{\Gamma_{q_N}(x + \alpha_N + \theta)}{\Gamma_{q_N}(x + \alpha_N)} = (1 - q_N)^{-\theta} \cdot (1 - q_N^x)^\theta \cdot \exp\left[O\left(\frac{N^{-1}}{1 - q_N^x}\right)\right], \tag{3.7}$$

where the constants in the big O notation depend on θ, A and $\mathtt{q}$.

Proof. For convenience we drop the dependence on N from the notation. Recall from (2.4) that

$$\frac{\Gamma_q(x + \alpha + \theta)}{\Gamma_q(x + \alpha)} = (1 - q)^{-\theta} \cdot \frac{(q^{x+\alpha}; q)_\infty}{(q^{x+\alpha+\theta}; q)_\infty}, \text{ where we have } (a; q)_\infty = \prod_{k=0}^\infty (1 - aq^k).$$

The first term matches the corresponding one in (3.7) and we focus on the second term.

We first observe that if $q, B \in (0,1)$ and $s \in [0,1]$ then

$$\frac{(B; q)_\infty}{(Bq^s; q)_\infty} \leq (1 - B)^s. \tag{3.8}$$

The latter is equivalent to showing that

$$f(s) := s\log(1-B) + \sum_{k=0}^{\infty}\log(1-Bq^{s+k}) - \sum_{k=0}^{\infty}\log(1-Bq^{k}) \geq 0 \text{ on } [0,1],$$

which is immediate from the observations: $f(0) = f(1) = 0$ and $f''(s) < 0$ for $s \in (0,1)$.

We next note that

$$\begin{aligned}\frac{(q^{x+\alpha};q)_\infty}{(q^{x+\alpha+\theta};q)_\infty} &= \prod_{i=1}^{\lfloor\theta\rfloor}(1-q^{x+\alpha+i-1})\cdot\frac{(q^{x+\alpha+\lfloor\theta\rfloor};q)_\infty}{(q^{x+\alpha+\theta};q)_\infty}\\ &= \prod_{i=1}^{\lfloor\theta\rfloor+1}(1-q^{x+\alpha+i-1})\cdot\frac{(q^{x+\alpha+\lfloor\theta\rfloor+1};q)_\infty}{(q^{x+\alpha+\theta};q)_\infty}.\end{aligned}$$

Combining the latter with (3.8) we conclude that

$$(1-q^x)^\theta \leq \frac{(q^{x+\alpha};q)_\infty}{(q^{x+\alpha+\theta};q)_\infty} \leq (1-q^{x+\alpha+\theta})^\theta \leq (1-q^x)^\theta\cdot\exp\left[\frac{\theta(1-q^{\alpha+\theta})}{1-q^x}\right],$$

where in the last inequality we used that $q^x \leq 1$ and the trivial inequality $(1+a)^\theta \leq e^{a\theta}$, for $a > 0$. The latter tower of inequalities implies our desired estimate. □

In the remainder of this section we present the proof of Proposition 3.1.3, using appropriately adapted arguments from [15]. For clarity we split it into several steps and we outline them here. In Step 1 we relate the formula for $\mathbb{P}_N$ to the value of the functional I_V from Section 3.1.1 at the empirical measure μ_N. In Steps 2, 3 and 4 we obtain a lower bound for the partition function Z_N. In the fifth step we replace the empirical measure μ_N with its convolution with the uniform measure on $[0, N^{-p}]$ with $p > 2$, and reduce the statement of the proposition to establishing a certain upper bound on $I_V[\mu_N] - I_V[\tilde{\mu_N}]$. In Steps 6,7 and 8 we establish the desired upper bound by employing the variational characterization of $\hat{\mu}$ from Section 3.1.1.

Step 1. We recall for the reader's convenience equation (2.3) below

$$\begin{aligned}&\mathbb{P}_N(\ell_1,\dots,\ell_N) = \frac{1}{Z_N}\cdot\prod_{1\leq i<j\leq N}H(\ell_i,\ell_j)\cdot\prod_{i=1}^{N}w(\ell_i;N), \text{ where}\\ &H(\ell_i,\ell_j) =\\ &q^{-2\theta\lambda_j}\frac{\Gamma_q(\lambda_j-\lambda_i+1)\Gamma_q(\lambda_j-\lambda_i+\theta)}{\Gamma_q(\lambda_j-\lambda_i)\Gamma_q(\lambda_j-\lambda_i+1-\theta)}\frac{\Gamma_q(\lambda_j+\lambda_i+v+1)\Gamma_q(\lambda_j+\lambda_i+v+\theta)}{\Gamma_q(\lambda_j+\lambda_i+v)\Gamma_q(\lambda_j+\lambda_i+v+1-\theta)},\end{aligned} \tag{3.9}$$

where we drop the dependence on N from the notation. The goal of this section is to show

$$\mathbb{P}_N(\ell_1,\dots,\ell_N) = Z_N^{-1}\cdot(1-q)^{-2\theta N(N-1)}\cdot\exp\left(N^2 I_V\left[\mu_N\right]+O(N\log(N))\right). \quad (3.10)$$

Notice that the definition of I_V in (3.2) makes sense for discrete (atomic) measures – here it is important that we integrate over $x \neq y$ as otherwise the integral would be infinite, and so the RHS of (3.10) is well-defined and finite.

From Assumption 2 in Section 2.2, we know that $w(\ell_i;N) = \exp(-NV(\ell_i) + O(\log N))$, and to conclude the proof of (3.10) what remains is to show that

$$\prod_{1\le i<j\le N} H(\ell_i,\ell_j) = (1-q)^{-2\theta N(N-1)} \prod_{1\le i<j\le N} (\ell_j-\ell_i)^{2\theta}\cdot\exp\left[O(N\log(N))\right]. \quad (3.11)$$

From Lemma 3.1 we know that

$$\prod_{1\le i<j\le N} H(\ell_i,\ell_j) = (1-q)^{-2\theta N(N-1)} \prod_{1\le i<j\le N} (\ell_j-\ell_i)^{2\theta}\cdot\exp\left[O\left(\frac{N^{-1}}{1-q^{\theta(j-i)}}\right)\right], \quad (3.12)$$

where we used that $\lambda_j-\lambda_i \ge \theta\cdot(j-i)$ by assumption. On the other hand,

$$\sum_{1\le i<j\le N}\frac{N^{-1}}{1-q^{\theta(j-i)}} \le \frac{1}{1-q^{\theta}}\cdot\sum_{i=1}^{N}\frac{1-q^{\theta}}{1-q^{i\theta}} \le \frac{1}{1-q^{\theta}}\cdot\sum_{i=1}^{N}\frac{1}{i\cdot c_0} \le C\cdot N\log(N), \quad (3.13)$$

where $c_0>0$ is a universal lower bound of q^N. Equations (3.12) and (3.13) imply (3.11).

Step 2. The goal of this and the next two steps is to obtain the following lower bound

$$Z_N \ge (1-q)^{-2\theta N(N-1)}\cdot\exp\left(N^2 I_V\left[\hat{\mu}\right]+O(N\log(N)^2)\right). \quad (3.14)$$

In this step we construct a particular element $\hat{\ell}=(\hat{\ell}_1,\dots,\hat{\ell}_N)\in\mathfrak{X}_N^{\theta}$ that depends on $\hat{\mu}$, and then in view of (3.10) we have the immediate lower bound

$$Z_N \ge (1-q)^{-2\theta N(N-1)}\cdot\exp\left(N^2 I_V\left[\mathrm{mes}\left[\hat{\ell}_1,\cdots,\hat{\ell}_N\right]\right]+O(N\log(N))\right), \quad (3.15)$$

where $\mathrm{mes}\left[\hat{\ell}_1,\cdots,\hat{\ell}_N\right] = \frac{1}{N}\sum_{i=1}^{N}\delta(\hat{\ell}_i)$.

Let y_i, $i=1,\dots,N$ be quantiles of $\sigma_{\mathfrak{q}}^{-1}\circ\hat{\mu}$ defined through

$$\int_0^{y_i}\left[\sigma_{\mathfrak{q}}^{-1}\circ\hat{\mu}\right](x)dx = \frac{i-1/2}{N}, \quad 1\le i\le N.$$

Since $\left[\sigma_{\mathfrak{q}}^{-1}\circ\hat{\mu}\right](x)\le\theta^{-1}$ we have $\theta(y_{i+1}-y_i)\ge N^{-1}$. Arguing as in the proof of [15, Proposition 5.6] we can find an element $\hat{\lambda}=(\hat{\lambda}_1,\dots,\hat{\lambda}_N)\in\mathbb{W}_N^{\theta}$ such that:

(1) if $\hat{\lambda}_i=\hat{x}_i+(i-1)\cdot\theta$, then $a_i(N)\le\hat{x}_i\le b_i(N)$;

(2) there is a constant U (independent of N) such that $|Ny_i - \hat{\lambda}_i| \leq U$ for all $1 \leq i \leq N$ except for $O(\log(N))$ ones.

We then define $\hat{\ell}$ through $\hat{\ell}_i = q^{-\hat{\lambda}_i} + uq^{\hat{\lambda}_i}$ and note that the first condition above ensures $\hat{\ell} \in \mathfrak{X}_N^\theta$.

Step 3. The goal of this step is to show that

$$N^2 I_V \left[\text{mes}\left[\hat{\ell}_1, \cdots, \hat{\ell}_N\right]\right] = N^2 I_V\left[\hat{\mu}\right] + O(N\log(N)^2). \tag{3.16}$$

Clearly, (3.16) and (3.15) give (3.14).

Setting $\sigma_N(x) = q^{-x} + uq^x$ we see that to show (3.16) it suffices to have the following equalities

$$\sum_{1\leq i<j\leq N} \log\left(\sigma_N(\hat{\lambda}_j) - \sigma_N(\hat{\lambda}_i)\right) = N^2 \iint\limits_{s<t} \log(t-s)\hat{\mu}(t)\hat{\mu}(s)dtds + O(N\log(N)^2), \tag{3.17}$$

$$N\sum_{i=1}^{N} V(\sigma_N(\hat{\lambda}_i)) = N^2 \int\limits_{\mathbb{R}} V(t)\hat{\mu}(t)dt + O(N\log(N)). \tag{3.18}$$

We defer the proof of (3.17) to Step 4 and focus on showing (3.18).

Set $z_i = \sigma_{\mathsf{q}}(y_i)$ for $i = 1, \ldots, N$ and observe that $\int_{z_i}^{z_{i+1}} \hat{\mu}(t)dt = N^{-1}$ for $i = 1, \ldots, N-1$. Then

$$\sum_{i=1}^{N} V\left(\sigma_N(\hat{\lambda}_i)\right) = N\sum_{i=1}^{N-1} \int\limits_{z_i}^{z_{i+1}} V\left(\sigma_N(\hat{\lambda}_i)\right)\hat{\mu}(t)dt + O(1). \tag{3.19}$$

Let I be the set of indices such that $\sigma_N(\hat{\lambda}_i), z_i, z_{i+1}$ for $i \in I$ are all inside $\cup_{j=1}^{k}[\hat{a}_j, \hat{b}_j]$ and at least N^{-1} away from the complement of this set, and such that $|Ny_i - \hat{\lambda}_i| \leq U$ from Step 2 holds. From Assumption 2 on $\hat{a}_j, \hat{b}_j$ we conclude that $N - |I| = O(\log N)$. Note that for $i \in I$

- $z_i = \sigma_N(\hat{\lambda}_i) + O(N^{-1})$;
- $V(z_i) - V\left(\sigma_N(\hat{\lambda}_i)\right) = (z_i - \sigma_N(\hat{\lambda}_i))V'(s) = O\left(N^{-1}\log(N)\right)$, where we used the mean value theorem and $V'(s) = O(\log(N))$ from Assumption 2.

In view of the above (3.19) implies

$$\sum_{i=1}^{N} V\left(\sigma_N(\hat{\lambda}_i)\right) = N\sum_{i\in I} \int\limits_{z_i}^{z_{i+1}} V(z_i)\hat{\mu}(t)dt + O(\log(N)). \tag{3.20}$$

A second application of the mean value theorem leads to

$$\sum_{i=1}^{N} V\left(\sigma_N(\hat{\lambda}_i)\right) = N \sum_{i\in I} \int_{z_i}^{z_{i+1}} \left[V(t) + (t-z_i)\, V'(\kappa(t))\right] \hat{\mu}(t)dt + O(\log(N)),$$

where $\kappa(t)$ is a point inside $\cup_{j=1}^{k}[\hat{a}_j, \hat{b}_j]$ at least N^{-1} from the complement of this set. Arguing as before that $V'(\kappa(t)) = O(\log(N))$, we see that

$$\sum_{i=1}^{N} V\left(\sigma_N(\hat{\lambda}_i)\right) = N \sum_{i\in I} \int_{z_i}^{z_{i+1}} \left[V(t) + (z_{i+1}-z_i)\, O(\log(N))\right] \hat{\mu}(t)dt + O(\log(N)) =$$

$$N \sum_{i\in I} \int_{z_i}^{z_{i+1}} V(t)\hat{\mu}(t)dt + N\frac{O(\log(N))}{N} \sum_{i=2}^{N-2} (z_{i+1}-z_i) = N \int_{\mathbb{R}} V(t)\hat{\mu}(t)dt + O(\log(N)). \tag{3.21}$$

The second equality above follows from the definition of z_i as quantiles of $\hat{\mu}$, and the last one follows from the fact that $\hat{\mu}\left(\cup_{i\in I}[z_i, z_{i+1}]\right) = |I|/N = 1 - O(N^{-1}\log(N))$ and $V = O(1)$ on the support of $\hat{\mu}$. Clearly, (3.21) implies (3.18).

Step 4. In this step we prove (3.17) and start by showing that

$$\sum_{1\le i<j\le N} \log\left(\sigma_N(\hat{\lambda}_j) - \sigma_N(\hat{\lambda}_i)\right) \le N^2 \iint_{s<t} \log(t-s)\hat{\mu}(t)\hat{\mu}(s)dtds + O(N\log(N)^2). \tag{3.22}$$

If I is as in Step 3 then we observe that

$$\sum_{1\le i<j\le N} \log\left(\sigma_N(\hat{\lambda}_j) - \sigma_N(\hat{\lambda}_i)\right) =$$

$$\sum_{i<j;i,j\in I} \log\left(\sigma_N(\hat{\lambda}_j) - \sigma_N(\hat{\lambda}_i)\right) + O(N\log(N)^2). \tag{3.23}$$

Indeed, the two sums differ by $O(N\log(N))$ summands, each of order $O(\log(N))$.

As discussed in Step 3 we have that $\sigma_N(\hat{\lambda}_i) = z_i + O(N^{-1})$ for $i \in I$. It follows, that we can find a positive constant C such that

$$\sum_{i<j;i,j\in I} \log\left(\sigma_N(\hat{\lambda}_j) - \sigma_N(\hat{\lambda}_i)\right) \le \sum_{i<j;i,j\in I} \log\left(z_j - z_i + CN^{-1}\right) \le$$

$$N^2 \sum_{i<j;i,j\in I} \int_{z_{i-1}}^{z_i} \int_{z_j}^{z_{j+1}} \log\left(t-s+CN^{-1}\right) \hat{\mu}(t)\hat{\mu}(s)dtds = \tag{3.24}$$

$$N^2 \iint\limits_{s<t} \log\left(t-s+CN^{-1}\right)\hat{\mu}(t)\hat{\mu}(s)dtds + O(N\log(N)^2).$$

In going from the first to the second line we used that z_i are quantiles of $\hat{\mu}$ and the monotonicity of log. In going from the second to the third, we note that the set difference over which the two integrals are taken has $\hat{\mu}\times\hat{\mu}$ measure $O(N^{-1}\log N)$ and the integrand is $O(\log(N))$ there.

We see that to conclude (3.22) it suffices to show that

$$\sum_{j=1}^{k} \iint\limits_{\hat{a}_j\le v<w\le\hat{b}_j} \log\left(\frac{w-v+CN^{-1}}{w-v}\right)dwdv = O\left(\frac{\log(N)}{N}\right).$$

The above is now immediate from the observation that for $c\ge 0$ we have

$$\int_a^b\int_v^b \log\left(w-v+c\right)dwdv = \int_a^b [(b-v+c)(\log(b-v+c)-1)-c\log(c)+c]dv =$$

$$c[1-\log(c)]r + \frac{c^2[1-2\log(c)]+(r+c)^2[2\log(r+c)-1]}{4} - br + r^2/2 + cr,$$

where $r=b-a$. The above identities show (3.22) and the reverse inequality can be established in an analogous way, which proves (3.17).

Step 5. In this step we show that we can replace μ_N in (3.10) with its convolution with the uniform measure on $[0,N^{-p}]$ with $p>2$, denoted by $\tilde{\mu}_N$. For that we extend V to $\cup_{j=1}^k[\hat{a}_j,\hat{b}_j+N^{-p}]$ by setting $V(x+\hat{b}_j)=V(\hat{b}_j)$ for $x\in[0,N^{-p}]$ and take two independent random variables $u,\tilde{u}$ uniformly distributed on $[0,N^{-p}]$. Then we have

$$\begin{aligned} I_V[\tilde{\mu}_N] &= \mathbb{E}_{u,\tilde{u}} \iint \log|x+u-y-\tilde{u}|\,\mu_N(dx)\mu_N(dy) - \mathbb{E}_u\int V(x+u)\mu_N(dx) = \\ I_V[\mu_N] &+ \frac{1}{N}\cdot\mathbb{E}_{u,\tilde{u}}\int\log|u-\tilde{u}|\mu_N(dx) - \mathbb{E}_u\int[V(x+u)-V(x)]\mu_N(dx) \\ &+ \mathbb{E}_{u,\tilde{u}} \iint\limits_{x\ne y} \log\left|\frac{x+u-y-\tilde{u}}{x-y}\right|\mu_N(dx)\mu_N(dy) = I_V[\mu_N] + O(N^{-1}\log(N)), \end{aligned} \tag{3.25}$$

where the last equality follows from the conditions on V from Assumption 2. The above shows that we can replace μ_N with $\tilde{\mu}_N$ in (3.10) without affecting the statement. Combining the latter with the lower bound of Z_N from (3.14), we conclude that there exists a constant $C'>0$ such that

$$\mathbb{P}_N(\ell_1,\dots,\ell_N) \le \exp\left(C'N\log(N)^2\right)\cdot\exp\left(N^2(I_V(\tilde{\mu}_N)-I_V(\hat{\mu}))\right). \tag{3.26}$$

We next claim that we have the following inequality

$$I_V[\tilde{\mu}_N] - I_V[\hat{\mu}] \leq -\theta \cdot \mathcal{D}^2(\tilde{\mu}_N, \hat{\mu}) + O(N^{-1}\log(N)). \tag{3.27}$$

We defer the proof of the above to the next step. In what follows we assume its validity and finish the proof of the proposition. It follows from (3.26) and (3.27) that for some $C'' > 0$ we have

$$\mathbb{P}_N(\ell_1, \dots, \ell_N) \leq \exp(C'' N \log(N)^2) \exp\left(-\theta \cdot N^2 \mathcal{D}^2(\tilde{\mu}_N, \hat{\mu})\right).$$

Notice that the number of N-tuples $\ell_1 < \cdots < \ell_N$ in $\mathfrak{X}_N^\theta$ is at most $\binom{M_N}{N}$. Since $M_N = O(N)$ we see that for some $C > 0$ and all $x > 0$ we have

$$\begin{aligned}\mathbb{P}_N\left(\mathcal{D}(\tilde{\mu}_N, \hat{\mu}) \geq x\right) &\leq \binom{M_N}{N} \exp\left(C'' N \log(N)^2 - \theta \cdot N^2 x^2\right) \\ &\leq \exp\left(CN\log(N)^2 - \theta \cdot x^2 N^2\right).\end{aligned}$$

This is the desired estimate.

Step 6. In this and the next two steps we establish (3.27). By definition of $\mathcal{D}$ we have

$$\begin{aligned}I_V[\tilde{\mu}_N] - I_V[\hat{\mu}] &= -\theta \cdot \mathcal{D}^2(\tilde{\mu}_N, \hat{\mu}) + \int_{\mathbb{R}} F_{\hat{\mu}}^V(x)\left(\tilde{\mu}_N(x) - \hat{\mu}(x)\right) dx = \\ &- \theta \cdot \mathcal{D}^2(\tilde{\mu}_N, \hat{\mu}) + \sum_{j=1}^{k} \int_{\hat{a}_j}^{\hat{b}_j} [F_{\hat{\mu}}^V(x) - f_i]\left(\tilde{\mu}_N(x) - \hat{\mu}(x)\right) dx + O(N^{-1}\log(N)),\end{aligned} \tag{3.28}$$

where we recall that $F_{\hat{\mu}}^V$ and f_i were defined in Section 3.1.1. The extra $O(N^{-1}\log(N))$ comes from two sources. Firstly, there is additional mass of $\tilde{\mu}_N$ that lies outside of $\cup_{j=1}^k [\hat{a}_j, \hat{b}_j]$ and we are excluding. The second source comes from the fact that the mass of μ_N and $\hat{\mu}$ on each $[\hat{a}_j, \hat{b}_j]$ are not exactly the same (thus the integral over the constant f_i is not zero). The first issue is resolved by Assumption 2 on the endpoints $\alpha_i(N), \beta_i(N)$, which estimates the missing weight as $O(N^{-1}\log(N))$. The second issue is resolved by our assumption that $|\hat{n}_i - \nu_i| = O(N^{-1}\log(N))$.

We recall from Section 3.1.1 that $S_j = \{\hat{a}_j \leq x \leq \hat{b_j} | 0 \leq \hat{\mu}(x) < \theta^{-1} \cdot f_{\mathfrak{q}}(\sigma_{\mathfrak{q}}^{-1}(x))^{-1}\}$ and also set $S_j' = \{\hat{a}_j \leq x \leq \hat{b}_j | \hat{\mu}(x) = 0\}$ and $S_j'' = \{\hat{a}_j \leq x \leq \hat{b}_j | \hat{\mu}(x) = \theta^{-1} \cdot f_{\mathfrak{q}}(\sigma_{\mathfrak{q}}^{-1}(x))^{-1}\}$. In view of (3.28) it suffices to show for each $j = 1, \dots, k$ that

$$\int_{S_j \setminus S_j'} [F_{\hat{\mu}}^V(x) - f_j]\left(\tilde{\mu}_N(x) - \hat{\mu}(x)\right) dx = 0, \quad \int_{S_j'} [F_{\hat{\mu}}^V(x) - f_j]\left(\tilde{\mu}_N(x) - \hat{\mu}(x)\right) dx \leq 0,$$

$$\int_{S''_j} [F_{\hat{\mu}}(x) - f_j] \left(\tilde{\mu}_N(x) - \hat{\mu}(x)\right) dx \leq 0 + O(N^{-1}\log(N)). \tag{3.29}$$

In what follows we fix $j \in \{1, \dots, k\}$ and show (3.29), dropping the dependence on j from all the notation. Let L be the subset of points $[\hat{a}, \hat{b}]$ for which the Lebesgue differentiation theorem for $\hat{\mu}$ holds. From (3.3) we know that a.e. on $(S \setminus S') \cap L$ the function $F^V_{\hat{\mu}}(x) - f$ vanishes, this proves the first equality in (3.29), since L is of full Lebesgue measure. We next observe that a.e. on S' we have $F^V_{\hat{\mu}}(x) - f \leq 0$ and $\hat{\mu}(x) = 0$ — this proves the second inequality in (3.29).

Let us denote by $R = \{x \in [\hat{a}, \hat{b}] : F^V_{\hat{\mu}}(x) > f\}$ and observe that

$$\int_{S''} [F^V_{\hat{\mu}}(x) - f] \left(\tilde{\mu}_N(x) - \hat{\mu}(x)\right) dx = \int_{R \cap S''} [F^V_{\hat{\mu}}(x) - f] \left(\tilde{\mu}_N(x) - \theta^{-1} \cdot f_{\mathtt{q}}(\sigma_{\mathtt{q}}^{-1}(x))^{-1}\right) dx.$$

To see the latter we first observe that on S'' we have $\hat{\mu}(x) = \theta^{-1} \cdot f_{\mathtt{q}}(x)^{-1}$. In addition, we know that a.e. point in S'' belongs to the support of $\hat{\mu}$, and so by (3.3) a.e. on S'' we have that $F^V_{\hat{\mu}}(x) - f \geq 0$. Finally, we can remove the points of equality as they do not contribute to the integral. Next,

$$\begin{aligned} &\int_{R \cap S''} [F^V_{\hat{\mu}}(x) - f] \left(\tilde{\mu}_N(x) - \theta^{-1} \cdot f_{\mathtt{q}}(\sigma_{\mathtt{q}}^{-1}(x))^{-1}\right) dx \\ &\quad = \int_{R} [F^V_{\hat{\mu}}(x) - f] \left(\tilde{\mu}_N(x) - \theta^{-1} \cdot f_{\mathtt{q}}(\sigma_{\mathtt{q}}^{-1}(x))^{-1}\right) dx. \end{aligned}$$

The above follows from the fact that $R \cap (S'')^c$ has zero Lebesgue measure, which we know from (3.3). We have thus reduced the proof of the proposition to establishing

$$\int_{R} [F^V_{\hat{\mu}}(x) - f] \left(\tilde{\mu}_N(x) - \theta^{-1} \cdot f_{\mathtt{q}}(\sigma_{\mathtt{q}}^{-1}(x))^{-1}\right) dx \leq 0 + O(N^{-1}\log(N)). \tag{3.30}$$

Step 7. In this and the next step we establish (3.30). We start by noting that if $\kappa = \mathtt{q}^{-\mathtt{M}-\theta}$ then because $\hat{\mu}(x)$ is bounded we have

$$\sup_{x,y \in \mathbb{R} : |x-y| \leq \kappa N^{-1}} \left|F^V_{\hat{\mu}}(x) - F^V_{\hat{\mu}}(y)\right| = O\left(N^{-1}\log(N)\right). \tag{3.31}$$

In particular, the above implies that $F^V_{\hat{\mu}}$ is continuous and so R is an open set. We denote $\sigma_{q_N^N}(y) = q_N^{-Ny} + u_N q_N^{Ny}$ and perform the change of variables $x = \sigma_{q_N^N}(y)$ to rewrite the LHS in (3.30) as

$$\int_{\sigma^{-1}_{q_N^N}(R)} \left[F^V_{\hat{\mu}}\left(\sigma_{q_N^N}(y)\right) - f\right] \left(\tilde{\mu}_N(\sigma_{q_N^N}(y)) - \theta^{-1} \cdot f_{\mathtt{q}}\left(\sigma_{\mathtt{q}}^{-1}(\sigma_{q_N^N}(y))\right)^{-1}\right) \sigma'_{q_N^N}(y) dy.$$

We observe that by Assumption 1 we have

$$f_{\mathfrak{q}}\left(\sigma_{\mathfrak{q}}^{-1}(\sigma_{q_N^N}(y))\right)^{-1}\cdot\sigma'_{q_N^N}(y)=1+O(N^{-1}).$$

If we set $O=\sigma_{q_N^N}^{-1}(R)$ and $\tilde{\rho}_N(y)=\tilde{\mu}_N(\sigma_{q_N^N}(y))\cdot\sigma'_{q_N^N}(y)$ we may rewrite the LHS in (3.30) as

$$\int_O\left[F_{\hat{\mu}}^V\left(\sigma_{q_N^N}(y)\right)-f\right]\left(\tilde{\rho}_N(y)-\theta^{-1}\right)dy+O(N^{-1}).$$

Since $\sigma_{q_N^N}(y)$ is an invertible diffeomorphism we know that O is an open subset of $\left[\sigma_{q_N^N}^{-1}(\hat{a}),\sigma_{q_N^N}^{-1}(\hat{b})\right]$. In particular, we can find a collection of disjoint open intervals $(s_i,t_i), i\in J$ with J countable such that $O=\cup_{i\in J}(s_i,t_i)$, upto the endpoints $\sigma_{q_N^N}^{-1}(\hat{a}),\sigma_{q_N^N}^{-1}(\hat{b})$.

Since the sum of the lengths of these intervals is at most $\mathtt{M}-1+\theta$ we have finitely many such that $t_i-s_i>\theta\cdot N^{-1}$. Let us further subdivide such segments into segments of length exactly $\theta\cdot N^{-1}$, which are contained in (t_i,s_i) as well as edge segments $(s_i,x],[y,t_i)$ with length at most $\theta\cdot N^{-1}$. In this way we obtain a finite collection $\{K_i:=[r_i,r_i+\theta/N]\}_{i\in J_1}$ and a countable collection $\{[c_i,d_i]\}_{i\in J_2}$ of intervals such that

$$\begin{aligned}&\int_O\left[F_{\hat{\mu}}^V\left(\sigma_{q_N^N}(y)\right)-f\right]\left(\tilde{\rho}_N(y)-\theta^{-1}\right)dy=\\&\sum_{i\in J_1}\int_{K_i}\left[F_{\hat{\mu}}^V\left(\sigma_{q_N^N}(y)\right)-f\right]\left(\tilde{\rho}_N(y)-\theta^{-1}\right)dy+\\&\sum_{i\in J_2}\int_{c_i}^{d_i}\left[F_{\hat{\mu}}^V\left(\sigma_{q_N^N}(y)\right)-f\right]\left(\tilde{\rho}_N(y)-\theta^{-1}\right)dy\end{aligned}\tag{3.32}$$

and also $d_i-c_i\leq\theta/N$, and at least one of the points c_i,d_i is a boundary point of O. Our goal for the remainder is to show that the sums over J_1 and J_2 are both dominated by $0+O(N^{-1}\log(N))$. This would conclude the proof of (3.30).

Notice that by the continuity of $F_{\hat{\mu}}^V\left(\sigma_{q_N^N}(y)\right)$ and the definition of O, we know that on boundary points of this set we have that $F_{\hat{\mu}}^V\left(\sigma_{q_N^N}(y)\right)=f$. In particular, for the sum over J_2 in (3.32)

$$\left|F_{\hat{\mu}}^V\left(\sigma_{q_N^N}(y)\right)-f\right|\leq\sup_{\substack{0\leq x,y\leq\mathtt{M}+\theta\\|x-y|\leq N^{-1}}}\left|F_{\hat{\mu}}^V\left(\sigma_{q_N^N}(y)\right)-F_{\hat{\mu}}^V\left(\sigma_{q_N^N}(y)\right)\right|.$$

Since $\sigma'_{q_N^N}(x)\leq 2\kappa$ on $[0,\mathtt{M}+\theta]$ for large N, we conclude from (3.31) that

$$\sup_{\substack{0\le x,y\le \mathtt{M}+\theta \\ |x-y|\le N^{-1}}} \left|F^V_{\hat{\mu}}\left(\sigma_{q_N^N}(x)\right) - F^V_{\hat{\mu}}\left(\sigma_{q_N^N}(y)\right)\right|$$
$$\le \sup_{\substack{x,y\in\mathbb{R} \\ |x-y|\le 2\kappa N^{-1}}} \left|F^V_{\hat{\mu}}(x) - F^V_{\hat{\mu}}(y)\right| = O(N^{-1}\log(N)). \tag{3.33}$$

We conclude that the sum over J_2 in (3.32) is bounded in absolute value by

$$O(N^{-1}\log(N)) \cdot \sum_{i\in J_2} \int_{c_i}^{d_i} \left(\tilde{\rho}_N(y) + \theta^{-1}\right) dy = O(N^{-1}\log(N)).$$

We are left with estimating the sum over J_1, which we do in the next step.

Step 8. To conclude the proof what remains is to show

$$\sum_{m\in J_1} \int_{r_m}^{r_m+\theta/N} \left[F^V_{\hat{\mu}}\left(\sigma_{q_N^N}(y)\right) - f\right]\left(\tilde{\rho}_N(y) - \theta^{-1}\right) dy \le 0 + O(N^{-1}\log(N)). \tag{3.34}$$

We first recall that by definition $\tilde{\mu}_N(x) = \sum_{i=1}^N N^{p-1}\cdot \mathbf{1}[\ell_i, \ell_i + N^{-p}](x)$, where $\ell = (\ell_1,\dots,\ell_N) \in \mathfrak{X}_N^\theta$ and $\mathbf{1}K$ stands for the indicator of the set K. In particular,

$$\tilde{\rho}_N(y) = \sum_{i=1}^N N^{p-1}\cdot \mathbf{1}A_i(y)\cdot \sigma'_{q_N^N}(y), \text{ where } A_i = \left[\lambda_i/N, \sigma^{-1}_{q_N^N}(\ell_i + N^{-p})\right]$$

and λ_i are such that $\ell_i = q_N^{-\lambda_i} + u_N q_N^{\lambda_i}$.

Since $\lambda_{i+1} - \lambda_i \ge \theta$, we know that each interval $K_m = [r_m, r_m + \theta/N]$ intersects at most two of the intervals A_i. If it intersects at most one we know that

$$\tilde{\rho}_N(K_m) \le \int_{A_i} \tilde{\rho}_N(y)dy = 1/N.$$

If it intersects two then they must be A_i and A_{i+1} for some i such that $\lambda_{i+1} - \lambda_i = \theta$. Let us note that $\sigma'_{q_N^N}(y) = \sigma'_{q_N^N}(\lambda_i/N) + O(N^{-1})$ whenever $y \in [\lambda_i/N, (\lambda_i + 2\theta)/N]$. In addition, we have

$$\sigma^{-1}_{q_N^N}(\ell_i + N^{-p}) = \lambda_i/N + N^{-p}\cdot\left[\frac{d}{dx}\sigma^{-1}_{q_N^N}\right](\ell_i) + O(N^{-2p}) \text{ and}$$
$$\sigma^{-1}_{q_N^N}(\ell_{i+1} + N^{-p}) = \lambda_{i+1}/N + N^{-p}\cdot\left[\frac{d}{dx}\sigma^{-1}_{q_N^N}\right](\ell_i) + O(N^{-p-1}).$$

Combining the above estimates, and setting $\gamma = N^{-p}\cdot\left[\frac{d}{dx}\sigma^{-1}_{q_N^N}\right](\ell_i)$, we see that

$$\tilde{\rho}_N(K_m) = \int_{K_m} N^{p-1} \left[\mathbf{1}A_i(y) + \mathbf{1}A_{i+1}(y)\right] \cdot \sigma'_{q_N^N}(y)dy =$$

$$\int_{K_m} N^{p-1} \left[\mathbf{1}[\lambda_i/N, \lambda_i/N + \gamma](y) + \mathbf{1}[\lambda_{i+1}/N, \lambda_{i+1}/N + \gamma](y)\right] \cdot \sigma'_{q_N^N}(\lambda_i/N)dy + O(N^{-2}). \tag{3.35}$$

The key observation is that the integral in the second line of (3.35) is precisely $\sigma'_{q_N^N}(\lambda_i/N) \cdot N^{p-1} \cdot \gamma = N^{-1}$. Consequently, we obtain the estimate

$$\tilde{\rho}_N(K_m) \leq N^{-1} + O(N^{-2}) \text{ for each } m \in J_1.$$

Combining the latter with the fact that $F_{\hat{\mu}}^V\left(\sigma_{q_N^N}(r_m)\right) \geq f$ we see that

$$\int_{K_m} \left[F_{\hat{\mu}}^V\left(\sigma_{q_N^N}(y)\right) - f\right] \left(\tilde{\rho}_N(y) - \theta^{-1}\right) dy \leq$$

$$\int_{K_m} \left[F_{\hat{\mu}}^V\left(\sigma_{q_N^N}(y)\right) - F_{\hat{\mu}}^V\left(\sigma_{q_N^N}(r_m)\right)\right] \left(\tilde{\rho}_N(y) - \theta^{-1}\right) dy + O(N^{-2}) \leq$$

$$\sup_{\substack{0 \leq x,y \leq \mathtt{M}+\theta \\ |x-y| \leq N^{-1}}} 2N^{-1} \left|F_{\hat{\mu}}^V\left(\sigma_{q_N^N}(x)\right) - F_{\hat{\mu}}^V\left(\sigma_{q_N^N}(y)\right)\right| + O(N^{-2}) =$$

$$O(N^{-2}\log(N)),$$

where the last equality follows from (3.33). Combining the above estimates over $m \in J_1$ we get

$$\sum_{m \in J_1} \int_{K_m} \left[F_{\hat{\mu}}^V\left(\sigma_{q_N^N}(y)\right) - f\right] \left(\tilde{\rho}_N(y) - \theta^{-1}\right) dy \leq 0 + O(N^{-2}\log(N)) \cdot |J_1|.$$

The above implies (3.34) since $(\theta/N) \cdot |J_1| \leq \mathtt{M} + \theta - 1$.

4. Nekrasov's equations

In this section we present the main algebraic component in our arguments, which we call the *Nekrasov's equations* – Theorem 4.1.1. In Section 4.2 we study the asymptotics of this equation and explain how it gives rise to a functional equation for the equilibrium measure from Section 3.

4.1. Formulation

As explained earlier the measure $\mathbb{P}_N$ in (2.3) can be understood as a discretization of the continuous β log-gas to shifted quadratic lattices. In [15] the authors consider

a different discretization (called *discrete β-ensembles*) where the particles occupy (appropriately shifted) integer lattices. They manage to obtain results about the global fluctuations of these particle systems and their analysis is based on appropriate discrete versions of the Schwinger–Dyson equations, which they also call the Nekrasov's equations. More recently, in [33] the same Nekrasov's equations were used to prove rigidity and edge universality for the models in [15].

Motivated by the success of the Nekrasov's equations for the discrete β-ensembles, we develop appropriate q-analogues that are applicable for the measures (2.3). The key result is given below and can be understood as a version of the Nekrasov's equation for shifted quadratic lattices.

Theorem 4.1.1. *Let $\mathbb{P}_N$ be a probability distribution on $\mathfrak{X}_N^\theta$ as in (2.3). Let $\mathcal{M} \subset \mathbb{C}$ be open and*

$$\bigcup_{i=1}^{k} \left(\left[q^{1-a_i-N_{i-1}\cdot\theta}, q^{-b_i-(N_i-1)\cdot\theta} \right] \cup \left[uq^{b_i+(N_i-1)\cdot\theta}, uq^{a_i-1+N_{i-1}\cdot\theta} \right] \right) \subset \mathcal{M}.$$

Suppose there exist two functions $\Phi^+(z)$ and $\Phi^-(z)$ that are analytic in $\mathcal{M}$ and such that

$$\frac{w(\sigma(z);N)}{w(\sigma(qz);N)} = \frac{(q^2z^2-uq^\theta)(z^2-u)\Phi^+(z)}{(q^\theta z^2-u)(q^2z^2-u)\Phi^-(z)}, \text{ whenever } \sigma(z), \sigma(qz) \in \cup_{i=1}^{k}[\alpha_i, \beta_i^-] \ , \tag{4.1}$$

where $\sigma(z) = z + uz^{-1}$. We also assume that $\Phi^\pm$ satisfy for each $i \in \{1, \dots, k\}$

$$\begin{aligned} \Phi^+\left(q^{-b_i-(N_i-1)\theta}\right) &= \Phi^-\left(q^{-a_i-N_{i-1}\cdot\theta}\right) = \Phi^+\left(uq^{a_i-1+N_{i-1}\cdot\theta}\right) \\ &= \Phi^-\left(uq^{b_i-1+(N_i-1)\cdot\theta}\right) = 0. \end{aligned} \tag{4.2}$$

If we define

$$\tilde{R}(z) = \Phi^-(z)\cdot\mathbb{E}_{\mathbb{P}_N}\left[\prod_{i=1}^{N}\frac{\sigma(q^\theta z)-\ell_i}{\sigma(z)-\ell_i}\right] + \Phi^+(z)\cdot\mathbb{E}_{\mathbb{P}_N}\left[\prod_{i=1}^{N}\frac{\sigma(q^{1-\theta}z)-\ell_i}{\sigma(qz)-\ell_i}\right] \tag{4.3}$$

then $\tilde{R}(z)$ is analytic in the same complex neighborhood $\mathcal{M}$. Moreover, if $\Phi^\pm(z)$ are polynomials of degree at most d, then so is $\tilde{R}(z)$.

Proof. As usual, see (2.2) we set $q^{-\lambda_i} + uq^{\lambda_i} = \ell_i$ for $i = 1, \dots, N$. Then we have

$$\begin{aligned} &\frac{\sigma(q^\theta z)-\ell_i}{\sigma(z)-\ell_i} = \frac{1}{q^\theta(q^{-\lambda_i}-uq^{\lambda_i})}\cdot(q^\theta z - q^{-\lambda_i})(q^\theta z - uq^{\lambda_i})\left(\frac{1}{z-q^{-\lambda_i}} - \frac{1}{z-uq^{\lambda_i}}\right) \\ &\frac{\sigma(q^{1-\theta}z)-\ell_i}{\sigma(qz)-\ell_i} = \end{aligned} \tag{4.4}$$

$$\frac{q^{\theta}}{(q^{-\lambda_i}-uq^{\lambda_i})}\cdot(q^{1-\theta}z-q^{-\lambda_i})(q^{1-\theta}z-uq^{\lambda_i})\left(\frac{1}{qz-q^{-\lambda_i}}-\frac{1}{qz-uq^{\lambda_i}}\right).$$

From the above we see that the possible singularities of $\tilde{R}(z)$ in $\mathcal{M}$ are simple poles at points q^{-m} and uq^{m} whenever $\lambda_i = m$ in the first line of (4.4) and $\lambda_i = m-1$ in the second line of (4.4).

We will separately compute the residue contribution coming from each $i = 1, \dots, N$, which we fix for the remainder. We also let $j \in \{1, \dots, k\}$ be the unique index such that $N_{j-1}+1 \le i \le N_j$. By definition, we know that λ_i varies in the set $\{C, C+1, \dots, D\}$, where $C = (i-1)\cdot\theta + a_j$ and $D = (i-1)\cdot\theta + b_j - 1$. If m lies in $\{C, C+1, \dots, D\}$, we see that the residue at q^{-m} is given by

$$\sum_{\ell\in\mathfrak{X}_N^{\theta}|\lambda_i=m} A(\ell,i,m) + \sum_{\ell\in\mathfrak{X}_N^{\theta}|\lambda_i=m-1} B(\ell,i,m), \tag{4.5}$$

where

$$A(\ell,i,m) = \frac{\Phi^-(q^{-m})(q^{\theta}q^{-m}-q^{-m})(q^{\theta}q^{-m}-uq^{m})}{q^{\theta}(q^{-m}-uq^{m})}\left[\mathbb{P}_N(\ell)\prod_{j\neq i}\frac{\sigma(q^{\theta-m})-\ell_j}{\sigma(q^{-m})-\ell_j}\right]$$

$$B(\ell,i,m) = \frac{\Phi^+(q^{-m})q^{\theta-1}(q^{1-\theta}q^{-m}-q^{-m+1})(q^{1-\theta}q^{-m}-uq^{m-1})}{(q^{-m+1}-uq^{m-1})} \cdot\left[\mathbb{P}_N(\ell)\prod_{j\neq i}\frac{\sigma(q^{1-\theta-m})-\ell_j}{\sigma(q^{1-m})-\ell_j}\right].$$

Let us fix $\ell_1, \dots, \ell_{i-1}$ and $\ell_{i+1}, \dots, \ell_N$ and set $\ell^+ = (\ell_1, \dots, \ell_{i-1}, q^{-m}+uq^{m}, \ell_{i+1}, \dots, \ell_N)$, $\ell^- = (\ell_1, \dots, \ell_{i-1}, q^{-m+1}+uq^{m-1}, \ell_{i+1}, \dots, \ell_N)$ – notice that ℓ^+, ℓ^- are not necessarily in $\mathfrak{X}_N^{\theta}$. We claim that $A(\ell^+,i,m)+B(\ell^-,i,m)=0$, where we set A and B to be zero if the argument is not in $\mathfrak{X}_N^{\theta}$. If true, we would obtain that the sum in (4.5) is zero and so R is analytic near q^{-m}. The latter statement is clear if both $\ell^{\pm}\notin\mathfrak{X}_N^{\theta}$, hence we assume at least one of them belongs to the state space.

If $m = a_j + (i-1)\cdot\theta$ then $B(\ell^-,i,m)=0$ since $\lambda_i = a_j - 1 + (i-1)\cdot\theta$ (and so $\ell^-\notin\mathfrak{X}_N^{\theta}$). In addition, $A(\ell^+,i,m)=0$, since either $i = N_{j-1}+1$ and then $\Phi^-(q^{-m}) = \Phi^-(q^{-a_j-N_{j-1}\theta}) = 0$ or $i \ge N_{j-1}+2$ and then $\lambda_{i-1} = a_j + (i-2)\cdot\theta$ so that the factor $\left(\sigma(q^{\theta-m})-\ell_{i-1}\right)$ vanishes. Similarly, we have $B(\ell^-,i,m) = 0 = A(\ell^+,i,m)$ if $m = b_j + (i-1)\cdot\theta$. If $\lambda_{i-1} = m-\theta$, we know that $B(\ell^-,i,m)=0$, since $\ell^-\notin\mathfrak{X}_N^{\theta}$, but also $A(\ell^+,i,m)=0$ as it has the factor $\left(\sigma(q^{\theta-m})-\ell_{i-1}\right)$. Similarly, we have $B(\ell^-,i,m) = 0 = A(\ell^+,i,m)$ if $\lambda_{i+1} = m+\theta$. We may thus assume that $\lambda_{i-1} \le m-1-\theta < m+1+\theta \le \lambda_{i+1}$, $a_j + 1 \le m-(i-1)\theta \le b_j - 1$ and that $\ell^{\pm}\in\mathfrak{X}_N^{\theta}$.

We next observe that

$$\frac{(q^{\theta}q^{-m}-q^{-m})(q^{\theta}q^{-m}-uq^{m})}{q^{\theta}(q^{-m}-uq^{m})}\cdot\frac{(q^{-m+1}-uq^{m-1})}{q^{\theta-1}(q^{1-\theta}q^{-m}-q^{-m+1})(q^{1-\theta}q^{-m}-uq^{m-1})} =$$

$$= -\frac{(q^{\theta}z^2-u)(q^2z^2-u)}{(q^2z^2-uq^{\theta})(z^2-u)}\Big|_{z=q^{-m}}.$$

Therefore, from the definition of Φ^+ and Φ^- we get

$$\frac{A(\ell^+,i,m)}{B(\ell^-,i,m)} = -\frac{\mathbb{P}_N(\ell^+)w(\sigma(q^{-m+1});N)}{\mathbb{P}_N(\ell^-)w(\sigma(q^{-m});N)}\prod_{j\neq i}\left[\frac{\sigma(q^{\theta-m})-\ell_j}{\sigma(q^{-m})-\ell_j}\cdot\frac{\sigma(q^{1-m})-\ell_j}{\sigma(q^{1-\theta-m})-\ell_j}\right]. \quad (4.6)$$

Our goal for the remainder is to show that the right side in (4.6) is equal to -1.

In view of (2.3) we have that $\frac{\mathbb{P}_N(\ell^+)w(\sigma(q^{-m+1});N)}{\mathbb{P}_N(\ell^-)w(\sigma(q^{-m});N)}$ equals

$$\prod_{1\le l<i} q^{-2\theta m}\frac{\Gamma_q(m-\lambda_l+1)\Gamma_q(m-\lambda_l+\theta)}{\Gamma_q(m-\lambda_l)\Gamma_q(m-\lambda_l+1-\theta)}\cdot q^{2\theta(m-1)}\frac{\Gamma_q(m-\lambda_l-1)\Gamma_q(m-\lambda_l-\theta)}{\Gamma_q(m-\lambda_l)\Gamma_q(m-\lambda_l+\theta-1)}\times$$
$$\prod_{1\le l<i}\frac{\Gamma_q(m+\lambda_l+v+1)\Gamma_q(m+\lambda_l+v+\theta)}{\Gamma_q(m+\lambda_l+v)\Gamma_q(m+\lambda_l+v+1-\theta)}\cdot\frac{\Gamma_q(m+\lambda_l-1+v)\Gamma_q(m+\lambda_l+v-\theta)}{\Gamma_q(m+\lambda_l+v)\Gamma_q(m+\lambda_l+v+\theta-1)}\times$$
$$\prod_{i<j\le N} q^{-2\theta\lambda_j}\frac{\Gamma_q(\lambda_j-m+1)\Gamma_q(\lambda_j-m+\theta)}{\Gamma_q(\lambda_j-m)\Gamma_q(\lambda_j-m+1-\theta)}\cdot q^{2\theta\lambda_j}\frac{\Gamma_q(\lambda_j-m+1)\Gamma_q(\lambda_j-m+2-\theta)}{\Gamma_q(\lambda_j-m+2)\Gamma_q(\lambda_j-m+1+\theta)}$$
$$\prod_{i<j\le N}\frac{\Gamma_q(\lambda_j+m+v+1)\Gamma_q(\lambda_j+m+v+\theta)}{\Gamma_q(\lambda_j+m+v)\Gamma_q(\lambda_j+m+v+1-\theta)}\cdot\frac{\Gamma_q(\lambda_j+m-1+v)\Gamma_q(\lambda_j+m+v-\theta)}{\Gamma_q(\lambda_j+m+v)\Gamma_q(\lambda_j+m+v+\theta-1)}. \quad (4.7)$$

Using (2.4) and that $q^v=u$ we can rewrite (4.7) as

$$\prod_{1\le l<i} q^{-2\theta}\frac{(1-q^{m-\lambda_l})(1-q^{m-\lambda_l+\theta-1})}{(1-q^{m-\lambda_l-1})(1-q^{m-\lambda_l-\theta})}\times\frac{(1-uq^{m+\lambda_l})(1-uq^{m+\lambda_l+\theta-1})}{(1-uq^{m+\lambda_l-1})(1-uq^{m+\lambda_l-\theta})}$$
$$\prod_{i<j\le N}\frac{(1-q^{\lambda_j-m})(1-q^{\lambda_j-m+1-\theta})}{(1-q^{\lambda_j-m+1})(1-q^{\lambda_j-m+\theta})}\times\frac{(1-uq^{\lambda_j+m})(1-uq^{\lambda_j+m+1-\theta})}{(1-uq^{\lambda_j+m+1})(1-uq^{\lambda_j+m+\theta})}. \quad (4.8)$$

We next observe that for $l<i$ we have

$$q^{-2\theta}\frac{(1-q^{m-\lambda_l})(1-q^{m-\lambda_l+\theta-1})}{(1-q^{m-\lambda_l-1})(1-q^{m-\lambda_l-\theta})}\times\frac{(1-uq^{m+\lambda_l})(1-uq^{m+\lambda_l+\theta-1})}{(1-uq^{m+\lambda_l-1})(1-uq^{m+\lambda_l-\theta})} =$$
$$\frac{(q^{-m}-q^{-\lambda_l})(q^{1-m-\theta}-q^{-\lambda_l})}{(q^{1-m}-q^{-\lambda_l})(q^{\theta-m}-q^{-\lambda_l})}\times\frac{(1-uq^{m+\lambda_l})(1-uq^{m+\lambda_l+\theta-1})}{(1-uq^{m+\lambda_l-1})(1-uq^{m+\lambda_l-\theta})} =$$
$$\frac{[\sigma(q^{-m})-\sigma(q^{-\lambda_l})]\cdot[\sigma(q^{1-\theta-m})-\sigma(q^{-\lambda_l})]}{[\sigma(q^{1-m})-\sigma(q^{-\lambda_l})]\cdot[\sigma(q^{\theta-m})-\sigma(q^{-\lambda_l})]},$$

where we used that $(z+uz^{-1}-t-ut^{-1})=(z-t)(1-uz^{-1}t^{-1})$.

One similarly establishes that for $i<j$ we have

$$\frac{(1-q^{\lambda_j-m})(1-q^{\lambda_j-m+1-\theta})}{(1-q^{\lambda_j-m+1})(1-q^{\lambda_j-m+\theta})}\times\frac{(1-uq^{\lambda_j+m})(1-uq^{\lambda_j+m+1-\theta})}{(1-uq^{\lambda_j+m+1})(1-uq^{\lambda_j+m+\theta})} =$$

$$\frac{(q^{-\lambda_j}-q^{-m})(q^{-\lambda_j}-q^{-m+1-\theta})}{(q^{-\lambda_j}-q^{-m+1})(q^{-\lambda_j}-q^{-m+\theta})}\times\frac{(1-uq^{\lambda_j+m})(1-uq^{\lambda_j+m+1-\theta})}{(1-uq^{\lambda_j+m+1})(1-uq^{\lambda_j+m+\theta})}=$$

$$\frac{[\sigma(q^{-\lambda_j})-\sigma(q^{-m})]\cdot[\sigma(q^{-\lambda_j})-\sigma(q^{1-\theta-m})]}{[\sigma(q^{-\lambda_j})-\sigma(q^{1-m})]\cdot[\sigma(q^{-\lambda_j})-\sigma(q^{\theta-m})]}.$$

The last two calculations together with (4.8) show the right side of (4.6) is equal to -1 as desired. This proves that $\tilde{R}$ is analytic near q^{-m}. One can use analogous arguments to show that R is also analytic near the points uq^m and so on all of $\mathcal{M}$.

Notice that if $\Phi^{\pm}(z)$ are polynomials of degree at most d then $\tilde{R}(z)$ is entire from the first part of the theorem, which grows as $O(|z|^d)$ as $|z|\to\infty$. By Liuoville's theorem $\tilde{R}(z)$ is a polynomial of degree at most d. □

Remark 4.1.2. Theorem 4.1.1 also holds if $u=0$, where (4.1) is replaced with $\frac{w(z;N)}{w(qz;N)}=q^{\theta}\cdot\frac{\Phi^{+}(z)}{\Phi^{-}(z)}$ and the second two equalities in (4.2) are removed. In this case (4.4) only produces possible poles at $z=q^{-m}$. From here the proof proceeds in the same way and can be found as [15, Theorem 4.2].

4.2. Asymptotics of the Nekrasov's equations

In this section we derive some properties of the equilibrium measure μ and R_μ, Q_μ from (2.12) using the asymptotics of the Nekrasov's equation (4.3) as $N\to\infty$ under Assumptions 1–4 and 6–7. We assume the same notation as in Section 2.2.

Lemma 4.2.1. *Suppose that Assumptions 1–4 and 6 from Section 2.2 hold. Then the functions R_μ and Q_μ^2 from (2.12) are analytic on $\mathcal{M}$. If $\Phi_N^{\pm}$ are polynomials of degree at most d then so is R_μ and Q_μ^2 is a polynomial of degree at most $2d$. If Assumption 7 also holds then R_μ and Q_μ^2 are real analytic on $\mathcal{M}\cap\mathbb{R}$.*

Proof. We observe that by Assumptions 4, and 6, the Nekrasov's equation (4.3) holds and so

$$\tilde{R}_N(z):=\Phi_N^{-}(z)\cdot\mathbb{E}_{\mathbb{P}_N}\left[\prod_{i=1}^{N}\frac{\sigma_N(q_N^{\theta}z)-\ell_i}{\sigma_N(z)-\ell_i}\right]+\Phi_N^{+}(z)\cdot\mathbb{E}_{\mathbb{P}_N}\left[\prod_{i=1}^{N}\frac{\sigma(q_N^{1-\theta}z)-\ell_i}{\sigma(q_Nz)-\ell_i}\right] \tag{4.9}$$

defines an analytic function on $\mathcal{M}$. For μ_N as in Section 3.1 define

$$\mathfrak{G}_N^{d}(z):=N\log q_N\cdot(z-u_Nz^{-1})\cdot\int_{\mathbb{R}}\frac{\mu_N(dx)}{z+u_Nz^{-1}-x}.$$

One readily observes that

$$\prod_{i=1}^{N} \frac{\sigma_N(q_N^{\theta} z) - \ell_i}{\sigma_N(z) - \ell_i} = \exp\left[\theta \mathfrak{G}_N^d(z) + O(N^{-1})\right],$$
$$\prod_{i=1}^{N} \frac{\sigma_N(q_N^{1-\theta} z) - \ell_i}{\sigma_N(q_N z) - \ell_i} = \exp\left[-\theta \mathfrak{G}_N^d(z) + O(N^{-1})\right],$$

where the constants in the big O notation are uniform as z varies over compact subsets of $\mathcal{M} \setminus \left\{\cup_{i=1}^{k}[\hat{a}_i, \hat{b}_i]\right\}$. In addition, by Theorem 3.1.1 we know that $\mathfrak{G}_N^d(z)$ converges in probability to $G_\mu(z + \mathtt{u} z^{-1})$. An application of the Bounded convergence theorem and Assumption 4 implies that

$$\lim_{N\to\infty} \tilde{R}_N(z) = R_\mu(z), \tag{4.10}$$

where the convergence is uniform over compact subsets of $\mathcal{M} \setminus \left\{\cup_{i=1}^{k}[\hat{a}_i, \hat{b}_i]\right\}$. Since $\tilde{R}_N(z)$ are analytic in $\mathcal{M}$ we conclude the same is true for $R_\mu(z)$. Next, since $Q_\mu^2(z) = R_\mu^2(z) - 4\Phi^+(z)\Phi^-(z)$, we conclude from Assumption 4 that $Q_\mu^2(z)$ is also analytic in $\mathcal{M}$. The real analyticity of R_μ and Q_μ^2 is a consequence of the one assumed for $\Phi^{\pm}$ in Assumption 7.

If $\Phi_N^{\pm}$ are polynomials of degree at most d then by Theorem 4.1.1 we know that so is $\tilde{R}_N(z)$. The uniform convergence of $\tilde{R}_N(z)$ over compact sets in $\mathcal{M} \setminus \left\{\cup_{i=1}^{k}[\hat{a}_i, \hat{b}_i]\right\}$ is equivalent to the convergence of the coefficients of the polynomials, and so $R_\mu(z)$ is a polynomial of degree at most d. Finally, the same argument shows $\Phi^{\pm}(z)$ are polynomials of degree at most d and $Q_\mu^2(z) = R_\mu^2(z) - 4\Phi^+(z)\Phi^-(z)$ is a polynomial of degree at most $2d$. □

Our next goal is to give a formula for the equilibrium measure μ in Theorem 3.1.1 in terms of the functions R_μ and $\Phi^{\pm}$ but we first introduce some notation that will be useful. From Assumption 7 we know that V is real analytic in an open neighborhood of $\cup_{i=1}^{k}[\hat{a}_i, \hat{b}_i]$ and from [51] we conclude that μ has a continuous density on each interval $[\hat{a}_i, \hat{b}_i]$. Borrowing terminology from [4], each of the intervals $[\hat{a}_i, \hat{b}_i]$ is split into three types of regions:

(1) Maximum (with respect to inclusion) closed intervals where $\mu(x) = 0$ are called *voids*.
(2) Maximal open intervals where $0 < \mu(x) < \theta^{-1} \cdot f_{\mathtt{q}}(\sigma_{\mathtt{q}}^{-1}(x))^{-1}$ are called *bands* (recall $f_{\mathtt{q}}$ was defined in Section 2.2).
(3) Maximal closed intervals where $\mu(x) = \theta^{-1} \cdot f_{\mathtt{q}}(\sigma_{\mathtt{q}}^{-1}(x))^{-1}$ are called *saturated regions*.

Lemma 4.2.2. *Suppose that Assumptions 1–4 and 6–7 from Section 2.2 hold. Then μ has density*[3]

$$\mu(y_0 + \mathtt{u}y_0^{-1}) = \frac{1}{\theta \log(\mathtt{q})\pi(y_0 - \mathtt{u}y_0^{-1})} \cdot arccos\left(\frac{R_\mu(y_0)}{2\sqrt{\Phi^-(y_0)\Phi^+(y_0)}}\right), \tag{4.11}$$

for $y_0 \geq 1$ such that $y_0 + \mathtt{u}y_0 \in \cup_{i=1}^k [\hat{a}_i, \hat{b}_i]$ and 0 otherwise.

Proof. We will assume that $\mathtt{u} > 0$, the case $\mathtt{u} = 0$ is simpler and can be handled similarly. As discussed earlier, Assumption 7 implies $\mu(x)$ is continuous on each interval $[\hat{a}_i, \hat{b}_i]$. By assumption there are unique $\hat{d}_i > \hat{c}_i \geq 1$ such that $\sigma(\hat{c}_i) = \hat{a}_i$ and $\sigma(\hat{d}_i) = \hat{b}_i$ for $i = 1, \dots, k$ where $\sigma(z) = z + \mathtt{u}z^{-1}$. Consequently, $\sigma^{-1}([\hat{a}_i, \hat{b}_i]) = [\hat{c}_i, \hat{d}_i] \cup [u\hat{d}_i^{-1}, u\hat{c}_i^{-1}]$ for $i = 1, \dots, k$ and all $2k$ of the latter intervals are disjoint. Let

$$\psi(y) := \begin{cases} \mu(\sigma(y)) & \text{if } \ y \in \cup_{i=1}^k \left\{ [\hat{c}_i, \hat{d}_i] \cup [u\hat{d}_i^{-1}, u\hat{c}_i^{-1}] \right\}, \\ 0 & \text{else.} \end{cases}$$

It follows from (2.11) that

$$(z - \mathtt{u}z^{-1}) G_\mu(z + \mathtt{u}z^{-1}) = \sum_{i=1}^k \int_{\hat{c}_i}^{\hat{d}_i} \psi(y)(1 - \mathtt{u}y^{-2}) \left[\frac{z}{z-y} - \frac{\mathtt{u}z^{-1}}{\mathtt{u}z^{-1} - y} \right] dy.$$

Using that

$$\frac{z}{z-y} = \frac{y}{z-y} + 1, \ \psi(y) = \psi(\mathtt{u}y^{-1}) \text{ and } \sum_{i=1}^k \int_{\mathtt{u}\hat{c}_i^{-1}}^{\mathtt{u}\hat{d}_i^{-1}} \psi(y)(1 - \mathtt{u}y^{-2}) dy = 1 \text{ we get}$$

$$\theta \log(\mathtt{q})(z - \mathtt{u}z^{-1}) G_\mu(z + \mathtt{u}z^{-1}) = \theta \log(\mathtt{q}) + F(z), \text{ where}$$

$$F(z) := \int_{\mathbb{R}} \frac{\theta \log(\mathtt{q}) \psi(y) |y - \mathtt{u}y^{-1}|}{z - y} dy.$$

Using [34, Theorem 2.1] and [69, Chapter 5, Theorem 93] we conclude that $F(x + iy)$ defines a regular function for $y > 0$ and

$$\begin{aligned} &\lim_{\varepsilon \to 0^+} F(x + i\varepsilon) = f(x) - ig(x) \text{ for a.e. } x \in \mathbb{R}, \text{ where } f, g \in L^2(\mathbb{R}) \text{ are given by} \\ &g(x) = \theta \log(\mathtt{q}) \pi \cdot \psi(x) |x - \mathtt{u}x^{-1}| \text{ and } f(x) = -P \int_{\mathbb{R}} \frac{g(t)}{t - x} dx, \end{aligned} \tag{4.12}$$

[3] Throughout the paper we denote by $\arccos(x)$ the function, which is π on $(-\infty, -1]$, 0 on $[1, \infty)$, and the usual arccosine function on $(-1, 1)$.

and P means that we take the integral in the principal value sense. Since g is continuous on $\cup_{i=1}^{k}[\hat{c}_i, \hat{d}_i]$, we can apply [56, Chapter 4] and conclude that f is continuous on $\cup_{i=1}^{k}(\hat{c}_i, \hat{d}_i)$.

Let us take $z = y_0 + i\varepsilon$ with $y_0 \in \cup_{i=1}^{k}(\hat{c}_i, \hat{d}_i)$ and let ε converge to 0^+ in (2.12). This gives

$$R_\mu(y_0) = \Phi^-(y_0)e^{\theta \log(\mathfrak{q})} \cdot \exp\left(f(y_0) - ig(y_0)\right) + \Phi^+(y_0) \cdot e^{-\theta \log(\mathfrak{q})} \exp\left(-f(y_0) + ig(y_0)\right).$$

The above defines a quadratic equation for $\exp\left(f(y_0) - ig(y_0)\right)$ and we conclude that

$$\{\exp\left(f(y_0) \pm ig(y_0)\right)\} = \left\{ \frac{R_\mu(y_0) \pm \sqrt{R_\mu^2(y_0) - 4\Phi^-(y_0)\Phi^+(y_0)}}{2\Phi^-(y_0)e^{\theta \log(\mathfrak{q})}} \right\}, \tag{4.13}$$

where the square root is with respect to the principal branch and assumed in $\mathbb{H}$ for negative values.

Suppose that $y_0 \in (\hat{c}_i, \hat{d}_i)$ and $\frac{R_\mu(y_0)}{2\sqrt{\Phi^-(y_0)\Phi^+(y_0)}} \in (-1, 1)$. Then the numbers in (4.13) are complex conjugates with non-zero imaginary part, and we have

$$\exp\left(f(y_0) + ig(y_0)\right) = \frac{R_\mu(y_0) + i\sqrt{-R_\mu^2(y_0) + 4\Phi^-(y_0)\Phi^+(y_0)}}{2\Phi^-(y_0)e^{\theta \log(\mathfrak{q})}}, \text{ since both lie in } \mathbb{H}.$$

Taking the argument on both sides of the above equation we see that

$$g(y_0) = \arccos\left(\frac{R_\mu(y_0)}{2\sqrt{\Phi^-(y_0)\Phi^+(y_0)}} \right) \in (0, \pi). \tag{4.14}$$

The above computation also shows that

$$y_0 + \mathfrak{u}y_0^{-1} \text{ belongs to a band of } \mu \text{ in } [\hat{a}_i, \hat{b}_i] \text{ if and only if } \frac{R_\mu(y_0)}{2\sqrt{\Phi^-(y_0)\Phi^+(y_0)}} \in (-1, 1). \tag{4.15}$$

If $y_0 \in (\hat{c}_i, \hat{d}_i)$ and $\frac{R_\mu(y_0)}{2\sqrt{\Phi^-(y_0)\Phi^+(y_0)}} \geq 1$ then the numbers in (4.13) are real and so $g(y_0) = 0$ or π, i.e. $y_0 + \mathfrak{u}y_0^{-1}$ belongs to a void or saturated region in $[\hat{a}_i, \hat{b}_i]$ for the measure μ, which we denote by $[s, t]$. Notice that $[s, t] \neq [\hat{a}_i, \hat{b}_i]$ by our assumption on the filling fractions ν_i. This implies that there is a band of μ in $[\hat{a}_i, \hat{b}_i]$ either ending at $s + \mathfrak{u}s^{-1}$ or starting from $t + \mathfrak{u}t^{-1}$. By continuity of g and (4.15) we see that $g(s) = 0$ or $g(t) = 0$, which implies that $g(y_0) = 0$. A similar argument shows that $g(y_0) = \pi$ if $y_0 \in (\hat{c}_i, \hat{d}_i)$ and $\frac{R_\mu(y_0)}{2\sqrt{\Phi^-(y_0)\Phi^+(y_0)}} \leq -1$. Combining the above statements with the definition of g concludes the proof of the lemma. □

We end the section by making a remark about the function $\mathfrak{G}_N^d(z)$ from the proof of Lemma 4.2.1, whose exponent is the observable we obtain from the Nekrasov's equation.

Remark 4.2.3. Let us assume for simplicity that $k = 1$ and set

$$\psi_N = \frac{1}{2N} \sum_{i=1}^{N} \left(\delta(q^{-\lambda_i}) + \delta(uq^{\lambda_i}) \right),$$

where λ_i is such that $q^{-\lambda_i} + uq^{\lambda_i} = \ell_i$ and we have dropped the dependence of q and u on N to ease notation. Then we have that

$$(z - uz^{-1}) \int_{\mathbb{R}} \frac{\mu_N(dx)}{z + uz^{-1} - x} = \int_{\mathbb{R}} \frac{(z - uz^{-1})\psi_N(dy)}{z + uz^{-1} - y - uy^{-1}} = \int_{\mathbb{R}} \psi_N(dy) \left[\frac{z}{z - y} - \frac{uz^{-1}}{uz^{-1} - y} \right].$$

Using that $\frac{z}{z-y} = \frac{y}{z-y} + 1$ we get

$$\int_{\mathbb{R}} \psi_N(dy) \left[\frac{z}{z - y} - \frac{uz^{-1}}{uz^{-1} - y} \right] = \int_{\mathbb{R}} \psi_N(dy) \left[\frac{y}{z - y} - \frac{y}{uz^{-1} - y} \right] =$$
$$\int_{\mathbb{R}} \psi_N(dy) \left[\frac{y}{z - y} - \frac{uy^{-1}}{uz^{-1} - uy^{-1}} \right] = 1 + \int_{\mathbb{R}} \frac{2y \cdot \psi_N(dy)}{z - y}.$$

The above computation shows that, upto a constant and negligible error, the observable $\mathfrak{G}_N^d(z)$ we obtain from the Nekrasov's equation is the Stieltjes transform of the (deformed) empirical measure

$$\frac{1}{N} \sum_{i=1}^{N} q^{-\lambda_i} \cdot \delta(q^{-\lambda_i}) + \frac{1}{N} \sum_{i=1}^{N} uq^{\lambda_i} \cdot \delta(uq^{\lambda_i}).$$

In [15] the Nekrasov's equation produced the exponent of the usual Stieltjes transform for the underlying particle system as an observable and the vanishing conditions on $\Phi^{\pm}$ the authors assumed, correspond to boundary conditions for that system. In our case, we see that in a sense we have two copies of particles sitting at $q^{-\lambda_i}$ and uq^{λ_i} and the vanishing assumptions in Theorem 4.1.1 play the role of boundary conditions for each copy. The authors are not aware of such a phenomenon occurring in other systems and would like to have a better conceptual understanding for its appearance.

5. Central limit theorem: Part I

Our goal in this and the next section is to study using Nekrasov's equation the fluctuations of the empirical measures μ_N, for which we proved the law of large numbers in Section 3. In Section 5.1 we introduce a $2m$-parameter deformation of the measures $\mathbb{P}_N$

and describe a certain map Υ_v. In Section 5.2 we state the main technical result of the section – Theorem 5.2.1, and deduce some corollaries from it. In Section 5.3 we explain how to employ Nekrasov's equation for the deformed measures. In Section 5.4 we give the proof of Theorem 5.2.1 modulo a certain asymptotic statement in equation (5.26), whose proof is the focus of Section 6.

5.1. Deformed measure

We adopt the same notation as in Section 2.2 and assume that Assumptions 1–6 hold. Introduce the usual random empirical measures μ_N on $\mathbb{R}$ through

$$\mu_N = \frac{1}{N}\sum_{i=1}^{N} \delta\left(\ell_i\right), \text{ where } (\ell_1, \ldots, \ell_N) \text{ is } \mathbb{P}_N\text{-distributed.} \tag{5.1}$$

We also define the (continuous, deterministic) probability measures

$$\hat{\mu}_N \text{ as in Assumption 5.} \tag{5.2}$$

It follows from Corollary 3.1.4 that $\mu_N - \hat{\mu}_N$ converges weakly in probability to 0. Our goal is to understand the fluctuations of $\mu_N - \hat{\mu}_N$.

Let us introduce the Stieltjes transforms of μ_N and $\hat{\mu}_N$ through

$$G_N^d(z) = \int_{\mathbb{R}} \frac{\mu_N(dx)}{z-x} \text{ and } G_N^c(z) = \int_{\mathbb{R}} \frac{\hat{\mu}_N(dx)}{z-x}. \tag{5.3}$$

Observe that the above formulas make sense whenever z does not lie in the support of the measures, and they define holomorphic functions there. Our study of $\mu_N - \hat{\mu}_N$ goes through understanding $G_N^d(z) - G_N^c(z)$ as $N \to \infty$. For that we introduce a deformed version of $\mathbb{P}_N$ following an approach that is similar to the one in [15].

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 $y \in \cup_{j=1}^{k} [\hat{a}_j, \hat{b}_j]$, and let the deformed distribution $\mathbb{P}_N^{\mathbf{t},\mathbf{v}}$ be defined through

$$\mathbb{P}_N^{\mathbf{t},\mathbf{v}}(\ell_1, \ldots, \ell_N) = Z(\mathbf{t}, \mathbf{v})^{-1} \prod_{1 \leq i < j \leq N} (\ell_i - \ell_j)^2 \prod_{i=1}^{N} \left[w(\ell_i; N) \prod_{a=1}^{k} \left(1 + \frac{t_a}{v_a - \ell_i}\right)\right]. \tag{5.4}$$

If $m = 0$ we have $\mathbb{P}_N^{\mathbf{t},\mathbf{v}} = \mathbb{P}_N$ is the undeformed measure. In general, $\mathbb{P}_N^{\mathbf{t},\mathbf{v}}$ may be a complex-valued measure but we always choose the normalization constant $Z(\mathbf{t}, \mathbf{v})$ so that $\sum_{\ell \in \mathfrak{X}} \mathbb{P}_N^{\mathbf{t},\mathbf{v}}(\ell) = 1$. In addition, we require that the numbers t_a are sufficiently close to zero so that $Z(\mathbf{t}, \mathbf{v}) \neq 0$.

Let us denote

$$\Delta G_N(z) = N(G_N^d(z) - G_N^c(z)), \text{ where } (\ell_1, \ldots, \ell_N) \text{ is } \mathbb{P}_N^{\mathbf{t},\mathbf{v}}\text{-distributed.} \tag{5.5}$$

Abusing notation we will suppress the dependence on $\mathbf{t}, \mathbf{v}$ of μ_N, G_N^d and ΔG_N when we replace $\mathbb{P}_N$ with $\mathbb{P}_N^{\mathbf{t},\mathbf{v}}$ in (5.1) and use the same letters. It will be clear from context, which formula we mean.

The definition of the deformed measure $\mathbb{P}^{\mathbf{t},\mathbf{v}}$ is motivated by the following observation.

Lemma 5.1.1. *Let ξ be a bounded random variable. For any $m \geq 1$ the mth mixed derivative*

$$\frac{\partial^m}{\partial t_1 \cdots \partial t_m} \mathbb{E}_{\mathbb{P}_N^{\mathbf{t},\mathbf{v}}}[\xi] \Big|_{t_a=0, 1 \leq a \leq m} \tag{5.6}$$

is the joint cumulant of the $m+1$ random variables $\xi, NG_N^d(v_1), \ldots, NG_N^d(v_m)$ with respect to $\mathbb{P}_N$.

Remark 5.1.2. The above result is analogous to Lemma 2.4 in [15], which in turn is based on earlier related work in random matrix theory [30,54]. We present a proof below for the sake of completeness.

Proof. One way to define the joint cumulant of $m+1$ bounded random variables $\xi_0, \ldots, \xi_m$ is through

$$\frac{\partial^{m+1}}{\partial t_0 \partial t_1 \cdots \partial t_m} \log \left(\mathbb{E} \exp \left(\sum_{i=0}^{m} t_i \xi_i \right) \right) \Big|_{t_i=0, 0 \leq i \leq m}.$$

Performing the differentiation with respect to t_0 we can rewrite the above as

$$\frac{\partial^m}{\partial t_1 \cdots \partial t_m} \frac{\mathbb{E}\left[\xi_0 \exp\left(\sum_{i=1}^m t_i \xi_i\right)\right]}{\mathbb{E}\left[\exp\left(\sum_{i=1}^m t_i \xi_i\right)\right]} \Big|_{t_i=0, 1 \leq i \leq m}.$$

Setting $\xi_0 = \xi$ and $\xi_i = NG_N^d(v_i)$ for $i = 1, \ldots, m$ and observing that

$$\exp\left(tNG_N^d(z)\right) = \prod_{i=1}^{N} \left(1 + \frac{t}{z - \ell_i}\right) + O(t^2) \text{ as } t \to 0,$$

we obtain the desired statement. □

In the remainder of this section we introduce some notation from the theory of hyperelliptic integrals. We will require the latter to formulate our main result in the next section.

Fix k simple positively oriented contours $\gamma_1, \ldots, \gamma_k$ such that each γ_j encloses the segment $[\hat{a}_j, \hat{b}_j]$ (and thus also $[r_j, s_j]$ from Assumption 5) for $j = 1, \ldots, k$. We assume that γ_j are pairwise disjoint and do not enclose each other.

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

$$\Omega : P(z) \to \left(\frac{1}{2\pi i} \oint\limits_{\gamma_1} \frac{P(z)dz}{\prod_{j=1}^{k} \sqrt{(z-r_j)(z-s_j)}}, \cdots, \frac{1}{2\pi i} \oint\limits_{\gamma_k} \frac{P(z)dz}{\prod_{j=1}^{k} \sqrt{(z-r_j)(z-s_j)}} \right). \tag{5.7}$$

Notice that the sum of the integrals in (5.7) equals (minus) the residue of $\frac{P(z)dz}{\prod_{j=1}^{k} \sqrt{(z-r_j)(z-s_j)}}$ at infinity, which is zero. Therefore, Ω defines a linear map between $(k-1)$-dimensional vector spaces. The map Ω is rather complicated, but it is known to be an isomorphism of vector spaces for $k \geq 2$ (see [29, Section 2.1]).

Using Ω we can now define a different map Υ_z as follows. The map Υ_z is defined in terms of the k contours γ_j and the points r_j, s_j for $j = 1, \ldots, k$. It is a linear map on the space of continuous functions $f(z)$ on $\gamma = \cup_{j=1}^{k} \gamma_j$, whose integral over γ is zero and is given by

$$\Upsilon_z[f] = f(z) + \frac{P(z)}{\prod_{j=1}^{k} \sqrt{(z-r_j)(z-s_j)}}, \tag{5.8}$$

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

$$\oint\limits_{\gamma_j} \Upsilon_z[f] dz = 0.$$

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

$$P = \Omega^{-1} \left(-\frac{1}{2\pi i} \oint\limits_{\gamma_1} f(z)dz, \cdots, -\frac{1}{2\pi i} \oint\limits_{\gamma_k} f(z)dz \right). \tag{5.9}$$

We emphasize that the map $f \to \Upsilon_z[f]$ does not depend on $\mathbf{t}, \mathbf{v}$.

We will require several properties of Υ_z, which can easily be deduced from the above definitions. We summarize them in the following proposition without proof.

Proposition 5.1. *The function Υ_z satisfies the following properties:*

(1) it is Lipschitz continuous in the uniform norm on the contours γ_j, $j = 1, \ldots, k$;
(2) if $\tilde{P}(z)$ is a polynomial of degree at most $k-2$ then $\Upsilon_z \left[\frac{\tilde{P}(z)}{\prod_{j=1}^{k} \sqrt{(z-r_j)(z-s_j)}} \right] = 0$;
(3) if Υ_z^N is defined in terms of $r_j(N), s_j(N)$ and $r_j(N) - r_j = O(N^{-1}\log(N)) = s_j(N) - s_j$ for $j = 1, \ldots, k$ then for any f we have $\Upsilon_z^N[f] - \Upsilon_z[f] = O(N^{-1}\log(N))$.

5.2. Main result

At this time we isolate the main technical result we prove about $\Delta G_N(z)$ and deduce a couple of easy corollaries from it. We continue with the notation from the previous section.

Theorem 5.2.1. *Suppose Assumptions 1–6 in Section 2.2 hold. Let* $U := \mathbb{C} \setminus \{\cup_{j=1}^{k} [\hat{a}_j, \hat{b}_j] \cup \{\pm 2\sqrt{\mathtt{u}}\}\}$ *and* $\Gamma = \cup_{j=1}^{k} \gamma_j \subset U$*, where each* γ_j *is a positively oriented contour that encloses the segment* $[\hat{a}_j, \hat{b}_j]$ *for* $j = 1, \ldots, k$*,* γ_j *are pairwise disjoint and do not enclose each other. We set* U^u *to be the single unbounded component of* $U \setminus \Gamma$*.*

Fix $m \in \mathbb{N}$ *and* $v_0, \ldots, v_m \in U^u$*. For* $m \geq 2$ *we have*

$$\frac{\partial^m}{\partial t_1 \cdots \partial t_m} \mathbb{E}_{\mathbb{P}_N^{\mathbf{t},\mathbf{v}}} \left[\Delta G_N(v_0) \right] \Big|_{t_a = 0, 1 \leq a \leq m} = O(N^{-1} \log(N)), \tag{5.10}$$

while for $m = 1$

$$\begin{aligned} &\frac{\partial}{\partial t_1} \mathbb{E}_{\mathbb{P}_N^{\mathbf{t},\mathbf{v}}} \left[\Delta G_N(v_0) \right] \Big|_{t_1 = 0} = O(N^{-1} \log(N)) + \\ &\Upsilon_{v_0} \left[\frac{1}{4\theta\pi i \cdot \prod_{j=1}^{k} \sqrt{(v_0 - \hat{r}_j)(v_0 - \hat{s}_j)}} \oint_{\Gamma} \frac{\prod_{j=1}^{k} \sqrt{(z - \hat{r}_j)(z - \hat{s}_j)}}{(z - v_1)^2 (z - v_0)} dz. \right] \end{aligned} \tag{5.11}$$

In the above Υ_{v_0} *is as in (5.8) for the contours* γ_j *and the points* $\hat{r}_j, \hat{s}_j$ *for* $j = 1, \ldots, k$ *as in Assumption 5. Finally, the constants in the big O notation are uniform as* $v_0, v_1, \ldots, v_k$ *vary over compact subsets of* U^u*.*

Remark 5.2.2. We will prove Theorem 5.2.1 for the case $\mathtt{u} > 0$. The case $\mathtt{u} = 0$ can be handled with minor modifications of the argument.

Theorem 5.2.3. *Assume the same notation as in Theorem 5.2.1. As* $N \to \infty$*, the random field* $N(G_N^d(z) - \mathbb{E}_{\mathbb{P}_N}\left[G_N^d(z)\right], z \in U$*, converges (in the sense of joint moments, uniformly in* z *in compact subsets of* U*) to a centered complex Gaussian random field with second moment*

$$\lim_{N \to \infty} N^2 \left(\mathbb{E}_{\mathbb{P}_N}\left[G_N^d(z_1) G_N^d(z_2)\right] - \mathbb{E}_{\mathbb{P}_N}\left[G_N^d(z_1)\right] \mathbb{E}_{\mathbb{P}_N}\left[G_N^d(z_2)\right] \right) =: \mathcal{C}_\theta(z_1, z_2), \text{ where} \tag{5.12}$$

$$\begin{aligned} \mathcal{C}_\theta(z_1, z_2) = \theta^{-1} \cdot \Upsilon_{z_2} \Bigg[&- \frac{1}{2(z_1 - z_2)^2} + \prod_{j=1}^{k} \frac{\sqrt{(z_1 - \hat{r}_j)(z_1 - \hat{s}_j)}}{\sqrt{(z_2 - \hat{r}_j)(z_2 - \hat{s}_j)}} \times \\ &\left(\frac{1}{(z_1 - z_2)^2} - \frac{1}{2(z_1 - z_2)} \sum_{j=1}^{k} \left(\frac{1}{z_1 - \hat{r}_j} + \frac{1}{z_1 - \hat{s}_j} \right) \right) \Bigg]. \end{aligned} \tag{5.13}$$

Remark 5.2.4. Since $\overline{G_N(z)} = G_N(\overline{z})$, we can use (5.13) to completely characterize the asymptotic covariance of the (recentered) random field $G_N(z)$.

Remark 5.2.5. When $m = 1$ the covariance $\mathcal{C}_\theta(z_1, z_2)$ can be written down explicitly as

$$\mathcal{C}_\theta(z_1, z_2) = -\frac{\theta^{-1}}{2(z_1 - z_2)^2}\left(1 - \frac{(z_1 - \hat{r}_1)(z_2 - \hat{s}_1) + (z_2 - \hat{r}_1)(z_1 - \hat{s}_1)}{2\sqrt{(z_1 - \hat{r}_1)(z_1 - \hat{s}_1)}\sqrt{(z_2 - \hat{r}_1)(z_2 - \hat{s}_1)}}\right).$$

When $m = 2$ one can also find an explicit form for $\mathcal{C}_\theta(z_1, z_2)$, involving the values of complete elliptic integrals, but we do not pursue it here, cf. [7].

Remark 5.2.6. In the continuous log-gas setting the covariance has the same form as (5.13), cf. [17,39,64]. A similar result also holds for the discrete β-ensembles in [15].

Proof. Fix $v_0, \dots, v_m \in U$ and Γ as in Theorem 5.2.1 so that $v_0, \dots, v_m \in U^u$. Setting $\xi = \Delta G_N(v_0)$ in Lemma 5.1.1 we know that the joint cumulant of $\Delta G_N(v_0), NG_N^d(v_1), \dots, NG_N^d(v_m)$ is given by

$$\frac{\partial^m}{\partial t_1 \partial t_2 \cdots \partial t_m} \mathbb{E}_{\mathbb{P}_N^{\mathbf{t},\mathbf{v}}}\left[\Delta G_N(v_0)\right]\Big|_{t_a = 0, 1 \le a \le m}.$$

Since cumulants remain unchanged under constant shifts, we see that the above formula is also the joint cumulant of $N(G_N^d(v_i) - \mathbb{E}_{\mathbb{P}_N}\left[G_N^d(v_i)\right])$ for $i = 0, \dots, m$. From Theorem 5.2.1 we see that as $N \to \infty$ all 3rd and higher order cumulants vanish, which proves the asymptotic Gaussianity of the field $N(G_N^d(z) - \mathbb{E}_{\mathbb{P}_N}\left[G_N^d(z)\right])$.

As $N(G_N^d(z) - \mathbb{E}_{\mathbb{P}_N}\left[G_N^d(z)\right])$ are centered for each N so is the limiting field. From (5.11) we also have the following formula for the limiting covariance (which is the second joint cumulant) of $N(G_N^d(z_1) - \mathbb{E}_{\mathbb{P}_N}\left[G_N^d(z_1)\right])$ and $N(G_N^d(z_2) - \mathbb{E}_{\mathbb{P}_N}\left[G_N^d(z_2)\right])$ for $z_1, z_2 \in U^u$

$$\Upsilon_{z_2}\left[\frac{1}{4\theta\pi i \cdot \prod_{j=1}^k \sqrt{(z_2 - \hat{r}_j)(z_2 - \hat{s}_j)}} \oint_\Gamma \frac{\prod_{j=1}^k \sqrt{(z - \hat{r}_j)(z - \hat{s}_j)}}{(z - z_1)^2(z - z_2)} dz\right].$$

Evaluating and adding the (minus) residues at $z = z_1$ and $z = z_2$ we obtain (5.13). □

Theorem 5.2.7. *Assume the same notation as in Theorem 5.2.1. For $m \ge 1$ let $f_1, \dots, f_m$ be real analytic functions in U and define*

$$\mathcal{L}_{f_i} = N\int_{\mathbb{R}} f_j(x)\mu_N(dx) - N\mathbb{E}_{\mathbb{P}_N}\left[\int_{\mathbb{R}} f_j(x)\mu_N(dx)\right] \text{ for } i = 1, \dots, m.$$

Then the random variables $\mathcal{L}_{f_i}$ converge jointly in the sense of moments to an m-dimensional centered Gaussian vector $X = (X_1, \dots, X_m)$ with covariance

$$Cov(X_i, X_j) = \frac{1}{(2\pi i)^2} \oint_{\Gamma} \oint_{\Gamma} f_i(s) f_j(t) \mathcal{C}_\theta(s,t) ds dt \text{ with } C_\theta(s,t) \text{ as in (5.13).}$$

Proof. Observe that when f is real analytic in U we have for all large N

$$\mathcal{L}_f = \frac{N}{2\pi i} \oint_{\Gamma} f(z)(G_N(z) - \mathbb{E}_{\mathbb{P}_N}[G_N(z)])dz,$$

where Γ is as in Theorem 5.2.1. Therefore, for any joint moment of $\mathcal{L}_{f_i}$ we have

$$\mathbb{E}_{\mathbb{P}_N}\left[\mathcal{L}_{f_{i_1}} \cdots \mathcal{L}_{f_{i_k}}\right] = \frac{1}{(2\pi i)^k} \oint_{\Gamma} \cdots \oint_{\Gamma} \mathbb{E}_{\mathbb{P}_N}\left[\prod_{m=1}^{k} N(G_N(z_m) - \mathbb{E}_{\mathbb{P}_N}[G_N(z_m)])\right] \prod_{m=1}^{k} f_{i_m}(z_m) dz_m. \quad (5.14)$$

Since cumulants of centered random variables are linear combinations of moments and vice versa, we conclude that all third and higher order cumulants of $\mathcal{L}_{f_i}$ vanish as $N \to \infty$ (here we used Theorem 5.2.3, which implies the third and higher order joint cumulants of $N(G_N(z_i) - \mathbb{E}_{\mathbb{P}_N}[G_N(z_i)])$ vanish uniformly when $z_i \in \Gamma$). This proves the Gaussianity of the limiting vector X. Since $\mathcal{L}_{f_i}$ are centered for each N the same is true for X. To get $Cov(X_i, X_j)$ we can set $k = 2$, $i_1 = i$ and $i_2 = j$ in (5.14) and send $N \to \infty$. In view of (5.12) we conclude that

$$Cov(X_i, X_j) = \frac{1}{(2\pi i)^2} \oint_{\Gamma} \oint_{\Gamma} f_i(s) f_j(t) \mathcal{C}_\theta(s,t) ds dt, \text{ where } \mathcal{C}_\theta(s,t) \text{ is as in (5.13).} \quad \square$$

5.3. Application of Nekrasov's equation

In this section we begin the proof of Theorem 5.2.1 emphasizing the contribution of the Nekrasov's equation. In what follows we use the same notation as in the previous section and Section 2.2, dropping the dependence on N from parameters.

The first key observation we make is that $\mathbb{P}_N^{\mathbf{t},\mathbf{v}}$ satisfies Nekrasov's equation with

$$\begin{aligned} \Phi_N^{+,\mathbf{t},\mathbf{v}}(z) &= \Phi_N^{+}(z) \prod_{a=1}^{m} \left[(v_a + t_a - \sigma_N(z))(v_a - \sigma_N(qz))\right], \\ \Phi_N^{-,\mathbf{t},\mathbf{v}}(z) &= \Phi_N^{-}(z) \prod_{a=1}^{m} \left[(v_a + t_a - \sigma_N(qz))(v_a - \sigma_N(z))\right], \text{ where} \end{aligned} \quad (5.15)$$

$\Phi_N^{\pm}$ are as in Assumption 4 and we recall that $\sigma_N(z) = z + uz^{-1}$. Notice that $\Phi_N^{\pm,\mathbf{t},\mathbf{v}}(z)$ are also analytic in $\mathcal{M}$. Denoting the RHS of the Nekrasov's equation for the measure $\mathbb{P}_N^{\mathbf{t},\mathbf{v}}$ by $R_N^{\mathbf{t},\mathbf{v}}(z)$ we see from Theorem 4.1.1 that $R_N^{\mathbf{t},\mathbf{v}}(z)$ is analytic in $\mathcal{M}$.

Using $q = \mathfrak{q}^{1/N} + O(N^{-2})$ and $u = \mathfrak{u} + O(N^{-1})$ we obtain asymptotic expansions for $\Phi_N^{\pm,\mathbf{t},\mathbf{v}}$

$$\begin{aligned}
\Phi_N^{\pm,\mathbf{t},\mathbf{v}}(z) &= [\Phi^{\pm}(z) + \Psi_N^{\pm}(z)] \prod_{a=1}^{m} [(v_a + t_a - \sigma(z))(v_a - \sigma(z))] + \mathrm{Rem}_{1,N}^{\pm,\mathbf{t},\mathbf{v}}(z),\\
\Psi_N^{+}(z) &:= \varphi_N^{+}(z) - \Phi^{+}(z) \cdot \sum_{a=1}^{m} \left[\frac{(\log \mathfrak{q}/N)\left(z - \mathfrak{u}z^{-1}\right) - (\mathfrak{u}-u)z^{-1}}{v_a - \sigma(z)} - \frac{(\mathfrak{u}-u)z^{-1}}{v_a + t_a - \sigma(z)} \right],\\
\Psi_N^{-}(z) &:= \varphi_N^{-}(z) - \Phi^{-}(z) \cdot \sum_{a=1}^{m} \left[\frac{(\log \mathfrak{q}/N)\left(z - \mathfrak{u}z^{-1}\right) - (\mathfrak{u}-u)z^{-1}}{v_a + t_a - \sigma(z)} - \frac{(\mathfrak{u}-u)z^{-1}}{v_a - \sigma(z)} \right],
\end{aligned} \tag{5.16}$$

where $\sigma(z) = z + \mathfrak{u}z^{-1}$ and $\mathrm{Rem}_{1,N}^{\pm,\mathbf{t},\mathbf{v}}(z) = O(N^{-2})$ uniformly over compact subsets of $\mathcal{M}$.

The second important observation is that we have the following asymptotic expansion

$$\begin{aligned}
\prod_{i=1}^{N} \frac{\sigma_N(q^{\theta}z) - \ell_i}{\sigma_N(z) - \ell_i} &= \exp\left[\theta \mathfrak{G}_N^{c}(z) + \theta \Delta\mathfrak{G}_N(z) + W_N^{-}(z) + \mathrm{Rem}_{2,N}^{-}(z)\right],\\
\prod_{i=1}^{N} \frac{\sigma_N(q^{1-\theta}z) - \ell_i}{\sigma_N(qz) - \ell_i} &= \exp\left[-\theta \mathfrak{G}_N^{c}(z) - \theta \Delta\mathfrak{G}_N(z) + W_N^{+}(z) + \mathrm{Rem}_{2,N}^{+}(z)\right],
\end{aligned} \tag{5.17}$$

$$\begin{aligned}
&\text{where } \mathfrak{G}_N^{d/c}(z) = N \log q \cdot (z - \mathfrak{u}z^{-1}) \cdot G_N^{d/c}(z + \mathfrak{u}z^{-1}),\\
&\qquad \Delta\mathfrak{G}_N(z) = \mathfrak{G}_N^{d}(z) - \mathfrak{G}_N^{c}(z) \text{ with } G_N^{c} \text{ and } G_N^{d} \text{ as in (5.3)};
\end{aligned} \tag{5.18}$$

$$\begin{aligned}
W_N^{-}(z) &= \theta \log q \Big[z(\theta/2) \cdot \partial_z \mathfrak{G}_N^{d}(z) + (u - \mathfrak{u})N\\
&\qquad \cdot \left[\partial_z G_N^{c}(z + \mathfrak{u}z^{-1}) - z^{-1} G_N^{c}(z + \mathfrak{u}z^{-1}) \right] \Big],\\
W_N^{+}(z) &= \theta \log q \Big[z(\theta/2 - 1) \partial_z \mathfrak{G}_N^{d}(z)\\
&\qquad - (u - \mathfrak{u})N \left[\partial_z G_N^{c}(z + \mathfrak{u}z^{-1}) - z^{-1} G_N^{c}(z + \mathfrak{u}z^{-1}) \right] \Big].
\end{aligned} \tag{5.19}$$

The remainders $\mathrm{Rem}_{2,N}^{\pm}(z) = O(N^{-2})$ are uniform in z on compact subsets of $\mathcal{O}$, which is the inverse image of U under the map $z \to z + \mathfrak{u}z^{-1}$. Explicitly, if $\hat{c}_j, \hat{d}_j$ are the points in $[1, \mathfrak{q}^{-\mathtt{M}}]$ with $\hat{c}_j + \mathfrak{u}\hat{c}_j^{-1} = \hat{a}_j$ and $\hat{d}_j + \mathfrak{u}\hat{d}_j^{-1} = \hat{b}_j$ then $\mathcal{O} := \mathbb{C} \setminus \left\{ \{0, \sqrt{\mathfrak{u}}, -\sqrt{\mathfrak{u}}\} \cup \cup_{j=1}^{k} [\hat{c}_j, \hat{d}_j] \cup \cup_{j=1}^{k} [\mathfrak{u}\hat{d}_j^{-1}, \mathfrak{u}\hat{c}_j^{-1}] \right\}$.

The third observation we require comes from Assumption 5 and Lemma 4.2.1, which imply:

$$\begin{aligned}&\text{If } R_N(z) := \Phi^-(z)e^{\theta\mathfrak{G}^c_N(z)} + \Phi^+(z)e^{-\theta\mathfrak{G}^c_N(z)},\\ &Q_N(z) := \Phi^-(z)e^{\theta\mathfrak{G}^c_N(z)} - \Phi^+(z)e^{-\theta\mathfrak{G}^c_N(z)},\\ &\text{then } R_N \text{ is analytic in } \mathcal{M} \text{ and } Q_N(z) = H_N(z)\cdot\prod_{j=1}^{k}\sqrt{(\sigma(z)-r_j)\,(\sigma(z)-s_j)},\end{aligned} \tag{5.20}$$

where H_N is analytic and non-vanishing in $\mathcal{M}$.

We detail the consequence of the above three observation. Nekrasov's equation for $\mathbb{P}_N^{\mathbf{t},\mathbf{v}}$ reads

$$R_N^{\mathbf{t},\mathbf{v}}(z) = \Phi_N^{-,\mathbf{t},\mathbf{v}}(z)\cdot\mathbb{E}_{\mathbb{P}_N^{\mathbf{t},\mathbf{v}}}\left[\prod_{i=1}^{N}\frac{\sigma_N(q^\theta z)-\ell_i}{\sigma_N(z)-\ell_i}\right] + \Phi_N^{+,\mathbf{t},\mathbf{v}}(z)\cdot\mathbb{E}_{\mathbb{P}_N^{\mathbf{t},\mathbf{v}}}\left[\prod_{i=1}^{N}\frac{\sigma_N(q^{1-\theta} z)-\ell_i}{\sigma_N(qz)-\ell_i}\right].$$

Combining the above with (5.16), (5.17) and (5.20) we conclude that

$$\begin{aligned}R_N^{\mathbf{t},\mathbf{v}}(z) &= \mathrm{Rem}_N^{\mathbf{t},\mathbf{v}}(z) + \prod_{a=1}^{m}[(v_a+t_a-\sigma(z))(v_a-\sigma(z))]\times\Big[\mathbb{E}_{\mathbb{P}_N^{\mathbf{t},\mathbf{v}}}\left[\theta\Delta\mathfrak{G}_N(z)\right]\cdot Q_N(z)\\ &\quad + \Psi_N^-(z)e^{\theta\mathfrak{G}^c_N(z)} + \Psi_N^+(z)e^{-\theta\mathfrak{G}^c_N(z)} + \tilde{W}_N(z) + R_N(z)\Big],\ \text{where}\end{aligned} \tag{5.21}$$

$$\begin{aligned}&\tilde{W}_N(z) = \Phi^-(z)e^{\theta\mathfrak{G}^c_N(z)}\tilde{W}_N^-(z) + \Phi^+(z)e^{-\theta\mathfrak{G}^c_N(z)}\tilde{W}_N^+(z),\ \text{with}\\ &\tilde{W}_N^-(z) = \theta\log q\left[z(\theta/2)\partial_z\mathfrak{G}^c_N(z) + (u-\mathfrak{u})N\cdot\left[\partial_z G^c_N(z+\mathfrak{u}z^{-1}) - z^{-1}G^c_N(z+\mathfrak{u}z^{-1})\right]\right],\\ &\tilde{W}_N^+(z) = \theta\log q\Big[z(\theta/2-1)\partial_z\mathfrak{G}^c_N(z) -\\ &(u-\mathfrak{u})N\left[\partial_z G^c_N(z+\mathfrak{u}z^{-1}) - z^{-1}G^c_N(z+\mathfrak{u}z^{-1})\right]\Big],\\ &\mathrm{Rem}_N^{\mathbf{t},\mathbf{v}}(z) = \mathrm{Rem}_{1,N}^{+,\mathbf{t},\mathbf{v}}(z)\mathbb{E}_{\mathbb{P}_N^{\mathbf{t},\mathbf{v}}}\left[\prod_{i=1}^{N}\frac{\sigma(q^\theta z)-\ell_i}{\sigma(z)-\ell_i}\right] +\\ &\mathrm{Rem}_{1,N}^{-,\mathbf{t},\mathbf{v}}(z)\mathbb{E}_{\mathbb{P}_N^{\mathbf{t},\mathbf{v}}}\left[\frac{\sigma_N(q^{1-\theta}z)-\ell_i}{\sigma_N(qz)-\ell_i}\right] +\\ &\prod_{a=1}^{m}[(v_a+t_a-\sigma(z))(v_a-\sigma(z))]\times[A_N+B_N+C_N+D_N]\ \text{and}\end{aligned} \tag{5.22}$$

$$\begin{aligned}A_N &= \Phi^-(z)\cdot\mathbb{E}_{\mathbb{P}_N^{\mathbf{t},\mathbf{v}}}\left[\prod_{i=1}^{N}\frac{\sigma_N(q^\theta z)-\ell_i}{\sigma_N(z)-\ell_i} - e^{\theta\Delta\mathfrak{G}_N(z)+\theta\mathfrak{G}^c_N(z)} - e^{\theta\mathfrak{G}^d_N(z)}\cdot\tilde{W}_N^-(z)\right] +\\ &\quad \Phi^+(z)\cdot\mathbb{E}_{\mathbb{P}_N^{\mathbf{t},\mathbf{v}}}\left[\frac{\sigma_N(q^{1-\theta}z)-\ell_i}{\sigma_N(qz)-\ell_i} - e^{-\theta\Delta\mathfrak{G}_N(z)-\theta\mathfrak{G}^c_N(z)} - e^{-\theta\mathfrak{G}^d_N(z)}\cdot\tilde{W}_N^+(z)\right],\\ B_N &= \Phi^-(z)\tilde{W}_N^-(z)\cdot\mathbb{E}_{\mathbb{P}_N^{\mathbf{t},\mathbf{v}}}\left[e^{\theta\mathfrak{G}^d_N(z)} - e^{\theta\mathfrak{G}^c_N(z)}\right]\Phi^+(z)\tilde{W}_N^+(z)\\ &\quad\cdot\mathbb{E}_{\mathbb{P}_N^{\mathbf{t},\mathbf{v}}}\left[e^{-\theta\mathfrak{G}^d_N(z)} - e^{-\theta\mathfrak{G}^c_N(z)}\right],\end{aligned}$$

$$C_N = \Phi^-(z)\cdot e^{\theta\mathfrak{G}^c_N(z)}\cdot \mathbb{E}_{\mathbb{P}_N^{\mathbf{t},\mathbf{v}}}\left[e^{\theta\Delta\mathfrak{G}_N(z)} - \theta\Delta\mathfrak{G}_N(z) - 1\right] +$$
$$\Phi^+(z)\cdot e^{-\theta\mathfrak{G}^c_N(z)}\cdot \mathbb{E}_{\mathbb{P}_N^{\mathbf{t},\mathbf{v}}}\left[e^{-\theta\Delta\mathfrak{G}_N(z)} + \theta\Delta\mathfrak{G}_N(z) - 1\right],$$

$$D_N = \Psi_N^-(z)\cdot \mathbb{E}_{\mathbb{P}_N^{\mathbf{t},\mathbf{v}}}\left[\frac{\sigma_N(q^\theta z)-\ell_i}{\sigma_N(z)-\ell_i} - e^{\theta\mathfrak{G}^c_N(z)}\right]$$
$$+\Psi_N^+(z)\cdot \mathbb{E}_{\mathbb{P}_N^{\mathbf{t},\mathbf{v}}}\left[\prod_{i=1}^N \frac{\sigma_N(q^{1-\theta}z)-\ell_i}{\sigma_N(qz)-\ell_i} - e^{-\theta\mathfrak{G}^c_N(z)}\right].$$

Let $\gamma_1,\dots,\gamma_k$ be as in the statement of the theorem and $v_0,\dots,v_m$ lie outside of $\Gamma = \cup_{j=1}^k \gamma_j$. We let γ_j', γ_j'' for $j=1,\dots,k$ be a positively oriented contours such that for each $j=1,\dots,k$

- γ_j' encloses the interval $[\hat{c}_j, \hat{d}_j]$ and excludes the points $\pm\sqrt{\mathfrak{u}}$ and 0,
- $\sigma(\gamma_j')$ is contained in the bounded component of $\mathbb{C}\setminus\gamma_j$,
- $\gamma_j'' := \omega(\gamma_j')$, where $\omega(z) = \frac{\mathfrak{u}}{z}$,
- $\{\gamma_j', \gamma_j''\}_{j=1}^k$ are all disjoint and contained in $\mathcal{M}$.

The existence of such contours is ensured by our assumptions on γ_j. Observe that by construction γ_j'' is a positively oriented contour that also excludes the points $\pm\sqrt{\mathfrak{u}}$ and 0, and encloses the interval $[\mathfrak{u}\hat{d}_j^{-1}, \mathfrak{u}\hat{c}_j^{-1}]$. For convenience we let $\Gamma_1 = \cup_{j=1}^k \gamma_j'$ and $\Gamma_2 = \cup_{j=1}^k \gamma_j''$.

We divide both sides of (5.21) by

$$2\pi i\cdot z\cdot(v_0-\sigma(z))\cdot\prod_{a=1}^m[(v_a+t_a-\sigma(z))(v_a-\sigma(z))]\cdot H_N(z)$$

and integrate over $\Gamma' := \Gamma_1\cup\Gamma_2$. Note that $R_N^{\mathbf{t},\mathbf{v}}(z)$ and $R_N(z)$ are both holomorphic inside the contours Γ_1,Γ_2 and so the integrals of the corresponding terms vanish. From the rest we get

$$\begin{aligned}
&\frac{1}{2\pi i}\oint_{\Gamma'}\frac{H_N(z)^{-1}z^{-1}Q_N(z)}{v_0-\sigma(z)}\mathbb{E}_{\mathbb{P}_N^{\mathbf{t},\mathbf{v}}}\left[\theta\Delta\mathfrak{G}_N(z)\right]dz = \\
&\frac{-1}{2\pi i}\oint_{\Gamma'}\frac{H_N(z)^{-1}z^{-1}dz}{v_0-\sigma(z)}\left[\frac{\mathrm{Rem}_N^{\mathbf{t},\mathbf{v}}(z)}{\prod_{a=1}^m[(v_a+t_a-\sigma(z))(v_a-\sigma(z))]}\right] \qquad (5.23)\\
&+\frac{-1}{2\pi i}\oint_{\Gamma'}\frac{H_N(z)^{-1}z^{-1}dz}{v_0-\sigma(z)}\left[\Psi_N^-(z)e^{\theta\mathfrak{G}^c_N(z)}+\Psi_N^+(z)e^{-\theta\mathfrak{G}^c_N(z)}+\tilde{W}_N(z)\right]
\end{aligned}$$

Equation (5.23) can be viewed as the main output of our application of the Nekrasov's equation. In the following section we use it to deduce Theorem 5.2.1.

5.4. Concluding the proof of Theorem 5.2.1

In this section we present the remainder of the proof of Theorem 5.2.1. Our arguments below will require a certain asymptotic bound – see (5.26), which will be established in Section 6. For clarity we split the proof into several steps.

Step 1. Our goal in this step is to rewrite (5.23) into a form that is more useful for our analysis.

Using the formula for Q_N from (5.20) and that $\Delta\mathfrak{G}_N(z) = \log q \cdot (z - \mathtt{u}z^{-1})\Delta G_N(\sigma(z))$ from (5.18) we see that the RHS of (5.23) equals

$$\frac{\log q}{2\pi i} \oint_{\Gamma'} \frac{\prod_{j=1}^{k} \sqrt{(\sigma(z) - r_j)(\sigma(z) - s_j)}}{v_0 - \sigma(z)} \cdot (1 - \mathtt{u}z^{-2}) \cdot \mathbb{E}_{\mathbb{P}_N^{\mathbf{t},\mathbf{v}}} [\theta \Delta G_N(\sigma(z))] \, dz$$

We perform a change of variables $\sigma(z) = w$ to rewrite the RHS of (5.23) as

$$2 \cdot \frac{\log q}{2\pi i} \oint_{\Gamma} \frac{\prod_{j=1}^{k} \sqrt{(w - r_j)(w - s_j)}}{v_0 - w} \cdot \mathbb{E}_{\mathbb{P}_N^{\mathbf{t},\mathbf{v}}} [\theta \Delta G_N(w)] \, dw,$$

where we used that $\sigma(\gamma_j') = \sigma(\gamma_j'')$ are contained γ_j and we can deform the image to the latter without affecting the value of the integral by Cauchy's theorem.

Note that $\mathbb{E}_{\mathbb{P}_N^{\mathbf{t},\mathbf{v}}} [\Delta G_N(w)]$ is analytic outside of the contour of integration and decays like $1/w^2$ when $|w| \to \infty$. Therefore, we can compute the integral as (minus) the residues at $w = v_0$ and $z = \infty$. The residue at v_0 is given by

$$-2 \cdot \log q \cdot \prod_{j=1}^{k} \sqrt{(v_0 - r_j)(v_0 - s_j)} \cdot \mathbb{E}_{\mathbb{P}_N^{\mathbf{t},\mathbf{v}}} [\theta \Delta G_N(v_0)],$$

while the residue at ∞ is a polynomial $P_N^{\mathbf{t},\mathbf{v}}(v_0)$ of degree at most $k-2$ in v_0, whose coefficients are rational functions in $\mathbf{t}, \mathbf{v}$. Substituting the above in (5.23) we obtain the formula

$$\begin{aligned} \mathbb{E}_{\mathbb{P}_N^{\mathbf{t},\mathbf{v}}} [\theta \Delta G_N(v_0)] &= \frac{N P_N^{\mathbf{t},\mathbf{v}}(v_0)}{D_N(v_0)} + \frac{N}{2\pi i} \oint_{\Gamma'} \frac{-D_N(v_0)^{-1} dz}{z H_N(z)(v_0 - \sigma(z))} \left[\frac{\mathrm{Rem}_N^{\mathbf{t},\mathbf{v}}(z)}{W_m^{\mathbf{t},\mathbf{v}}(z)} \right] + \\ &\frac{N}{2\pi i} \oint_{\Gamma'} \frac{-D_N(v_0)^{-1} dz}{z H_N(z)(v_0 - \sigma(z))} \left[\Psi_N^-(z) e^{\theta \mathfrak{G}_N^c(z)} + \Psi_N^+(z) e^{-\theta \mathfrak{G}_N^c(z)} + \tilde{W}_N(z) \right], \end{aligned} \tag{5.24}$$

where

$$D_N(v) = 2N \log q \prod_{j=1}^{k} \sqrt{(v - r_j)(v - s_j)} \text{ and}$$
$$W_m^{\mathbf{t},\mathbf{v}}(z) = \prod_{a=1}^{m} [(v_a + t_a - \sigma(z))(v_a - \sigma(z))]. \tag{5.25}$$

Step 2. In this step we isolate an asymptotic estimate that we require to finish the proof.

Fact 5.4.1. For each $m \geq 0$ we have

$$\frac{\partial^m}{\partial t_1 \cdots \partial t_m} \frac{1}{2\pi i} \oint_{\Gamma'} \frac{D_N(v_0)^{-1} dz}{z H_N(z)(v_0 - \sigma(z))} \left[\frac{\mathrm{Rem}_N^{\mathbf{t},\mathbf{v}}(z)}{W_m^{\mathbf{t},\mathbf{v}}(z)} \right] \Bigg|_{t_a = 0, 1 \leq a \leq m} = O(N^{-2}), \tag{5.26}$$

where the constant in the big O notation is uniform as $v_0, v_1, \ldots, v_k$ vary over compacts in U^u.

The proof of Fact 5.4.1 will be presented in Section 6. In the remainder of the section we assume its validity and finish the proof of Theorem 5.2.1.

Step 3. Let us fix $m \geq 1$, differentiate both sides of (5.24) with respect to $t_1, \ldots, t_m$ and set $t_a = 0$ for $a = 1, \ldots, m$. Using (5.26) we get

$$\begin{aligned} &\partial_{t_1} \cdots \partial_{t_m} \mathbb{E}_{\mathbb{P}_N^{\mathbf{t},\mathbf{v}}} [\theta \Delta G_N(v_0)] \Big|_{t_a = 0, 1 \leq a \leq m} = \\ &\partial_{t_1} \cdots \partial_{t_m} \left[\frac{N P_N^{\mathbf{t},\mathbf{v}}(v)}{D_N(v_0)} + \frac{1}{2\pi i} \oint_{\Gamma'} \frac{-D_N(v_0)^{-1} dz}{z H_N(z)(v_0 - \sigma(z))} \right. \\ &\left. \times \left[N \Psi_N^-(z) e^{\theta \mathfrak{G}_N^c(z)} + N \Psi_N^+(z) e^{-\theta \mathfrak{G}_N^c(z)} + \tilde{W}_N(z) \right] \right] \Bigg|_{t_a = 0, 1 \leq a \leq m} + O(N^{-1}). \end{aligned} \tag{5.27}$$

The only functions in (5.27), which depend on $\mathbf{t}$ are $P_N^{\mathbf{t},\mathbf{v}}(v), \Psi_N^{\pm}(z)$, see (5.16). Since any mixed partial derivatives of $\Psi_N^{\pm}(z)$ vanish, we conclude that for $m \geq 2$ we have

$$\begin{aligned} &\partial_{t_1} \cdots \partial_{t_m} \mathbb{E}_{\mathbb{P}_N^{\mathbf{t},\mathbf{v}}} [\theta \Delta G_N(v_0)] \Big|_{t_a = 0, 1 \leq a \leq m} \\ &= \partial_{t_1} \cdots \partial_{t_m} \left[\frac{N P_N^{\mathbf{t},\mathbf{v}}(v_0)}{D_N(v_0)} \right] \Bigg|_{t_a = 0, 1 \leq a \leq m} + O(N^{-1}). \end{aligned} \tag{5.28}$$

We may now apply $\Upsilon_{v_0}^N$ from (5.8) for the contours $\gamma_1, \ldots, \gamma_k$ and the points r_j, s_j for $j = 1, \ldots, k$ to both sides of the above equation. Indeed, we notice that the integral of G_N^d around γ_i is deterministic and equals $n_i(N)/N$. On the other hand, the integral of

G_N^c around γ_i equals the total mass of $\hat{\mu}_N(x)$ inside γ_i, which is $n_i(N)/N$ by assumption. We conclude that the integral of $\Delta G_N(v)$ around each loop γ_i vanishes. The integral over the first term on the right side vanishes by Property (2) in Proposition 5.1. By linearity, we see that the integral over the term represented by $O(N^{-1})$ over Γ must also vanish. Applying $\Upsilon_{v_0}^N$ and using Property (1) in Proposition 5.1 we get

$$\partial_{t_1} \cdots \partial_{t_m} \mathbb{E}_{\mathbb{P}_N^{\mathbf{t},\mathbf{v}}} \left[\Delta G_N(v_0) \right] \Big|_{t_a=0, 1 \le a \le m} = O(N^{-1}),$$

which proves the case $m \ge 2$.

If $m = 1$

$$\partial_{t_1} \mathbb{E}_{\mathbb{P}_N^{\mathbf{t},\mathbf{v}}} \left[\theta \Delta G_N(v_0) \right] \Big|_{t_1=0} = O(N^{-1}) + \partial_{t_1} \frac{N P_N^{\mathbf{t},\mathbf{v}}(v_0)}{D_N(v_0)} \Big|_{t_1=0} + \frac{-D_N(v_0)^{-1}}{2\pi i} \oint_{\Gamma'} f(z) dz,$$

where $f(z) :=$

$$\frac{\Phi^-(z) e^{\theta \mathfrak{G}_N^c(z)} \left[(1 - \mathfrak{u} z^{-2}) \log \mathfrak{q} - (\mathfrak{u} - u) z^{-2} \right] - \Phi^+(z) e^{-\theta \mathfrak{G}_N^c(z)} (\mathfrak{u} - u) z^{-2}}{H_N(z)(v_0 - \sigma(z))(v_1 - \sigma(z))^2}. \tag{5.29}$$

Applying (5.20) we obtain

$$\frac{-D_N(v_0)^{-1}}{2\pi i} \oint_{\Gamma'} f(z) dz =$$

$$\frac{-D_N(v_0)^{-1}}{2\pi i} \oint_{\Gamma'} \frac{2^{-1} \log \mathfrak{q} [R_N(z) + Q_N(z)](1 - \mathfrak{u} z^{-2}) - (\mathfrak{u} - u) z^{-2} R_N(z)}{H_N(z)(v_0 - \sigma(z))(v_1 - \sigma(z))^2} dz.$$

Notice that the terms with $R_N(z)$ integrate to 0 by analyticity, and so we may remove them. Substituting $Q_N(z)$ from (5.20) and $D_N(v_0)$ from (5.25) we get

$$\frac{-D_N(v_0)^{-1}}{2\pi i} \oint_{\Gamma'} f(z) dz = \frac{-1}{8\pi i \cdot \prod_{j=1}^k \sqrt{(v_0 - r_j)(v_0 - s_j)}} \times$$
$$\oint_{\Gamma'} \frac{\prod_{j=1}^k \sqrt{(\sigma(z) - r_j)(\sigma(z) - s_j)}}{(v_0 - \sigma(z))(v_1 - \sigma(z))^2} \cdot (1 - \mathfrak{u} z^{-2}) dz.$$

We perform the change of variables $w = \sigma(z)$ and deform the resulting contours to Γ to obtain

$$\partial_{t_1}\mathbb{E}_{\mathbb{P}_N^{\mathbf{t},\mathbf{v}}}\left[\theta\Delta G_N(v_0)\right]\Big|_{t_1=0} = O(N^{-1}) + \frac{1}{4\pi i\cdot\prod_{j=1}^{k}\sqrt{(v_0-r_j)(v_0-s_j)}}\times \oint_{\Gamma}\frac{\prod_{j=1}^{k}\sqrt{(w-r_j)(w-s_j)}}{(w-v_0)(v_1-w)^2}dw + \partial_{t_1}\frac{NP_N^{\mathbf{t},\mathbf{v}}(v_0)}{D_N(v_0)}\Big|_{t_1=0}. \tag{5.30}$$

As before we can apply $\Upsilon_{v_0}^N$ to both sides of the above equation. The only difference with respect to the $m\geq 2$ case is the second term on the right side. Notice that it is analytic in the unbounded component of Γ and decays like $|v_0|^{-k-1}$ as $|v_0|\to\infty$. Consequently, there is no residue at infinity and the integral over Γ is zero. Arguing as in the case $m\geq 2$ we get

$$\partial_{t_1}\mathbb{E}_{\mathbb{P}_N^{\mathbf{t},\mathbf{v}}}\left[\theta\Delta G_N(v_0)\right]\Big|_{t_1=0} = O(N^{-1}) + \Upsilon_{v_0}^N\left[\frac{1}{4\pi i\cdot\prod_{j=1}^{k}\sqrt{(v_0-r_j)(v_0-s_j)}}\oint_{\Gamma}\frac{\prod_{j=1}^{k}\sqrt{(w-r_j)(w-s_j)}}{(w-v_0)(v_1-w)^2}dw\right].$$

Finally, we can replace r_j, s_j with $\hat{r}_j, \hat{s}_j$ and $\Upsilon_{v_0}^N$ with Υ_{v_0}, which produces an error $O(N^{-1}\log(N))$ by Assumption 5 and Property (3) in Proposition 5.1.

6. Central limit theorem: Part II

In this section we will prove (5.26), which is the missing ingredient necessary to complete the proof of Theorem 5.2.1. In what follows we will continue to use the same notation as in Section 5. Before we go into the main argument we introduce a bit of additional notation and isolate a basic result, which will be used several times throughout.

If $X_1,\dots,X_n$ are bounded random variables, we denote by $M_c(X_1,\dots,X_n)$ their joint cumulant. From Lemma 5.1.1 we know that for any bounded random variable ξ we have

$$\begin{aligned}\frac{\partial^n\mathbb{E}_{\mathbb{P}_N^{\mathbf{t},\mathbf{v}}}[\xi]}{\partial t_1\cdots\partial t_n}\Big|_{t_a=0,1\leq a\leq n} &= M_c(\xi, NG_N(v_1),\dots,NG_N(v_n))\\ &= M_c(\xi,\Delta G_N(v_1),\dots,\Delta G_N(v_n)),\end{aligned} \tag{6.1}$$

where the second equality follows from the fact that cumulants are unchanged under shifts. To ease notation later in the text we set for a subset $A=\{a_1,\dots,a_k\}\subset\{1,\cdots,n\}$

$$M_N(\xi|v_a,A) := M_c(\xi,\Delta G_N(v_{a_1}),\dots,\Delta G_N(v_{a_k})).$$

6.1. Estimating the remainders

In this section we reduce (5.26) to the following statement, whose proof is given in Section 6.2.

Proposition 6.1.1. *Assume the notation from Theorem 5.2.1. If $l \geq 1$ and $v_1, \dots, v_l \in U$ then*

$$\mathbb{E}_{\mathbb{P}_N}\left[\prod_{a=1}^{l} |\Delta G_N(v_a)|\right] = O(1), \tag{6.2}$$

where the constant in the big O notation is uniform as $v_1, \dots, v_l$ vary in compact subsets of U.

We assume the validity of Proposition 6.1.1 and proceed with the proof of (5.26). Our goal is to prove that for $m \geq 0$ we have

$$\frac{\partial^m}{\partial t_1 \cdots \partial t_m}\left[\frac{N^2 \cdot \mathrm{Rem}_N^{\mathbf{t},\mathbf{v}}(z)}{W_m^{\mathbf{t},\mathbf{v}}(z)}\right]\Bigg|_{t_a=0, 1\leq a\leq m} = O(1) \tag{6.3}$$

uniformly as z and $v_1, \dots, v_m$ vary over Γ' and compacts in U^u respectively. This implies (5.26).

In view of (5.22) we have

$$\begin{aligned}\frac{N^2 \cdot \mathrm{Rem}_N^{\mathbf{t},\mathbf{v}}(z)}{W_m^{\mathbf{t},\mathbf{v}}(z)} = \mathbb{E}_{\mathbb{P}_N^{\mathbf{t},\mathbf{v}}}\left[\xi_N(z)\cdot \Delta G_N(\sigma(z))^2\right] + \mathbb{E}_{\mathbb{P}_N^{\mathbf{t},\mathbf{v}}}\left[\xi_N'(z)\cdot \partial_z \Delta G_N(\sigma(z))\right] + \\ \mathbb{E}_{\mathbb{P}_N^{\mathbf{t},\mathbf{v}}}\left[\mathfrak{c}_N(z;\mathbf{t},\mathbf{v})\right] + \mathbb{E}_{\mathbb{P}_N^{\mathbf{t},\mathbf{v}}}\left[\mathfrak{c}_N'(z;\mathbf{t},\mathbf{v})\cdot \Delta G_N(\sigma(z))\right].\end{aligned} \tag{6.4}$$

In (6.4) $\xi_N(z), \xi_N'(z)$ are random analytic functions in z, which do not depend on $\mathbf{t}, \mathbf{v}$ and that are $O(1)$ uniformly over compacts in $\mathcal{O} \cap \mathcal{M}$ and N, recall that $\mathcal{O}$ is the inverse image of U under the map $z \to z + \mathtt{u} z^{-1}$. In addition, $\mathfrak{c}_N(z;\mathbf{t},\mathbf{v}), \mathfrak{c}_N'(z;\mathbf{t},\mathbf{v})$ are linear combinations of random analytic function in z, independent of $\mathbf{t}, \mathbf{v}$, that are also $O(1)$ uniformly over compacts in $\mathcal{O} \cap \mathcal{M}$. The coefficients of these linear combination are infinitely differentiable functions in t_i, whose derivatives, evaluated at $t_1 = \cdots = t_m = 0$, are all uniformly bounded as $v_1, \dots, v_m$ vary over compacts in U, z varies over compacts in $\mathcal{O} \cap \mathcal{M}$ and $|\sigma(z) - v_i|$ for $i = 1, \dots, m$ are bounded away from 0.

We now fix $m \geq 1$ and differentiate both sides of (6.4) with respect to $t_1, \dots, t_m$ and set $t_1 = \cdots = t_m = 0$ (the case $m = 0$ will be treated separately). For the terms involving random variables we use (6.1) to rewrite the result as a cumulant. Observe that we need to apply Leibniz rule when we differentiate $\mathbb{E}_{\mathbb{P}_N^{\mathbf{t},\mathbf{v}}}\left[\mathfrak{c}_N(z;\mathbf{t},\mathbf{v})\right]$ or $\mathbb{E}_{\mathbb{P}_N^{\mathbf{t},\mathbf{v}}}\left[\mathfrak{c}_N'(z;\mathbf{t},\mathbf{v})\right]$; therefore, we will obtain a sum depending on how many times we differentiated one of the coefficients in $\mathfrak{c}_N(z;\mathbf{t},\mathbf{v})$ or $\mathfrak{c}_N'(z;\mathbf{t},\mathbf{v})$ and how many times the expectations $\mathbb{E}_{\mathbb{P}_N^{\mathbf{t},\mathbf{v}}}$. We obtain the following result

$$\frac{\partial^m}{\partial t_1 \cdots \partial t_m}\left[\frac{N^2 \cdot \mathrm{Rem}_N^{\mathbf{t},\mathbf{v}}(z)}{W_m^{\mathbf{t},\mathbf{v}}(z)}\right]\Bigg|_{t_a=0, 1\leq a\leq m} = M_N\left(\xi_N(z)\Delta G_N(\sigma(z))^2 | v_a, \{1, \dots, m\}\right) +$$

$$M_N\left(\xi_N'(z)\partial_z\Delta G_N(\sigma(z))|v_a,\{1,\dots,m\}\right)+\sum_{A\subset\{1,\dots,m\}}\Big[M_N\left(\partial_A\mathfrak{c}_N(z;\mathbf{0},\mathbf{v})|v_a,A^c\right)+$$
$$M_N(\partial_A\mathfrak{c}_N'(z;\mathbf{0},\mathbf{v})\cdot\Delta G_N(\sigma(z))|v_a,A^c)\Big]. \tag{6.5}$$

Using that cumulants are linear combinations of moments and Proposition 6.1.1, we conclude that each term in (6.5) is $O(1)$ uniformly as $v_1,\dots,v_m$ vary over compacts in U, z varies over compacts in $\mathcal{O}\cap\mathcal{M}$, $|\sigma(z)-v_i|$ for $i=1,\dots,m$ are bounded away from 0 and $N\to\infty$. One might be cautious about the term involving $\partial_z\Delta G_N(\sigma(z))$; however, by Cauchy's Theorem the uniform moment bound we have for $\Delta G_N(\sigma(z))$ implies one for its derivative.

Since $\Gamma'\subset\mathcal{O}\cap\mathcal{M}$ we conclude (6.3) for the case $m\geq 1$. If $m=0$, then (6.4) reads

$$N^2\cdot\mathrm{Rem}_N(z)=\mathbb{E}_{\mathbb{P}_N}\left[\xi_N(z)\cdot\Delta G_N(\sigma(z))^2\right]+\mathbb{E}_{\mathbb{P}_N}\left[\xi_N'(z)\cdot\partial_z\Delta G_N(\sigma(z))\right]+$$
$$\mathbb{E}_{\mathbb{P}_N}\left[\mathfrak{c}_N(z)\right]+\mathbb{E}_{\mathbb{P}_N}\left[\mathfrak{c}_N'(z)\Delta G_N(\sigma(z))\right]. \tag{6.6}$$

Combining that $\xi_N(z),\xi_N'(z),\mathfrak{c}_N(z)$ and $\mathfrak{c}_N'(z)$ are uniformly bounded over compacts in $\mathcal{O}\cap\mathcal{M}$ with Proposition 6.1.1 we conclude that (6.6) is $O(1)$ as $N\to\infty$. This proves (6.3) for all $m\geq 0$.

6.2. Self-improving estimates and the proof of Proposition 6.1.1

In this section we prove Proposition 6.1.1. For clarity we split the proof into several steps.

Step 1. In the first step we derive a weak a priori estimate on $\mathbb{E}_{\mathbb{P}_N}\left[\prod_{a=1}^m|\Delta G_N(v_a)|\right]$, which will be iteratively improved in the steps below until we reach the desired estimate of the proposition. More precisely, we show that for each $n\in\mathbb{N}$, compact subset $K\subset U$ and $v_1,\dots,v_n\in K$ we have

$$\mathbb{E}_{\mathbb{P}_N}\left[\prod_{i=1}^n|\Delta G_N(v_i)|\right]=O\left(N^{n/2+1/2}\right), \tag{6.7}$$

where the constant in the big O notation depends on K and n.

Recall from Section 5.1 that $\Delta G_N(v)=N\left(G_N^d(v)-G_N^c(v)\right)$, where

$$G_N^d(v)=\int_{\mathbb{R}}\frac{\mu_N(dx)}{v-x}\text{ and }G_N^c(v)=\int_{\mathbb{R}}\frac{\hat{\mu}_N(dx)}{v-x}.$$

Using Hölder's inequality, we can reduce (6.7) to showing that for all $v\in K$ we have

$$\mathbb{E}_{\mathbb{P}_N}\left[N^n\left|G_N^d(v)-G_N^c(v)\right|^n\right]=O\left(N^{n/2+1/2}\right) \tag{6.8}$$

Fix $\eta > 0$ small enough so that the η neighborhood of $S = \cup_{j=1}^{k}[\hat{a}_j, \hat{b}_j]$ is disjoint from K. Let $h(x)$ be a smooth function, whose support is inside the η-neighborhood of S, and such that $h(x) = 1$ on an $\eta/2$-neighborhood of S. Note that for all N sufficiently large we have that μ_N and $\hat{\mu}_N$ are both supported in the $\eta/2$-neighborhood of S. Setting $g(x) = (v - x)^{-1}$ we have

$$G_N^d(v) - G_N^c(v) = \int_{\mathbb{R}} g(x)h(x)\mu_N(dx) - \int_{\mathbb{R}} g(x)h(x)\mu(dx).$$

We can apply Corollary 3.1.4 for the function $g \cdot h$ with $a = r \cdot N^{1/2n-1/2}$, $r > 0$ and $p = 3$ to get

$$\mathbb{P}_N\left(\left|G_N^d(v) - G_N^c(v)\right| \geq c_1 r N^{-1/2+1/2n} + c_2 N^{-3}\right) \leq \exp\left(CN\log(N) - 2\theta\pi^2 r^2 N^{1+1/n}\right),$$

which implies (6.8).

Step 2. In this step we reduce the proof of the proposition to the establishment of the following self-improvement estimate claim.

Claim: Suppose that for some $n, M \in \mathbb{N}$ we have that

$$\mathbb{E}_{\mathbb{P}_N}\left[\prod_{a=1}^{m}|\Delta G_N(v_a)|\right] = O(1) + O\left(N^{m/2+1-M/2}\right) \text{ for } m = 1, \ldots, 4n+4, \tag{6.9}$$

where the constants in the big O notations are uniform as v_a vary over compact subsets of U for $a = 1, \ldots, 4n+4$. Then we have

$$\mathbb{E}_{\mathbb{P}_N}\left[\prod_{a=1}^{m}|\Delta G_N(v_a)|\right] = O(1) + O\left(N^{m/2+1-(M+1)/2}\right) \text{ for } m = 1, \ldots, 4n. \tag{6.10}$$

The proof of the above claim will be established in the following steps. For now we assume its validity and conclude the proof of the proposition.

Notice that (6.7) implies that (6.9) holds for the pair $n = 2l$ and $M = 1$. The conclusion is that (6.9) holds for the pair $n = 2l - 1$ and $M = 2$. Iterating the argument an additional l times we conclude that (6.9) holds with $n = l - 1$ and $M = l + 2$, which implies the proposition.

Step 3. In this step we prove that

$$\begin{aligned} &M_N(\Delta G_N(v_0)|v_a, \{1, \ldots, m\}) = O(1) + O\left(N^{m/2+1-M/2}\right) \\ &\text{for } m = 1, \ldots, 4n+2 \text{ and } \mathbb{E}_{\mathbb{P}_N}\left[\Delta G_N(v_0)\right] = O(1) + O\left(N^{1-M/2}\right), \text{ where} \end{aligned} \tag{6.11}$$

the constants in the big O notation are uniform over $v_0, \dots, v_m$ in compact subsets of U.

We start by fixing a compact subsets $\mathcal{V} \subset U$, which is invariant under conjugation and let $\Gamma' = \Gamma_1 \cup \Gamma_2$ be as in Section 5.3 with $\sigma(\Gamma_1)$ and the bounded components of $U \setminus \sigma(\Gamma_1)$ disjoint from $\mathcal{V}$. For $m = 1, \dots, 4n+2$, we differentiate both sides of (5.24) with respect to $t_1, \dots, t_m$ and set $t_1 = \cdots = t_m = 0$. Combining (6.1), (5.28) and (5.30) the result we obtain is

$$M_N(\Delta G_N(v_0)|v_a, \{1,\dots,m\}) = \frac{NP_N^m(v_0)}{D_N(v_0)} + \frac{\mathbf{1}_{\{m=1\}}}{4\pi i \cdot \prod_{j=1}^k \sqrt{(v_0 - r_j)(v_0 - s_j)}} \times \oint_{\sigma(\Gamma_1)} \frac{\prod_{j=1}^k \sqrt{(w-r_j)(w-s_j)}}{(w-v_0)(w-v_1)^2} dw + \frac{-D_N(v_0)^{-1}}{2\pi i} \oint_{\Gamma'} \frac{N \cdot \mathrm{Rem}_N^m(z) dz}{z H_N(z)(v_0 - \sigma(z))}, \text{ where} \tag{6.12}$$

$$\begin{aligned} \mathrm{Rem}_N^m(z) &= \frac{\partial^m}{\partial t_1 \cdots \partial t_m} \frac{\mathrm{Rem}_N^{\mathbf{t},\mathbf{v}}(z)}{W_m^{\mathbf{t},\mathbf{v}}(z)}\bigg|_{t_a=0, 1\le a\le m} P_N^m(v_0) \\ &= \frac{\partial^m}{\partial t_1 \cdots \partial t_m} P_N^{\mathbf{t},\mathbf{v}}(v_0)\bigg|_{t_a=0,1\le a\le m}. \end{aligned} \tag{6.13}$$

By the same arguments following (5.30) we may apply the map $\tilde{\Upsilon}_{v_0}^N$ from (5.8) for the contours $\sigma(\gamma_1'), \dots, \sigma(\gamma_k')$ and the points r_j, s_j for $j = 1, \dots, k$ to both sides of (6.12) to get

$$\begin{aligned} M_N(\Delta G_N(v_0)|v_a, \{1,\dots,m\}) = \tilde{\Upsilon}_{v_0}^N \left[\frac{-D_N(v_0)^{-1}}{2\pi i} \oint_{\Gamma'} \frac{N \cdot \mathrm{Rem}_N^m(z) dz}{z H_N(z)(v_0 - \sigma_N(z))} \right] + \\ \tilde{\Upsilon}_{v_0}^N \left[\frac{\mathbf{1}_{\{m=1\}}}{4\pi i \cdot \prod_{j=1}^k \sqrt{(v_0 - r_j)(v_0 - s_j)}} \oint_{\sigma(\Gamma_1)} \frac{\prod_{j=1}^k \sqrt{(w-r_j)(w-s_j)}}{(w-v_0)(v_1-w)^2} dw \right]. \end{aligned} \tag{6.14}$$

Combining (6.14) and an application of $\tilde{\Upsilon}_{v_0}^N$ to both sides of (5.24) we get

$$\begin{aligned} M_N(\Delta G_N(v_0)|v_a, \{1,\dots,m\}) &= O(1) + \tilde{\Upsilon}_{v_0}^N \left[\frac{-D_N(v_0)^{-1}}{2\pi i} \oint_{\Gamma'} \frac{N \cdot \mathrm{Rem}_N^m(z) dz}{z H_N(z)(v_0 - \sigma(z))} \right], \\ \mathbb{E}_{\mathbb{P}_N}[\Delta G_N(v_0)] &= O(1) + \tilde{\Upsilon}_{v_0}^N \left[\frac{-D_N(v_0)^{-1}}{2\pi i} \oint_{\Gamma'} \frac{N \cdot \mathrm{Rem}_N(z) dz}{z H_N(z)(v_0 - \sigma(z))} \right]. \end{aligned} \tag{6.15}$$

The constants in the big O notation are uniform over $v_0, v_1, \dots, v_m$ in compact subsets of $\mathcal{V}$.

At this time we recall (6.5), which states that for $m = 1, \dots, 4n+2$ we have

$$\begin{aligned} N^2 \cdot \mathrm{Rem}_N^m(z) = \sum_{A \subset \{1,\dots,m\}} & \left[M_N \left(\partial_A \mathfrak{c}_N(z; \mathbf{0}, \mathbf{v}) | v_a, A^c \right) \right. \\ & \left. + M_N(\partial_A \mathfrak{c}'_N(z; \mathbf{0}, \mathbf{v}) \cdot \Delta G_N(\sigma(z)) | v_a, A^c) \right] \\ & + M_N \left(\xi_N(z) \Delta G_N(\sigma(z))^2 | v_a, \{1, \dots, m\} \right) \\ & + M_N \left(\xi'_N(z) \partial_z \Delta G_N(\sigma(z)) | v_a, \{1, \dots, m\} \right). \end{aligned}$$

Recall that $\xi_N(z)$, $\xi'_N(z)$, $\partial_A \mathfrak{c}_N(z; \mathbf{0}, \mathbf{v})$ and $\partial_A \mathfrak{c}'_N(z; \mathbf{0}, \mathbf{v})$ are all $O(1)$ if $v_1, \dots, v_m \in \mathcal{V}$ and $z \in \Gamma'$. The latter and (6.9) imply

$$\mathrm{Rem}_N^m(z) = O\left(N^{-2}\right) + O\left(N^{m/2 - M/2}\right), \text{ for } m = 1, \dots, 4n+2. \tag{6.16}$$

By combining (6.6) and (6.9) we get that (6.16) holds for $m = 0$ as well. Finally, (6.15), (6.16) and Property (1) in Proposition 5.1 together imply (6.11).

Step 4. In this step we will establish the validity of (6.10) except for a single case, which will be handled separately in the next step.

Notice that by Hölders inequality we have

$$\sup_{v_1, \dots, v_m \in \mathcal{V}} \mathbb{E}_{\mathbb{P}_N} \left[\prod_{a=1}^m |\Delta G_N(v_a)| \right] \leq \sup_{v \in \mathcal{V}} \mathbb{E}_{\mathbb{P}_N} \left[|\Delta G_N(v)|^m \right],$$

and so to finish the proof it suffices to show that for $m = 1, \dots, 4n$ we have

$$\mathbb{E}_{\mathbb{P}_N} \left[|\Delta G_N(v)|^m \right] = O(1) + O\left(N^{m/2 + 1/2 - M/2}\right). \tag{6.17}$$

Since centered moments are linear combinations of products of joint cumulants, we deduce from the first line in (6.11) that for $m = 1, \dots, 4n+2$ we have

$$\sup_{v_0, \dots, v_{m-1} \in \mathcal{V}} \mathbb{E}_{\mathbb{P}_N} \left[\prod_{a=0}^{m-1} \left(\Delta G_N(v_a) - \mathbb{E}_{\mathbb{P}_N}[\Delta G_N(v_a)] \right) \right] = O(1) + O\left(N^{(m-1)/2 + 1 - M/2}\right). \tag{6.18}$$

Combining the latter with the first and second lines of (6.11) we see that

$$\sup_{v_0, \dots, v_{m-1} \in \mathcal{V}} \mathbb{E}_{\mathbb{P}_N} \left[\prod_{a=0}^{m-1} \Delta G_N(v_a) \right] = O(1) + O\left(N^{(m-1)/2 + 1 - M/2}\right). \tag{6.19}$$

If $m = 2m_1$ then we can set $v_0 = \cdots v_{m_1 - 1} = v$ and $v_{m_1} = \cdots = v_{2m_1 - 1} = \overline{v}$ in (6.18), which yields

$$\sup_{v\in\mathcal{V}} \mathbb{E}_{\mathbb{P}_N}\left[|\Delta G_N(v)|^m\right] = O(1) + O\left(N^{m/2+1/2-M/2}\right) \text{ for } m = 2,4,6,\dots,4n+2\,. \quad (6.20)$$

We next let $m = 2m_1+1$ be odd and notice that by the Cauchy–Schwarz inequality and (6.20)

$$\begin{aligned}&\sup_{v\in\mathcal{V}} \mathbb{E}_{\mathbb{P}_N}\left[|\Delta G_N(v)|^{2m_1+1}\right] \le \\ &\sup_{v\in\mathcal{V}} \left[\mathbb{E}_{\mathbb{P}_N}\left[|\Delta G_N(v)|^{2m_1+2}\right]\right]^{1/2} \cdot \left[\mathbb{E}_{\mathbb{P}_N}\left[|\Delta G_N(v)|^{2m_1}\right]\right]^{1/2} = \\ &O(1) + O\left(N^{m_1+1-M/2}\right) + O\left(N^{m_1/2+3/4-M/4}\right).\end{aligned} \quad (6.21)$$

We note that the bottom line of (6.21) is $O(1) + O\left(N^{m_1+1-M/2}\right)$ except when $M = 2m_1+2$, since

$$m_1/2 + 3/4 - M/4 \le \begin{cases} m_1 + 1 - M/2 & \text{when } M \le 2m_1+1, \\ 0 & \text{when } M \ge 2m_1+3. \end{cases}$$

Consequently, (6.20) and (6.21) together imply (6.17) except when $M = 2m_1+2$ and $m = 2m_1+1$. We will handle this case in the next step.

Step 5. In this last step we will show that (6.17) holds even when $M = 2m_1+2$ and $4n > m = 2m_1+1$. In the previous step we showed in (6.20) that $\sup_{v\in\mathcal{V}} \mathbb{E}_{\mathbb{P}_N}\left[|\Delta G_N(v)|^{2m_1+2}\right] = O\left(N^{1/2}\right)$, and below we will improve this estimate to

$$\sup_{v\in\mathcal{V}} \mathbb{E}_{\mathbb{P}_N}\left[|\Delta G_N(v)|^{2m_1+2}\right] = O(1). \quad (6.22)$$

The trivial inequality $x^{2m_1+2} + 1 \ge |x|^{2m_1+1}$ together with (6.22) implies

$$\sup_{v\in\mathcal{V}} \mathbb{E}_{\mathbb{P}_N}\left[|\Delta G_N(v)|^{2m_1+1}\right] = O(1).$$

Consequently, we have reduced the proof of the claim to establishing (6.22).

Let us list the relevant estimates we will need

$$\begin{aligned}&\mathbb{E}_{\mathbb{P}_N}\left[\prod_{a=1}^{2m_1+4} |\Delta G_N(v_a)|\right] = O\left(N^{3/2}\right),\ \mathbb{E}_{\mathbb{P}_N}\left[\prod_{a=1}^{2m_1+2} |\Delta G_N(v_a)|\right] = O\left(N^{1/2}\right), \\ &\mathbb{E}_{\mathbb{P}_N}\left[\prod_{a=1}^{j} |\Delta G_N(v_a)|\right] = O(1) \text{ for } 0 \le j \le 2m_1,\ \mathbb{E}_{\mathbb{P}_N}\left[\prod_{a=1}^{2m_1+3} |\Delta G_N(v_a)|\right] = O(N).\end{aligned} \quad (6.23)$$

The above identities follow from (6.20) and (6.21). All constants are uniform over $v_a \in \mathcal{V}$. Below we feed the improved estimates of (6.23) into Steps 3 and 4, which would ultimately yield (6.22).

In Step 3, we note that we have the following version of (6.16)

$$\mathrm{Rem}_N^m(z) = O\left(N^{-1}\right) \text{ whenever } 0 \leq m \leq 2m_1 + 1. \tag{6.24}$$

The latter statement follows from (6.5) for $2m_1 + 1 \geq m \geq 1$ and (6.6) for $m = 0$, combined with

$$\mathbb{E}_{\mathbb{P}_N}\left[\prod_{a=1}^{m+2} |\Delta G_N(v_a)|\right] = O(N) \text{ for } m = 0, \ldots, 2m_1 + 1,$$

the latter being a consequence of (6.23). Using (6.24) instead of (6.16) in Step 3, we obtain the following improvement over (6.11)

$$\begin{aligned} &M_c(\Delta G_N(v), \ldots, \Delta G_N(v_m)) = O(1), \text{ for } m = 1, \ldots, 2m_1 + 1 \text{ and} \\ &\mathbb{E}_{\mathbb{P}_N}\left[\Delta G_N(v)\right] = O(1). \end{aligned} \tag{6.25}$$

We next repeat the arguments in Step 4, and note that by using (6.25) in place of (6.11) we obtain the following improvement over (6.19)

$$\sup_{v_0, \ldots, v_{2m_1+1} \in \mathcal{V}} \mathbb{E}_{\mathbb{P}_N}\left[\prod_{a=0}^{2m_1+1} \Delta G_N(v_a)\right] = O(1). \tag{6.26}$$

Setting $v_0 = \cdots v_{m_1} = v$ and $v_{m_1+1} = \cdots = v_{2m_1+1} = \overline{v}$ in (6.26) we get (6.22).

7. q-Racah tiling models and ensembles

As discussed in Section 1 our main motivation for studying discrete log-gases on shifted quadratic lattices comes from the q-Racah tiling model that was introduced in [16]. In Section 7.1 we give a formal definition of the model and in Section 7.2 we state the main results we prove about it in Theorems 7.2.2 and 7.2.4. In Section 7.3 we explain how the model is related to a certain random particle system that we call the q-Racah ensemble and state a law of large numbers and central limit theorem for the latter as Theorems 7.4.4 and 7.4.5 in Section 7.4.

7.1. The q-Racah tiling model

7.1.1. Lozenge tilings

Denote by $\Omega_{a\times b\times c}$ the set of all tilings of a hexagon with side lengths a, b, c by rhombi (or alternatively boxed plane partitions), see Fig. 4. Denote the horizontal rhombi by

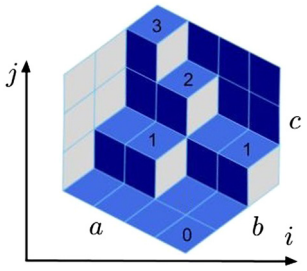

Fig. 4. Tiling of a $3 \times 3 \times 3$ hexagon.

$\diamondsuit$ and introduce coordinate axes (i, j). Given two parameters q and κ we define the probability of an element $\mathcal{T} \in \Omega_{a\times b\times c}$ through

$$\mathbb{P}(\mathcal{T}) = \frac{w(\mathcal{T})}{\sum_{\mathcal{S}\in\Omega_{a\times b\times c}} w(\mathcal{S})}, \text{ where } w(\mathcal{T}) = \prod_{\diamondsuit\in\mathcal{T}} w(\diamondsuit), \text{ and}$$
$$w(\diamondsuit) = \kappa^2 q^{j-(c+1)/2} - q^{-j+(c+1)/2}. \tag{7.1}$$

In the above formula the product is over all horizontal lozenges $\diamondsuit$ that belong to $\mathcal{T}$ and j denotes the j-th coordinate of the topmost point of $\diamondsuit$. We call the probability measure in (7.1) the *q-Racah tiling model*.

It was shown in [16] that the *partition function* (the sum of all weights $w(\mathcal{T})$ or the normalization term in (7.1)) has a nice *product form*, which generalizes the famous MacMahon formula for the number of boxed plane partitions [67]. Note that the number of horizontal rhombi in all tilings of a given hexagon is the same, hence $\mathbb{P}$ is invariant under multiplication of $w(\diamondsuit)$ by a constant.

In order for (7.1) to define an honest probability measure, one requires that the weights $w(\mathcal{T})$ be non-negative. This imposes certain restrictions on the parameters q, κ and there are three possible cases that lead to positive weights:

(i) *imaginary q-Racah case:* q is a positive real number and κ is a purely imaginary number;
(ii) *real q-Racah case:* q is a positive real number and κ is a real number that cannot lie inside the interval $\left[q^{-a+1/2}, q^{(b+c-1)/2}\right]$ if $q > 1$ or the interval $\left[q^{(b+c-1)/2}, q^{(a-1)/2}\right]$ if $q < 1$;
(iii) *trigonometric q-Racah case:* q and κ are complex numbers on the unit circle, i.e. $q = e^{i\alpha}$, $\kappa = e^{i\beta}$, where α, β must be such that $-\alpha(b+c-1)/2+\beta$ and $\alpha(a-1/2)+\beta$ must lie in the same interval of the form $[\pi k, \pi(k+1)]$, $k \in \mathbb{Z}$.

The names of the above cases are related to those of the classical orthogonal polynomials that appear in the analysis. In this paper, we will only consider the real q-Racah case with $q \in (0, 1)$ and $\kappa \in \left[0, q^{(b+c-1)/2}\right)$ although most of our arguments can be extended to other cases.

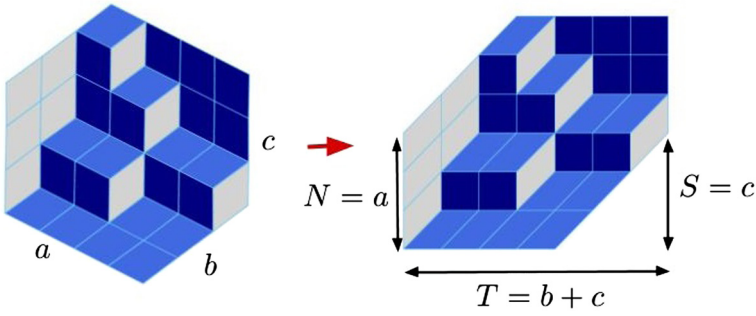

Fig. 5. Affine transformation of lozenges.

If we let $\kappa \to 0$ then we get the *q-Hahn case* $w(\diamondsuit) = q^{-j}$. In this case the probability of a plane partition is proportional to q^{-V}, where V denotes the *volume* of the plane partition, i.e. the number of cubes that it contains. If we send $\kappa \to \infty$ we get that the probability of a partition is proportional to q^{V}. In this sense, one can view our model as an interpolation between the models q^{V} and q^{-V}. Finally, if one sends $\kappa \to 0$ and $q \to 1$, one recovers the uniform measure on boxed plane partitions.

7.1.2. Particle configurations

In what follows we describe an alternative formulation of our model that is more suitable for stating our results. We perform a simple affine transformation of the hexagon and lozenges, detailed in Fig. 5.

Let us introduce new parameters N, T, S that are related to a, b, c through $N = a$, $T = b + c$ and $S = c$. Each tiling in $\Omega_{a\times b\times c}$ naturally corresponds to a family of N non-intersecting up-right paths as shown in Fig. 6. For each $0 \le t \le T$ we draw a vertical line through the point $(t, 0)$ and denote by $x_1^t < x_2^t < \cdots < x_N^t$ the intersection of the line with the N up-right paths. We interpret the intersection points as particles and will typically use the same letter to refer to a particle and its location. In this way, we can view a tiling as an N-point (or particle) configuration, which varies in time $t = 0, \dots, T$. Observe that when $t = 0$ the configuration consists of the points $\{0, 1, \dots, N-1\}$ and when $t = T$ the configuration consists of the points $\{S, S+1, \dots, S+N-1\}$.

Given a random configuration $\{x_k^t\}$ we define the random *height function*

$$h : \mathbb{Z}_{\ge 0} \times \left(\mathbb{Z} + \frac{1}{2}\right) \to \mathbb{Z}_{\ge 0} \text{ as } h(t, s) = |\{k \in \{1, \dots, N\} : x_t^k < s\}|. \tag{7.2}$$

In terms of the tiling in Fig. 6 the height function is defined at the vertices of rhombi, and it counts the number of particles below a given vertex. The latter definition is in agreement with the standard three-dimensional interpretation of the tiling as a stack of boxes [43].

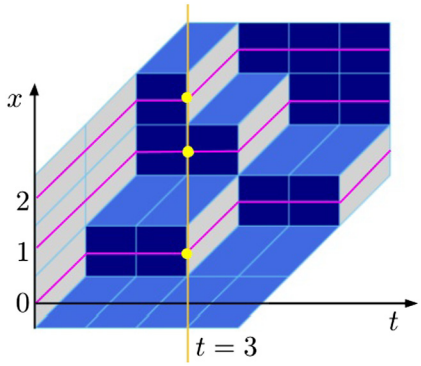

Fig. 6. Modified hexagon and up-right path configuration (in purple). The yellow dots are the particles at time $t = 3$ and we have $x_1^t = 1, x_2^t = 3$ and $x_3^t = 4$. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

7.2. Main results for the q-Racah tiling model

Our results are about the global fluctuations of a random lozenge tiling with distribution (7.1) when the parameters q, κ and the sizes of the hexagon N, T, S scale in a particular fashion that we detail below.

Definition 7.2.1. We assume that we are given real numbers $\mathtt{N}, \mathtt{T}, \mathtt{S}, \mathtt{q}$ and $\mathtt{k}$ such that

$$\mathtt{N}, \mathtt{T}, \mathtt{S}, \mathtt{q} > 0, \qquad \mathtt{k} \geq 0, \qquad \mathtt{q} < 1, \qquad \mathtt{N} < \mathtt{T}, \qquad \mathtt{S} < \mathtt{T}, \qquad \mathtt{k}^2 \mathtt{q}^{-\mathtt{T}} < 1.$$

Given such a choice of parameters and $\varepsilon \in (0,1)$ we let $\mathbb{P}_\varepsilon$ be the probability measure in (7.1) with

$$q = \mathtt{q}^\varepsilon + O(\varepsilon^2), \qquad N = \mathtt{N}\varepsilon^{-1} + O(1), \qquad T = \mathtt{T}\varepsilon^{-1} + O(1),$$
$$S = \mathtt{S}\varepsilon^{-1} + O(1), \qquad \kappa = \mathtt{k} + O(\varepsilon).$$

7.2.1. Limit shape

Our first result concerns the hydrodynamic limit of the height function h, with distribution $\mathbb{P}_\varepsilon$, under the parameter scaling in Definition 7.2.1 when ε converges to zero. On a macroscopic scale the random height function concentrates around a deterministic *limit shape*, i.e.

$$\varepsilon \cdot h\left(\lfloor x\varepsilon^{-1} \rfloor, \lfloor y\varepsilon^{-1} \rfloor + 1/2\right) \to \hat{h}(x,y) \text{ as } \varepsilon \to 0^+, \tag{7.3}$$

where (x,y) are the new global continuous coordinates, $\hat{h}(x,y)$ is the function whose graph is the limit shape and the convergence is in probability. The new coordinates (x,y) are assumed to belong to the limiting hexagon $\mathcal{P}$, which is parametrized by $\mathtt{N}, \mathtt{S}, \mathtt{T}$ the same way that our discrete hexagon was parametrized by N, S, T; see the central

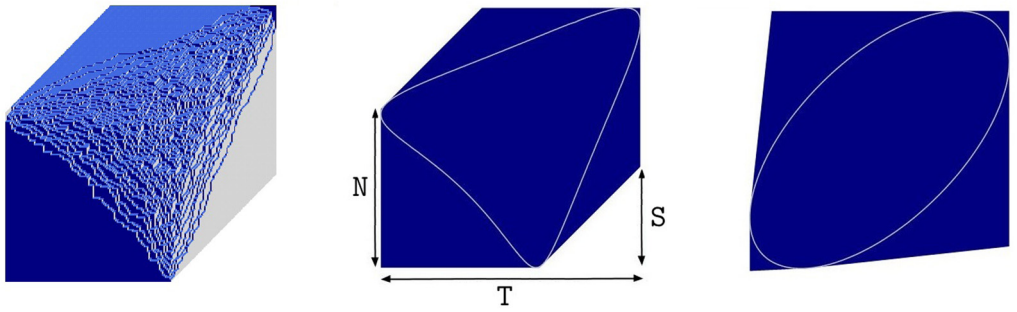

Fig. 7. The left part shows a simulation of a tiling. The middle part shows the hexagon $\mathcal{P}$ and the liquid region $\mathcal{D}$ is the region inside the gray curve. The right part denotes the image of $\mathcal{P}$ and $\mathcal{D}$ under the map $(x, y) \to (\mathtt{q}^{-x}, \mathtt{q}^{-y} + \mathtt{k}^2\mathtt{q}^{-\mathtt{S}-x+y})$.

part of Fig. 7. The limit shape can then be understood as a continuous function on $\mathcal{P}$ and we describe it next.

With parameters as in Definition 7.2.1 we define ϕ as

$$\phi(x,y) = \arccos\left(\frac{\left(\mathtt{q}^{-\mathtt{N}}-1\right)\left(1-\mathtt{q}^{-\mathtt{N}-\mathtt{T}}\right)\left(\mathtt{q}^{-2x}-\mathtt{k}^2\mathtt{q}^{-y-\mathtt{S}}\right)^2 + A + B}{2\sqrt{AB}}\right), \text{ where}$$

$$A = \left(\mathtt{q}^{-x}-\mathtt{q}^{-\mathtt{S}-\mathtt{N}}\right)\left(\mathtt{q}^{-x}-\mathtt{k}^2\mathtt{q}^{-\mathtt{T}}\right)\left(\mathtt{q}^{-x}-\mathtt{q}^{-y-\mathtt{N}}\right)\left(\mathtt{q}^{-x}-\mathtt{k}^2\mathtt{q}^{-y-\mathtt{S}}\right),$$

$$B = \mathtt{q}^{-2\mathtt{N}-\mathtt{T}}\left(\mathtt{q}^{-x}-1\right)\left(\mathtt{q}^{-x}-\mathtt{k}^2\mathtt{q}^{-t+\mathtt{N}}\right)\left(\mathtt{q}^{-x}-\mathtt{q}^{-y-\mathtt{S}+\mathtt{T}}\right)\left(\mathtt{q}^{-x}-\mathtt{k}^2\mathtt{q}^{-\mathtt{S}+\mathtt{N}}\right).$$

If the expression inside the arccosine is greater than 1, then we set $\phi = 0$ and if it less than -1, then we set $\phi = \pi$. In terms of the above function ϕ we define the limit shape $\hat{h}$ as

$$\hat{h}(x,y) = \frac{1}{\pi}\int_0^y \phi(x,u)du, \quad \text{for } (x,y) \in \mathcal{P}. \tag{7.4}$$

With the above notation we can state our limit shape result.

Theorem 7.2.2. *Suppose that* $\mathtt{N}, \mathtt{T}, \mathtt{S}, \mathtt{q}$, $\mathtt{k}$ *and* $\mathbb{P}_\varepsilon$ *are as in Definition 7.2.1 and that* h *is distributed according to* $\mathbb{P}_\varepsilon$. *Then for any* $(x,y) \in \mathcal{P}$ *and* $\eta > 0$ *we have*

$$\lim_{\varepsilon \to 0^+} \mathbb{P}_\varepsilon\left(\left|\varepsilon \cdot h\left(\lfloor x\varepsilon^{-1}\rfloor, \lfloor y\varepsilon^{-1}\rfloor + 1/2\right) - \hat{h}(x,y)\right| > \eta\right) = 0. \tag{7.5}$$

Remark 7.2.3. The formula for $\phi(x,y)$ was derived in Theorem 8.1 in [16]. We remark that while an explicit formula for the limit shape was obtained in [16], it was not proved that the height function actually converges to it. Theorem 7.2.2 constitutes a proof of this fact.

An important feature of our model is that the limit shape develops *frozen facets* where the function $\hat{h}(x,y)$ is linear. In terms of the tiling a frozen facet corresponds to a region where asymptotically only one type of lozenge is present. In addition, there is a connected open *liquid region* $\mathcal{D} \subset \mathcal{P}$, which interpolates the facets. Explicitly, the liquid region $\mathcal{D}$ is given by the set of points $(x,y) \in \mathcal{P}$ where the expression inside the arccosine in the definition of ϕ is in $(-1,1)$, i.e.

$$\mathcal{D} = \left\{ (x,y) \in \mathcal{P} : \left[\left(\mathtt{q}^{-\mathtt{N}} - 1\right) \left(1 - \mathtt{q}^{-\mathtt{N}-\mathtt{T}}\right) \left(\mathtt{q}^{-2x} - \mathtt{k}^2 \mathtt{q}^{-y-\mathtt{S}}\right)^2 + A + B \right]^2 < 4AB \right\}.$$

If $(x,y) \in \mathcal{D}$ then the limiting height function $\hat{h}$ is curved near (x,y): asymptotically inside the liquid region one observes all three types of lozenges, see e.g. [25,43,44] for further discussion regarding frozen and liquid regions in related contexts. In addition, the local distribution of the tiling near $(x,y) \in \mathcal{D}$ is described asymptotically by a certain ergodic translation-invariant Gibbs measure on lozenge tilings of the whole plane. Such a measure is unique up to fixed proportions of lozenges of all three types [65], and these proportions depend on the slope of $\hat{h}$ at the point (x,y). We refer the reader to [16] for a more detailed discussion of this fact for the model we consider, and also to [43,45,59,65] for analogous results in general dimer models.

7.2.2. Central limit theorem

Before stating our central limit theorem for the measures $\mathbb{P}_\varepsilon$ we introduce a transformation of our particle configuration from Section 7.1. This transformation is (in some sense) the natural way to view the particle system, and it allows us to identify its global asymptotic fluctuations with a $1D$ section of the two-dimensional Gaussian free field (GFF for short).

Given a point configuration $\{(t, x_k^t)\}$ we define a new point configuration $\{(U,V)\}$ through

$$U(t,k) = q^{-t} \text{ and } V(t,k) = q^{-x_k^t} + \kappa^2 q^{x_k^t - S - t} \text{ for } 0 \le t \le T \text{ and } 1 \le k \le N. \tag{7.6}$$

Similarly to before, we define a random height function for the new particle system

$$\mathcal{H} : \left\{q^0, q^{-1}, \dots, q^{-T}\right\} \times \mathbb{R} \to \mathbb{Z}_{\ge 0} \text{ as } \mathcal{H}(q^{-t}, v) = \left|\{k \in \{1, \dots, N\} : V(t,k) < v\}\right|. \tag{7.7}$$

One can formulate an equivalent statement to Theorem 7.2.2 for the height function $\mathcal{H}$. I.e. there will be asymptotically a deterministic limiting height function $\hat{\mathcal{H}}$, near which $\mathcal{H}$ concentrates with high probability. Moreover, if we set $\sigma_{\mathtt{q}}(x,y) = (\mathtt{q}^{-x}, \mathtt{q}^{-y} + \mathtt{k}^2 \mathtt{q}^{-\mathtt{S}-x+y})$ then we have the explicit relationship $\hat{h}(x,y) = \hat{\mathcal{H}}\left(\sigma_{\mathtt{q}}(x,y)\right)$.

The function $\sigma_{\mathtt{q}}$ maps the liquid region $\mathcal{D}$ bijectively to a new region $\mathcal{D}'$, parametrized through

$$\mathcal{D}' = \left\{(u,v) \in \mathbb{R}^2 : \tilde{Q}(u,v) < 0\right\},$$

where $\tilde{Q}(u,v) = \tilde{A}u^2 + \tilde{B}v^2 + \tilde{C}uv + \tilde{D}u + \tilde{E}v + \tilde{F}$, and $\tilde{A}, \tilde{B}, \tilde{C}, \tilde{D}, \tilde{E}, \tilde{F}$ are explicitly computable constants that depend on $\mathtt{q}, \mathtt{S}, \mathtt{T}, \mathtt{N}$ and $\mathtt{k}$. Consequently, $\mathcal{D}'$ is an ellipse, see the right of Fig. 7.

Our next goal is to define a *complex structure* on the limit shape surface — this is a bijective diffeomorphism $\Omega : \mathcal{D}' \to \mathbb{H}$. The significance of this map is that the fluctuations of $\mathcal{H}$ will be asymptotically described by the pullback of the GFF on $\mathbb{H}$ under the map Ω. The function $\Omega(u,v)$ is algebraic and it satisfies the following quadratic equation

$$a_2\Omega^2 + a_1\Omega + a_0 = 0, \text{ where} \tag{7.8}$$

a_2, a_1, a_0 are explicit *linear* functions of u and v and are such that $a_1^2 - 4a_2a_0 = \mathtt{q}^{2\mathtt{N}} \cdot \tilde{Q}(u,v)$ (see Section 8.2 for the details). Whenever $(u,v) \in \mathcal{D}'$ the polynomial (7.8) has two complex conjugate roots and we define $\Omega(u,v)$ to be the one that lies in $\mathbb{H}$.

We are now ready to state our main theorem for the q-Racah tiling model, giving the asymptotics of the global $1D$ fluctuations of $\mathbb{P}_\varepsilon$ in terms of the two-dimensional Gaussian free field. In Section 8.1 we recall the definition and basic properties of the GFF.

Theorem 7.2.4. *Suppose that* $\mathtt{N}, \mathtt{T}, \mathtt{S}, \mathtt{q}, \mathtt{k}$ *and* $\mathbb{P}_\varepsilon$ *are as in Definition 7.2.1 and that* $\mathcal{H}$ *is as in (7.7) for the distribution* $\mathbb{P}_\varepsilon$*. Fix* $u \in (1, \mathtt{q}^{-\mathtt{T}})$ *and let* $t(\varepsilon)$ *be a sequence of integers such that* $q^{-t(\varepsilon)} = u + O(\varepsilon)$*. Then the centered random height function*

$$\sqrt{\pi}\left(\mathcal{H}(q^{-t}, v) - \mathbb{E}_{\mathbb{P}_\varepsilon}\left[\mathcal{H}(q^{-t}, v)\right]\right)$$

converges to the $1d$ *section of the pullback of the Gaussian free field with Dirichlet boundary conditions on the upper half-plane* $\mathbb{H}$ *with respect to the map* Ω *in the following sense: For any set of polynomials* $f_i \in \mathbb{R}[x]$ *for* $i = 1, \dots, m$ *the joint distribution of*

$$\int_{\mathbb{R}} \sqrt{\pi}\left(\mathcal{H}(q^{-t}, v) - \mathbb{E}_{\mathbb{P}_\varepsilon}\left[\mathcal{H}(q^{-t}, v)\right]\right) f_i(v)dv, \qquad i = 1, \dots, m, \tag{7.9}$$

converges to the joint distribution of the similar averages

$$\int_{a(u)}^{b(u)} \mathcal{F}(\Omega(u,y)) f_i(y)dy, \qquad i = 1, \dots, m$$

of the pullback of the GFF. In the above formula $a(u), b(u)$ *are the* v*-coordinates of the two points where the vertical line through* u *intersects the ellipse* $\tilde{Q}(u,v) = 0$*.*

Equivalently, the variables in (7.9) converge jointly to a Gaussian vector $(X_1, \dots, X_m)$ *with mean zero and covariance*

$$\mathbb{E}[X_i X_j] = \int_{a(u)}^{b(u)} \int_{a(u)}^{b(u)} f_i(x) f_j(y) \left(-\frac{1}{2\pi} \log \left| \frac{\Omega(u,x) - \Omega(u,y)}{\Omega(u,x) - \overline{\Omega}(u,y)} \right| \right) dxdy. \tag{7.10}$$

As discussed in Section 1 the GFF is assumed to be a universal scaling limit in tiling models, which motivates its appearance in our setup. Another reason one might expect to see the GFF in our tiling model comes from its connection to the β-log gas with $\beta = 2$. We will elaborate this idea later in Section 9, but essentially there is a natural way to view the particle configuration on a fixed vertical slice as a discrete log-gas on a quadratic gases. Log-gases appear naturally in random matrix theory, and in that context there are several models that are known to converge to the GFF [10,11,14].

We end the section by remarking that Theorem 7.2.4 admits a natural two-dimensional generalization, which we formulate as Conjecture 8.4.1 in Section 8.4. At this time our methods only provide access to the global fluctuations at fixed vertical sections of the model, and so we cannot establish the full $2D$ result. Nevertheless, we provide some numerical simulations that give evidence for the validity of the conjecture.

7.3. The q-Racah ensemble

In this section we define the *q-Racah ensemble*.

Definition 7.3.1. Let $q \in (0,1)$, $M \in \mathbb{Z}_{\geq 0}$, α, β, $\delta \in \mathbb{R}$ and $\gamma = q^{-M-1}$. For $x \in \{0, 1, \dots, M\}$ we introduce the following weight function

$$w^{qR}(x) = \frac{(\alpha q, \beta\delta q, \gamma q, \gamma\delta q; q)_x}{(q, \alpha^{-1}\gamma\delta q, \beta^{-1}\gamma q, \delta q; q)_x} \frac{(1 - \gamma\delta q^{2x+1})}{(\alpha\beta q)^x (1 - \gamma\delta q)}, \tag{7.11}$$

where $(y_1, \dots, y_i; q)_k = (y_1; q)_k \cdots (y_i; q)_k$, and $(y; q)_k = (1 - y)(1 - yq) \cdots (1 - yq^{k-1})$ is the q-Pochhammer symbol.

Remark 7.3.2. The weight w^{qR} is the weight function of the q-Racah orthogonal polynomials, see e.g. [48, Section 3.2]. One can more generally have $\alpha = q^{-M-1}$ or $\beta\delta = q^{-M-1}$ instead of $\gamma = q^{-M-1}$. Our choice is dictated by the fact that under the substitutions $\gamma = q^{-M-1}$ and $\delta = 0$ the q-Racah weight reduces to the q-Hahn weight.

With the above notation we can define the q-Racah ensemble as follows.

Definition 7.3.3. Fix $N \in \mathbb{N}$ and let $\alpha, \beta, \gamma, \delta, q$ and M be as in Definition 7.3.1 with $M \geq N - 1$. Denote by $\mathfrak{X}$ the collection of N-tuples of integers

$$\mathfrak{X} = \{(\lambda_1, \dots, \lambda_N) \in \mathbb{Z}^N : 0 \leq \lambda_1 < \lambda_2 < \cdots < \lambda_N \leq M\}.$$

The q-Racah ensemble is a probability measure $\mathbb{P}^{qR}$ on the set $\mathfrak{X}$, given by

$$\mathbb{P}^{qR}(\lambda_1,\dots,\lambda_N)=\frac{1}{Z(N,M,\alpha,\beta,\gamma,\delta,q)}\prod_{1\le i<j\le N}\left(\sigma(q^{-\lambda_i})-\sigma(q^{-\lambda_j})\right)^2\cdot\prod_{i=1}^{N}w^{qR}(\lambda_i), \tag{7.12}$$

where $\sigma(z)=z+\gamma\delta qz^{-1}$ and Z is a normalization constant that makes the sum over $\mathfrak{X}$ equal to 1.

Observe that for general choice of parameters the expressions in (7.12) need not be non-negative. Consequently, we need to restrict the space of parameters so that $\mathbb{P}^{qR}$ is an honest probability measure. We isolate one possible choice that accomplishes this in the following definition.

Definition 7.3.4. We assume that the parameters $\alpha,\beta,\gamma,\delta,q\in\mathbb{R}$ and $M,N\in\mathbb{Z}$ are such that

$$M\ge N-1\ge 0,\ 1>q>0,\ \alpha,\beta>0,\ \delta\ge 0,\ \gamma=q^{-M-1},\ 1>\beta\delta,\ \beta\ge\gamma,\ \alpha\ge\gamma.$$

One readily verifies that the above choice of parameters makes (7.12) non-negative on all of $\mathfrak{X}$.

We end this section by detailing the connection between $\mathbb{P}^{qR}$ and the measure on tilings from Section 7.1 in the following theorem.

Theorem 7.3.5. *Fix $a,b,c\ge 1$ and set $N=a$, $T=b+c$ and $S=c$. Let $\mathbb{P}$ denote the probability distribution of (7.1) with parameters $q\in(0,1)$ and $\kappa\in\left[0,q^{(T-1)/2}\right)$. Fix $t\in\{0,1,\dots,T\}$ and let $(x_1^t,\dots,x_N^t)$ denote the random N-point configuration of Section 7.1.2. We have that*

(1) if $t<S$ and $t<T-S$ then the distribution of $(x_1^t,\dots,x_N^t)$ is $\mathbb{P}^{qR}$ with $M=t+N-1$, $\alpha=q^{-S-N}$, $\beta=q^{S-T-N}$, $\gamma=q^{-t-N}$ and $\delta=\kappa^2q^{-S+N}$;

(2) if $S-1<t<T-S+1$ then the distribution of $(x_1^t,\dots,x_N^t)$ is $\mathbb{P}^{qR}$ with $M=S+N-1$, $\alpha=q^{-t-N}$, $\beta=q^{t-T-N}$, $\gamma=q^{-S-N}$ and $\delta=\kappa^2q^{-t+N}$;

(3) if $T-S+1<t<S$ then the distribution of $(T-t-S+x_1^t,\dots,T-t-S+x_N^t)$ is $\mathbb{P}^{qR}$ with $M=T-S+N-1$, $\alpha=q^{-T-N+t}$, $\beta=q^{-t-N}$, $\gamma=q^{-T-N+S}$ and $\delta=\kappa^2q^{-T+t+N}$;

(4) if $S-1<t$ and $T-S-1<t$ then the distribution of $(T-t-S+x_1^t,\dots,T-t-S+x_N^t)$ is $\mathbb{P}^{qR}$ with $M=T-t+N-1$, $\alpha=q^{-T-N+S}$, $\beta=q^{-S-N}$, $\gamma=q^{-T-N+t}$ and $\delta=\kappa^2q^{-T+S+N}$.

In all cases the parameter q in the definition of $\mathbb{P}^{qR}$ is the same as the one that is given.

Proof. This is essentially [16, Theorem 4.1] and we refer to the same paper for the details. □

7.4. Global asymptotics for q-Racah ensembles

In this section we state a law of large numbers and a central limit theorem for the q-Racah ensembles — Theorems 7.4.4 and 7.4.5 below.

We begin by explaining how we are scaling the parameters in the q-Racah ensemble.

Definition 7.4.1. We assume that we have parameters $\mathtt{a}, \mathtt{b}, \mathtt{c}, \mathtt{d}, \mathtt{q}$ and $\mathtt{M}$ such that

$$1 > \mathtt{q} > 0,\ \mathtt{a}, \mathtt{b}, \mathtt{M} > 0,\ \mathtt{d} \geq 0,\ 1 > \mathtt{bd},\ \mathtt{c} = \mathtt{q}^{-\mathtt{M}},\ \mathtt{b} \geq \mathtt{c},\ \mathtt{a} \geq \mathtt{c},\ \mathtt{cq} > 1.$$

For future reference we denote the set of parameters $\mathtt{a}, \mathtt{b}, \mathtt{c}, \mathtt{d}, \mathtt{q}$ and $\mathtt{M}$ that satisfy the above conditions by $\mathtt{P}$ and view it as a subset of $\mathbb{R}^6$ with the subspace topology.

In addition, we assume that we have a sequence of parameters $\alpha_N, \beta_N, \gamma_N, \delta_N, q_N$ and M_N that satisfy the conditions in Definition 7.3.4 and such that for some constant $A > 0$ we have

$$\max\left(N\left|q_N - \mathtt{q}^{1/N}\right|, |\alpha_N - \mathtt{a}|, |\beta_N - \mathtt{b}|, |\gamma_N - \mathtt{c}|, |\delta_N - \mathtt{d}|, \left|N^{-1}M_N - \mathtt{M}\right|\right) \leq AN^{-1}.$$

We let $\mathbb{P}_N$ be the measure from Definition 7.3.3 with parameters $\alpha_N, \beta_N, \gamma_N, \delta_N, q_N, M_N$ and N.

Definition 7.4.2. Suppose we are given parameters a, b, c, d and q such that

$$1 > q > 0,\ a, b, c > 0,\ d \geq 0,\ 1 > bd,\ b \geq c,\ a \geq c,\ cq \geq 1.$$

We define the following polynomials

$$\Phi^+(z) = (z-a)(z-bd)(z-c)(z-cd), \qquad \Phi^-(z) = (z-1)(az-cd)(bz-c)(z-d)$$
$$R(z) = \Phi^+(z) + \Phi^-(z) - (abq-1)(q^{-1}-1)(z^2-cd)^2, \qquad Q(z)^2 = R(z)^2 - 4\Phi^-(z)\Phi^+(z).$$

With the above data we define

$$\mu(x) = \begin{cases} \frac{1}{\pi} \cdot \arccos\left(\frac{R(q^{-x})}{2\sqrt{\Phi^-(q^{-x})\Phi^+(q^{-x})}}\right) & \text{when } x \in (0, -\log_q(c)), \\ 0 & \text{otherwise.} \end{cases} \tag{7.13}$$

If the expression inside the arccosine is bigger than 1 we set $\mu = 0$ and if it is less than -1 we set $\mu = 1$. The square root is the usual one as on $(0, -\log_q(c))$ both $\Phi^-(q^{-x})$ and $\Phi^+(q^{-x})$ are positive.

We also isolate the following fact.

Lemma 7.4.3. *The polynomial Q^2 from Definition 7.4.2 factors completely over $\mathbb{R}$. If we enumerate its roots in increasing order $x_1, \ldots, x_8$ we get*

$$x_1 = x_2 = -\sqrt{cd}, \qquad d \le x_3 \le x_4 \le cd, \qquad x_5 = x_6 = \sqrt{cd}, \qquad 1 \le x_7 \le x_8 \le c. \tag{7.14}$$

Moreover, we have $x_3 = \frac{cd}{x_8}$ *and* $x_4 = \frac{cd}{x_7}$.

Proof. One readily checks that

$$z^{-4}q^2Q(z)^2 = \left(z - cdz^{-1}\right)^2 \cdot Q_0\left(z + cdz^{-1}\right), \text{ where } Q_0(x) = a_2x^2 + a_1x + a_0, \text{ with} \tag{7.15}$$

- $a_2 = (abq^2 - 1)^2$
- $a_1 = -2q(a^2b^2dq^2 + ab^2cdq^2 + a^2b^2q^2 + a^2bcq^2 + ab^2dq^2 + abcdq^2 + a^2bq^2 - 2ab^2dq - 2abcdq + abcq^2 - 2a^2bq - 2abcq - 2abdq - 2bcdq + abd - 2abq - 2acq + bcd + ab + ac + bd + cd + a + c)$,
- $a_0 = 4a^2b^2cdq^3 - 4a^2b^2cdq^4 + 4ab^2cd^2q^3 + a^2b^2d^2q^2 + 4a^2b^2dq^3 + 4a^2bcdq^3 - 2ab^2cd^2q^2 + 4ab^2cdq^3 + 4abc^2dq^3 + b^2c^2d^2q^2 - 2a^2b^2dq^2 - 2a^2bcdq^2 + 4a^2bcq^3 - 2ab^2cdq^2 - 2ab^2d^2q^2 - 2abc^2dq^2 - 2abcd^2q^2 + 4abcdq^3 - 2b^2cd^2q^2 - 2bc^2d^2q^2 + a^2b^2q^2 - 2a^2bcq^2 - 2a^2bdq^2 + a^2c^2q^2 - 2ab^2dq^2 - 16abcdq^2 - 2ac^2dq^2 + b^2d^2q^2 - 2bd^2dq^2 - 2bcd^2q^2 + c^2d^2q^2 - 2a^2bq^2 - 2a^2cq^2 + 4abcdq - 2abcq^2 - 2abcq^2 - 2ac^2q^2 - 2acdq^2 + 4bcd^2q - 2bcdq^2 - 2c^2dq^2 + a^2q^2 + 4abdq + 4acdq - 2acq^2 + 4bcdq + 4c^2dq + c^2q^2 + 4acq + 4cdq - 4cd.$

Consequently, what remains is to show that Q_0 has two real roots y_1 and y_2 such that $1 + cd \le y_1 \le y_2 \le c + d$. Indeed, if the latter is true we would have that $z + cdz^{-1} = y_1$ (resp. $z + cdz^{-1} = y_2$) has two real roots x_3, x_8 (resp. x_4, x_7) and these satisfy the conditions of the lemma.

A direct calculation shows that the discriminant of Q_0 equals

$$D = a_1^2 - 4a_0a_2 = 16(q-1)(bq-1)(bdq-1)(cq-1)(aq-1)(aq-d)(abq-1)(abq-c) \ge 0.$$

Thus Q_0 indeed has two real roots and to show that they both lie in the interval $[1 + cd, c + d]$ it suffices to show that $Q_0'(1 + cd) \le 0$ and $Q_0'(c + d) \ge 0$. We notice that

$$F_1 = Q_0'(1 + cd) = 2a_2(1 + cd) + a_1(1 + cd) \text{ and } F_2 = Q_0'(c + d) = 2a_2(c + d) + a_1(c + d)$$

are both linear functions of γ and by assumption $q^{-1} \le c \le \min(a, b)$. In particular, it suffices to check that $F_1 \le 0$ when $c = q^{-1}$ and $c = b$, while $F_2 \ge 0$ when $c = q^{-1}$ and $c = a$.

When $c = q^{-1}$ we have that

$$F_1 = 2(1 - q^{-1})(bq - 1)(abq^2 - 1)(aq - d) \le 0 \text{ and}$$
$$F_2 = 2(1 - q^{-1})(aq - 1)(abq^2 - 1)(bdq - 1) \ge 0.$$

When $c = b$ we have

$$F_1 = 2(bq-1)(aq-1)G_1 \text{ where } G_1 = ab^2dq^2 + abq^2 - 2abq - 2bdq + bd + 1,$$

while when $c = a$ we have

$$F_2 = 2(bq-1)(aq-1)G_2 \text{ where } G_2 = a^2bq^2 + abdq^2 - 2abdq - 2aq + a + d.$$

What remains is to show that $G_1 \leq 0$ and $G_2 \geq 0$. Note that G_1 and G_2 are linear functions in d and $d \in [0, b^{-1})$. Thus, it suffices to check that $G_1 \leq 0$ and $G_2 \geq 0$ when $d = 0$ and $d = b^{-1}$.

When $d = 0$ we have $G_1 = abq(q-1) + (1-abq) \leq 0$ and $G_2 = aq(abq-1) + a(1-q) \geq 0$. When $d = b^{-1}$ we have $G_1 = 2(q-1)(abq-1) \leq 0$ and $G_2 = b^{-1}\left((abq-1)^2 + ab(q-1)^2\right) \geq 0$. □

The formula for Q that we will use in the paper is

$$Q(x) = (abq - q^{-1}) \cdot (z^2 - cd) \cdot \sqrt{(z-x_3)(z-x_4)} \cdot \sqrt{(z-x_7)(z-x_8)} \tag{7.16}$$

7.4.1. Law of large numbers

In this section we state a law of large numbers theorem for the q-Racah ensembles as Theorem 7.4.4 below. Its proof will be established in Section 9.2. We assume we have the same parameters and measures $\mathbb{P}_N$ as in Definition 7.4.1 and define the empirical measures μ_N

$$\mu_N = \frac{1}{N}\sum_{i=1}^{N} \delta\left(\frac{\lambda_i}{N}\right) \text{ where } (\lambda_1, \ldots, \lambda_N) \text{ is } \mathbb{P}_N\text{-distributed.} \tag{7.17}$$

Theorem 7.4.4. *Under the assumptions in this section, we have that the measures μ_N concentrate (in probability) near $\mu(x)dx$, where $\mu(x)$ is as in Definition 7.4.2 with parameters $a = \mathtt{a}, b = \mathtt{b}, c = \mathtt{c}, d = \mathtt{d}$ and $q = \mathtt{q}$. More precisely, for each Lipschitz function $f(x)$ defined in a real neighborhood of the interval $[0, \mathtt{M}]$ and each $\varepsilon > 0$ the random variables*

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

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

7.4.2. Central limit theorem

In this section we state a central limit theorem for the q-Racah ensembles as Theorem 7.4.5 below. Its proof will be established in Section 9.2. We assume the same parameters and measures $\mathbb{P}_N$ as in Definition 7.4.1. It turns out that to better see the Gaussian structure of the q-Racah ensemble it is convenient to consider a transformed

particle system, given by $(y_1, \dots, y_N)$ with $y_i = \sigma_N(q_N^{-\lambda_i})$ and $\sigma_N(z) = z + \gamma_N \delta_N q_N z^{-1}$. Then the transformed empirical measure of the system is given by

$$\rho_N = \frac{1}{N}\sum_{i=1}^{N} \delta(y_i) = \frac{1}{N}\sum_{i=1}^{N} \delta\left(\sigma_N(q_N^{-\lambda_i})\right) \text{ where } (\lambda_1, \dots, \lambda_N) \text{ is } \mathbb{P}_N\text{-distributed.} \tag{7.18}$$

Theorem 7.4.5. *Take $m \geq 1$ polynomials $f_1, \dots, f_m \in \mathbb{R}[x]$. Let ρ_N be as in (7.18) and define*

$$\mathcal{L}_{f_i} = N \int_{\mathbb{R}} f_j(x)\rho_N(dx) - N\mathbb{E}_{\mathbb{P}_N}\left[\int_{\mathbb{R}} f_j(x)\rho_N(dx)\right] \text{ for } i = 1, \dots, m.$$

Then the random variables $\mathcal{L}_{f_i}$ converge jointly in the sense of moments to an m-dimensional centered Gaussian vector $X = (X_1, \dots, X_m)$ with covariance

$$Cov(X_i, X_j) = \frac{1}{(2\pi i)^2} \oint_{\Gamma} \oint_{\Gamma} f_i(s) f_j(t) \mathcal{C}(s,t) ds dt,$$

where Γ is a positively oriented contour, which encloses the interval $[1 + \mathtt{cd}, \mathtt{c} + \mathtt{d}]$. The covariance kernel $\mathcal{C}(s,t)$ is given by

$$\mathcal{C}(s,t) = -\frac{1}{2(s-t)^2}\left(1 - \frac{(s-a_-)(t-a_+) + (t-a_-)(s-a_+)}{2\sqrt{(s-a_-)(s-a_+)}\sqrt{(t-a_-)(t-a_+)}}\right), \tag{7.19}$$

where $a_- = x_3 + x_8$, $a_+ = x_4 + x_7$ and $x_1, \dots, x_8$ are the ordered roots of the polynomial $Q(z)^2$ from Definition 7.4.2 with parameters $a = \mathtt{a}, b = \mathtt{b}, c = \mathtt{c}, d = \mathtt{d}$ and $q = \mathtt{q}$, cf. Lemma 7.4.3.

8. Global asymptotics for the q-Racah tiling model

As discussed in Section 7.2 the $1D$ global fluctuations of our model are asymptotically described by an appropriate pullback of the Gaussian free field in $\mathbb{H}$. In Section 8.1 we provide some preliminaries on the two-dimensional Gaussian free field. In Section 8.2 we describe the complex structure Ω on the liquid region $\mathcal{D}'$ of our tiling model and show that Ω defines a bijection between $\mathcal{D}'$ and $\mathbb{H}$. In Section 8.3 we give the proof of Theorems 7.4.4 and 7.4.5. Finally, in Section 8.4 we state our 2D conjecture.

8.1. Gaussian free field

In this section we briefly recall the formulation and some basic properties of the Gaussian free field (GFF). Our discussion will follow the exposition in [14, Section 4.5]

and for a more thorough background on the subject we refer to [66], [29, Section 4], [36, Section 2], and the references therein.

Definition 8.1.1. The Gaussian free field with Dirichlet boundary conditions in the upper half-plane $\mathbb{H}$ is a (generalized) centered Gaussian field $\mathcal{F}$ on $\mathbb{H}$ with covariance given by

$$\mathbb{E}\left[\mathcal{F}(z)\mathcal{F}(w)\right] = -\frac{1}{2\pi}\log\left|\frac{z-w}{z-\overline{w}}\right|, \ z,w\in\mathbb{H}. \tag{8.1}$$

We remark that $\mathcal{F}$ can be viewed as a probability Gaussian measure on a suitable class of generalized functions on $\mathbb{H}$; however, one cannot define the value of $\mathcal{F}$ at a given point $z\in\mathbb{H}$ (this is related to the singularity of (8.1) at $z=w$).

Even though $\mathcal{F}$ does not have a pointwise value; one can define the (usual distributional) pairing $\mathcal{F}(\phi)$, whenever ϕ is a smooth function of compact support, and the latter is a mean zero normal random variable. In general, one can characterize the distribution of $\mathcal{F}$ through pairings with test functions as follows. If $\{\phi_k\}$ is any sequence of compactly supported smooth functions on $\mathbb{H}$ then the pairings $\{\mathcal{F}(\phi_k)\}$ form a sequence of centered normal variables with covariance

$$\mathbb{E}\left[\mathcal{F}(\phi_k)\mathcal{F}(\phi_l)\right] = \int\limits_{\mathbb{H}^2}\phi_k(z)\phi_l(w)\left(-\frac{1}{2\pi}\log\left|\frac{z-w}{z-\overline{w}}\right|\right)|dz|^2|dw|^2.$$

An important property of $\mathcal{F}$ that will be useful for us is that it can be integrated against smooth functions on smooth curves $\gamma\subset\mathbb{H}$. We isolate the statement in the following lemma.

Lemma 8.1.2. *[14, Lemma 4.6] Let $\gamma\subset\mathbb{H}$ be a smooth curve and μ a measure on $\mathbb{H}$, whose support is γ and whose density with respect to the natural (arc length) measure on γ is a given by a smooth function $g(z)$ such that*

$$\iint\limits_{\gamma\times\gamma} g(z)g(w)\left(-\frac{1}{2\pi}\log\left|\frac{z-w}{z-\overline{w}}\right|\right)dzdw<\infty. \tag{8.2}$$

Then

$$\int\limits_{\mathbb{H}}\mathcal{F}d\mu = \int\limits_{\gamma}\mathcal{F}(u)g(u)du$$

is a well-defined Gaussian centered random variable of variance given by (8.2). Moreover, if we have two such measures μ_1 and μ_2 (with two curves γ_1 and γ_2 and two densities g_1 and g_2), then $X_1=\int_{\gamma_1}\mathcal{F}(u)g_1(u)du$, $X_2=\int_{\gamma_2}\mathcal{F}(u)g_2(u)du$ are jointly Gaussian with covariance

$$\mathbb{E}[X_1 X_2] = \iint\limits_{\gamma_1 \times \gamma_2} g_1(z) g_2(w) \left(-\frac{1}{2\pi} \log \left| \frac{z-w}{z-\overline{w}} \right| \right) dz dw.$$

Another property of $\mathcal{F}$ that we require is that it behaves well under bijective maps, which leads to the notion of pullback.

Definition 8.1.3. Given a domain D and a bijection $\Omega : D \to \mathbb{H}$, the pullback $\mathcal{F} \circ \Omega$ is a generalized centered Gaussian field on D with covariance

$$\mathbb{E}\left[\mathcal{F}(\Omega(z))\mathcal{F}(\Omega(w))\right] = -\frac{1}{2\pi} \log \left| \frac{\Omega(z) - \Omega(w)}{\Omega(z) - \overline{\Omega}(w)} \right|, \ z, w \in D.$$

Integrals of $\mathcal{F} \circ \Omega$ with respect to measures can be computed through

$$\int\limits_D (\mathcal{F} \circ \Omega) d\mu = \int\limits_{\mathbb{H}} \mathcal{F} d\Omega(\mu),$$

where $d\Omega(\mu)$ stands for the pushforward of the measure μ.

The above definition immediately implies the following analogue of Lemma 8.1.2.

Lemma 8.1.4. *[14, Lemma 4.8] In the notation of Definition 8.1.3, let μ be a measure on D whose support is a smooth curve γ and whose density with respect to the natural (length) measure on γ is given by a smooth function $g(z)$ such that*

$$\iint\limits_{\gamma \times \gamma} g_1(z) g_2(w) \left(-\frac{1}{2\pi} \log \left| \frac{\Omega(z) - \Omega(w)}{\Omega(z) - \overline{\Omega}(w)} \right| \right) dz dw < \infty. \tag{8.3}$$

Then

$$\int\limits_D (\mathcal{F} \circ \Omega) d\mu = \int\limits_{\gamma} \mathcal{F}(\Omega(u)) g(u) du$$

is a well-defined Gaussian centered random variable of variance given by (8.3). Moreover, if we have two such measures μ_1 and μ_2 (with two curves γ_1 and γ_2 and two densities g_1 and g_2), then $X_1 = \int_{\gamma_1} \mathcal{F}(\Omega(u)) g_1(u) du$, $X_2 = \int_{\gamma_2} \mathcal{F}(\Omega(u)) g_2(u) du$ are jointly Gaussian with covariance

$$\mathbb{E}[X_1 X_2] = \iint\limits_{\gamma_1 \times \gamma_2} g_1(z) g_2(w) \left(-\frac{1}{2\pi} \log \left| \frac{\Omega(z) - \Omega(w)}{\Omega(z) - \overline{\Omega}(w)} \right| \right) dz dw.$$

We end this section by remarking that the Gaussian free field is conformally invariant: if ϕ is an automorphism of $\mathbb{H}$ (i.e. $\phi(z) = \frac{az+b}{cz+d}$ with $a, b, c, d \in \mathbb{R}$ and $ad - bc = 1$) then the distributions of $\mathcal{F}$ and $\mathcal{F} \circ \phi$ are the same.

8.2. Complex structure

In this section we adopt the same notation as in Section 7.2 and formulate the map Ω. We first observe that if $(x,y) \in \mathcal{P}$ then we have that $\phi(x,y) = \pi \cdot \mu\left(y/\mathtt{N}\right)$, where μ is as in Definition 7.4.2 for the parameters $q = \mathtt{q}^{\mathtt{N}}$, $a = \mathtt{q}^{-\mathtt{S}-\mathtt{N}}$, $b = \mathtt{q}^{\mathtt{S}-\mathtt{T}-\mathtt{N}}$, $c = \mathtt{q}^{-x-\mathtt{N}}$ and $d = \mathtt{k}^2\mathtt{q}^{-\mathtt{S}+\mathtt{N}}$. If we have that $R, Q, \Phi^{\pm}$ are as in Definition 7.4.2 with the same parameters then the liquid region $\mathcal{D}$ is given by

$$\mathcal{D} = \left\{(x,y) \in \mathcal{P} : Q(\mathtt{q}^{-y})^2 < 0\right\} = \left\{x \in (0,\mathtt{T}) \text{ and } y \in (-\log_{\mathtt{q}}(x_7), -\log_{\mathtt{q}}(x_8))\right\}, \quad (8.4)$$

where x_7, x_8 stand for the two roots of Q^2 in $(1,c)$ — see Lemma 7.4.3. From (7.15) we know

$$z^{-4}q^2Q(z)^2 = \left(z - cdz^{-1}\right)^2 \cdot Q_0\left(z + cdz^{-1}\right),$$

and if we set $u = \mathtt{q}^{-x}$ and $v = \mathtt{q}^{-y} + \mathtt{k}^2\mathtt{q}^{-\mathtt{S}-x+y}$ then we see that

$$Q_0\left(\mathtt{q}^{-y} + cd\mathtt{q}^{y}\right) = \tilde{Q}(u,v) = \tilde{A}u^2 + \tilde{B}v^2 + \tilde{C}uv + \tilde{D}u + \tilde{E}v + \tilde{F},$$

where $\tilde{A}, \tilde{B}, \tilde{C}, \tilde{D}, \tilde{E}, \tilde{F}$ are explicit constants that depend only on $\mathtt{q}, \mathtt{S}, \mathtt{N}, \mathtt{T}$ and $\mathtt{k}^2$ and not on x, y. Combining the last two observations, we see that

$$\mathcal{D}' = \left\{(u,v) : \tilde{Q}(u,v) < 0\right\} = \left\{(u,v) : u \in (1,\mathtt{q}^{-\mathtt{T}}) \text{ and } v \in \left(x_7 + cdx_7^{-1}, x_8 + cdx_8^{-1}\right)\right\}, \quad (8.5)$$

where we recall from Section 7.2 that $\mathcal{D}'$ is the image of $\mathcal{D}$ under the map $\sigma_{\mathtt{q}}(x,y) = (\mathtt{q}^{-x}, \mathtt{q}^{-y} + \mathtt{k}^2\mathtt{q}^{-\mathtt{S}-x+y})$. In particular, we have that $\mathcal{D}'$ is an ellipse.

We next consider the quadratic equation

$$\begin{aligned} &P(w;u,v) := a_2(u,v)w^2 + a_1(u,v)w + a_0(u,v) = 0, \text{ where } a_2 = \mathtt{q}^{\mathtt{N}}(v - 1 - \mathtt{k}^2\mathtt{q}^{-\mathtt{S}}u), \\ &a_1 = v\mathtt{q}^{\mathtt{N}}(\mathtt{q}^{-\mathtt{T}} - 1) + \left(u(\mathtt{q}^{-\mathtt{S}} - \mathtt{q}^{\mathtt{N}}) - \mathtt{q}^{-\mathtt{S}+\mathtt{N}} - \mathtt{q}^{-\mathtt{T}} + 2\mathtt{q}^{\mathtt{N}}\right) + u\mathtt{k}^2\mathtt{q}^{\mathtt{N}}(\mathtt{q}^{-\mathtt{T}} + \mathtt{q}^{-\mathtt{S}+\mathtt{N}} - 2\mathtt{q}^{-\mathtt{S}-\mathtt{T}}) \\ &\quad + \mathtt{k}^2\mathtt{q}^{-\mathtt{T}+\mathtt{N}}(\mathtt{q}^{-\mathtt{S}} - \mathtt{q}^{\mathtt{N}}) \text{ and } a_0 = (u-1)(\mathtt{q}^{-\mathtt{T}} - \mathtt{q}^{\mathtt{N}})(\mathtt{q}^{-\mathtt{S}} - 1)(1 - \mathtt{k}^2\mathtt{q}^{-\mathtt{T}+\mathtt{N}}). \end{aligned} \quad (8.6)$$

For the above equation one calculates $a_1^2 - 4a_2a_0 = \mathtt{q}^{2\mathtt{N}} \cdot \tilde{Q}(u,v)$ and so for $(u,v) \in \mathcal{D}'$ we have that the equation has two complex conjugate roots. We define the map $\Omega : \mathcal{D}' \to \mathbb{H}$ as

$$\Omega(u,v) = w(u,v) \text{ such that } P(w(u,v);u,v) = 0 \text{ and } w(u,v) \in \mathbb{H} \text{ for } (u,v) \in \mathcal{D}' \quad (8.7)$$

and from our earlier discussion Ω is well-defined and algebraic.

In the remainder of this section we show that Ω defines a bijective diffeomorphism between $\mathcal{D}'$ and $\mathbb{H}$ satisfies an important property that is used in the proof of Theorem 7.2.4 in the next section.

For convenience we denote

$$\begin{aligned}
\lambda_3 &= (1-\mathtt{q}^{\mathtt{N}})(1-\mathtt{k}^2\mathtt{q}^{\mathtt{N}})(\mathtt{q}^{-\mathtt{S}}-\mathtt{q}^{-\mathtt{T}}),\\
\lambda_2 &= \mathtt{k}^2\mathtt{q}^{\mathtt{N}}\left(\mathtt{q}^{-\mathtt{S}-\mathtt{T}}-\mathtt{q}^{-\mathtt{S}+\mathtt{N}}+\mathtt{q}^{-\mathtt{S}}-\mathtt{q}^{-\mathtt{T}}\right)-\mathtt{q}^{-\mathtt{S}}+\mathtt{q}^{\mathtt{N}},\\
\lambda_1 &= -(\mathtt{q}^{-\mathtt{T}}-\mathtt{q}^{\mathtt{N}})(1-\mathtt{k}^2\mathtt{q}^{-\mathtt{T}+\mathtt{N}})(\mathtt{q}^{-\mathtt{S}}-1) \text{ and}\\
\lambda_0 &= -(\mathtt{q}^{-\mathtt{T}}-1)(\mathtt{q}^{-\mathtt{T}}-\mathtt{q}^{\mathtt{N}})(1-\mathtt{k}^2\mathtt{q}^{-\mathtt{T}+\mathtt{N}})(\mathtt{q}^{-\mathtt{S}}-1).
\end{aligned}$$

Also we define the map $f:\mathbb{H}\to\mathbb{R}^2$ through $f(r+is)=(f_1(r,s),f_2(r,s))$ with

$$\begin{aligned}
&f_1(r,s)=1+\frac{\lambda_3(r^2+s^2)}{\lambda_2(r^2+s^2)+2r\lambda_1+\lambda_0},\ f_2(r,s)=1+\mathtt{k}^2\mathtt{q}^{-\mathtt{T}}+\\
&+\frac{\mathtt{k}^2\mathtt{q}^{\mathtt{N}}(\mathtt{q}^{-\mathtt{S}}-1)(\mathtt{q}^{-\mathtt{T}}-\mathtt{q}^{-\mathtt{S}})(1-\mathtt{k}^2\mathtt{q}^{-\mathtt{T}})}{\lambda_2}-\frac{\lambda_1\lambda_3(\lambda_2+\mathtt{k}^2\mathtt{q}^{-\mathtt{S}+\mathtt{N}}(\mathtt{q}^{-\mathtt{T}}+2r-1))}{\mathtt{q}^{\mathtt{N}}\lambda_2(\lambda_2(r^2+s^2)+2r\lambda_1+\lambda_0)}.
\end{aligned}\tag{8.8}$$

We observe that

$$\tilde{Q}(f_1(r,s),f_2(r,s))=-\frac{4\cdot\mathtt{q}^{-2\mathtt{N}}\lambda_3^2\lambda_1^2s^2}{(\lambda_2(r^2+s^2)+2r\lambda_1+\lambda_0)^2}<0,$$

and so f maps $\mathbb{H}$ in $\mathcal{D}'$. One directly checks that $f\circ\Omega$ and $\Omega\circ f$ are the identities on $\mathcal{D}'$ and $\mathbb{H}$ respectively, which shows that Ω has our desired properties.

Remark 8.2.1. Let us give some ideas about how the formula for Ω was discovered. Once the appropriate physical coordinates u,v for the system are found, which lead to the liquid region $\mathcal{D}'$ being an ellipse, one suspects that the map Ω should be given by the solution in $\mathbb{H}$ of *some* quadratic equation $a_2w^2+a_1w+a_0$, whose discriminant $D=a_1^2-4a_0a_2$ is negative precisely on $\mathcal{D}'$. In particular, we expect that $a_1^2-4a_0a_2=\lambda\tilde{Q}(u,v)$ for some positive parameter λ.

In [62] the complex structure for the uniform tiling case (this is $\kappa=0$ and $q=1$ in our model) was given by a quadratic equation, whose coefficients are *linear* in the coordinates of the system. By analogy we guess that $a_i=a_i^1u+a_i^2v+a_i^3$ for $i=1,2,3$ in our case as well, where the new coefficients do not depend on u and v. This gives us a nine parameter system.

When searching for a map Ω one has a choice of which point of the boundary of $\mathcal{D}'$ should be sent to infinity. In our case, we choose the boundary point at $(u,1+u\mathtt{k}^2\mathtt{q}^{-\mathtt{S}})$ with $u=1+\lambda_3\cdot\lambda_2^{-1}$ to be sent to infinity, which gives us 2 equations for our 9 parameters. In addition, the relationship $a_1^2-4a_0a_2=\lambda\tilde{Q}(u,v)$ gives an additional 6 equations (comparing the coefficients in front of u^iv^j) and an extra parameter λ. Overall we have a ten parameter system with eight equations.

The resulting system has a 2-parameters set of solutions. The extra freedom comes from multiplying a_2, a_1, a_0 by the same positive constant and also from multiplying a_2 and dividing a_0 by the same positive constant. Observe that the resulting complex structures are all equivalent modulo a multiplication by a positive constant, which is an automorphism of $\mathbb{H}$. Our particular, choice for the parameters is dictated by the *product form* of the coefficient a_0 in (8.6).

Remark 8.2.2. As mentioned in Section 1.1.2 there is a natural complex coordinate one can define on the liquid region $\mathcal{D}$, called *complex slope*. Let us explain how to construct it briefly – see [42,44] for more details. Suppose $(x, y) \in \mathcal{D}$ and set (p_1, p_2, p_3) to be the normal vector to the limit shape $\hat{h}$ at (x, y) such that $p_1 + p_2 + p_3 = 1$. Then the complex slope $z(x, y)$ is the unique point in $\mathbb{H}$ such that the triangle $(0, 1, z)$ has angles $(\pi p_1, \pi p_2, \pi p_3)$. In the case of the uniform tilings of the hexagon (and more general domains) it is known that there is an *algebraic* relationship between $z(x, y)$ and the complex structure $\Omega(x, y)$, whose pullback establishes the connection with the GFF on $\mathbb{H}$, [44,62]. For the q-Racah tiling model an expression for $z(x, y)$ was obtained in [16, Section 8.1] and it is related to $\Omega(x, y)$ from (8.7) as follows. If we set

$$U = U(x, y) = \frac{z(x, y)\mathtt{q}^x - \mathtt{k}^2\mathtt{q}^{-\mathtt{S}+2y}}{1 - z(x, y)\mathtt{k}^2\mathtt{q}^{-\mathtt{S}+2y-x}} \text{ and } \Omega = \Omega(x, y) \text{ then}$$

$$U = \frac{\Omega(\mathtt{q}^{-\mathtt{S}} - \mathtt{q}^{\mathtt{N}}) + (\mathtt{q}^{-\mathtt{S}} - 1)(\mathtt{q}^{-\mathtt{T}} - \mathtt{q}^{\mathtt{N}}) - \mathtt{k}^2\mathtt{q}^{\mathtt{N}}(\Omega + \mathtt{q}^{-\mathtt{T}} - \mathtt{q}^{\mathtt{N}})(\Omega\mathtt{q}^{-\mathtt{S}} + \mathtt{q}^{-\mathtt{T}-\mathtt{S}} - \mathtt{q}^{-\mathtt{T}})}{(\Omega\mathtt{q}^{\mathtt{N}} + \mathtt{q}^{-\mathtt{T}} - \mathtt{q}^{\mathtt{N}})(\Omega + \mathtt{q}^{-\mathtt{S}} - 1) - \mathtt{k}^2\mathtt{q}^{\mathtt{N}-\mathtt{T}}[(\mathtt{q}^{-\mathtt{S}} - \mathtt{q}^{\mathtt{N}})\Omega + (\mathtt{q}^{-\mathtt{T}} - \mathtt{q}^{\mathtt{N}})(\mathtt{q}^{-\mathtt{S}} - 1)]}.$$

We end the section with the following result that will be required in the next section.

Lemma 8.2.3. *Suppose that $u, v_1, v_2 \in \mathbb{R}$ are such that $(u, v_1), (u, v_2) \in \mathcal{D}'$. Then we have*

$$-\log\left|\frac{\Omega(u, v_1) - \Omega(u, v_2)}{\Omega(u, v_1) - \overline{\Omega}(u, v_2)}\right| = \log\left|\frac{\sqrt{(v_1 - a)(b - v_2)} + \sqrt{(v_2 - a)(b - v_1)}}{\sqrt{(v_1 - a)(b - v_2)} - \sqrt{(v_2 - a)(b - v_1)}}\right|, \tag{8.9}$$

where $a < b$ denote the intersection points of the vertical line through u with the ellipse $\tilde{Q}(u, v) = 0$.

Proof. Note that if $\phi(z)$ is an automorphism of $\mathbb{H}$, i.e. $\phi(z) = \frac{m \cdot z + n}{k \cdot z + l}$ with $m, n, k, l \in \mathbb{R}$ and $ml - nk = 1$, then the LHS of (8.9) is the same upon replacing $\Omega(u, v_i)$ with $\phi(\Omega(u, v_i))$ for $i = 1, 2$. Set

$$m = \frac{-a_1 + v_i\mathtt{q}^{\mathtt{N}}(\mathtt{q}^{-\mathtt{T}} - 1)}{\sqrt{2a_0(u, v_i)}},\ n = -\sqrt{2a_0(u, v_i)},\ k = \frac{1}{\sqrt{2a_0(u, v_i)}},\ l = 0, \tag{8.10}$$

and observe that the above do not change if we take $i = 1$ or 2. Moreover, by our choice of parameters we know that m, n, k, l satisfy the earlier conditions and we

let ϕ denote the automorphism corresponding to this quadruple. Setting $D(u,v_i) = a_1(u,v_i)^2 - 4a_0(u,v_i)a_2(u,v_i)$ we see that

$$\phi(\Omega(u,v_i)) = \frac{(-a_1 + v_i\mathtt{q}^{\mathtt{N}}(\mathtt{q}^{-\mathtt{T}}-1))(-a_1+\sqrt{D(u,v_i)}) - 4a_0a_1}{-a_1+\sqrt{D}} = v_i\mathtt{q}^{\mathtt{N}}(\mathtt{q}^{-\mathtt{T}}-1) - \sqrt{D},$$

where in the second equality we multiplied the numerator and denominator by $-a_1-\sqrt{D}$ and used that $a_1^2 - D = 4a_0a_2$. Recalling that $\sqrt{D(u,v_i)} = \mathtt{q}^{\mathtt{N}}\sqrt{\tilde{Q}(u,v_i)} = \mathtt{q}^{\mathtt{N}}(\mathtt{q}^{-\mathtt{T}} - 1)\sqrt{(v_i-a)(v_i-b)}$, where a,b are as in the statement of the lemma, we see that

$$\frac{\Omega(u,v_1)-\Omega(u,v_2)}{\Omega(u,v_1)-\overline{\Omega}(u,v_2)} = \frac{(v_1-v_2)+\sqrt{(v_1-a)(v_1-b)}-\sqrt{(v_2-a)(v_2-b)}}{(v_1-v_2)+\sqrt{(v_1-a)(v_1-b)}+\sqrt{(v_2-a)(v_2-b)}}.$$

Taking absolute value on both sides above and squaring we get

$$\left|\frac{\Omega(u,v_1)-\Omega(u,v_2)}{\Omega(u,v_1)-\overline{\Omega}(u,v_2)}\right|^2 = \frac{(b-v_1)(v_2-a)+(b-v_2)(v_1-a)-2\sqrt{(v_1-a)(b-v_1)(v_2-a)(b-v_2)}}{(b-v_1)(v_2-a)+(b-v_2)(v_1-a)+2\sqrt{(v_1-a)(b-v_1)(v_2-a)(b-v_2)}}.$$

If we take logarithms on both sides of the above and multiply the result by $-1/2$ we get (8.9). □

8.3. Proof of Theorems 7.2.2 and 7.2.4

8.3.1. Proof of Theorem 7.2.2

We suppose that we have a sequence ε_k, which converges to 0^+ and also sequences $q(\varepsilon_k), N(\varepsilon_k), T(\varepsilon_k), S(\varepsilon_k)$ and $\kappa(\varepsilon_k)$ as in Definition 7.2.1. Let us define $t(\varepsilon_k) = \lfloor x\varepsilon_k^{-1}\rfloor$ and observe that in this notation we have for all large k that

$$\varepsilon_k \cdot h\left(\lfloor x\varepsilon_k^{-1}\rfloor, \lfloor y\varepsilon_k^{-1}\rfloor + 1/2\right) = \varepsilon_k \cdot N \cdot \int_{\mathbb{R}} \mathbf{1}_{\left\{r<y\varepsilon_k^{-1}\right\}}\mu^t(dr), \tag{8.11}$$

where $\mu^t = N^{-1}\sum_{i=1}^N \delta\left(x_i^t/N\right)$. By possibly passing to a subsequence we may assume that the parameters t,S,T fall into one of the four cases in Theorem 7.3.5. These cases need to be handled separately, but as the arguments are analogous we assume that we are in the case $t < \min(S, T-S)$.

It follows from Theorem 7.3.5, (8.11) and the definition of $\hat{h}$ that for large k we have

$$p(\varepsilon_k) := \mathbb{P}_{\varepsilon_k}\left(\left|\varepsilon_k \cdot h\left(\lfloor x\varepsilon_k^{-1}\rfloor, \lfloor y\varepsilon_k^{-1}\rfloor + 1/2\right) - \hat{h}(x,y)\right| > \eta\right) = \mathbb{P}_N\left(\left|\varepsilon_k \cdot N \cdot \int_{\mathbb{R}} \mathbf{1}_{\left\{r<y\varepsilon_k^{-1}N^{-1}\right\}}\mu_N(dr) - \int_0^y \mu\left(r/\mathtt{N}\right)dr\right| > \eta\right), \tag{8.12}$$

where $\mathbb{P}_N$, μ_N are as in the statement of Theorem 7.4.4 for the parameters $q_N = q$, $M_N = t + N - 1$, $\alpha_N = q_N^{-S-N}$, $\beta_N = q_N^{S-T-N}$, $\gamma_N = q_N^{-t-N}$ and $\delta_N = \kappa^2 q_N^{-S+N}$ and μ is as in Definition 7.4.2 for the parameters $q = \mathtt{q}^{\mathtt{N}}$, $a = \mathtt{q}^{-\mathtt{S}-\mathtt{N}}$, $b = \mathtt{q}^{\mathtt{S}-\mathtt{T}-\mathtt{N}}$, $c = \mathtt{q}^{-x-\mathtt{N}}$ and $d = \mathtt{k}^2\mathtt{q}^{-\mathtt{S}+\mathtt{N}}$.

For $\tilde{\delta} > 0$ we let $f_{\tilde{\delta}}$ be a smooth function, such that $f_{\tilde{\delta}} = 1$ on $[0, y/\mathtt{N}]$, its support lies in $(-\tilde{\delta}, y/\mathtt{N} + \tilde{\delta})$, $f_{\tilde{\delta}}(x) \in [0,1]$ for all x. Then choosing $\tilde{\delta}$ sufficiently small we have for all large k that

$$p(\varepsilon_k) < \mathbb{P}_N \left(\left| \mathtt{N} \cdot \int_{\mathbb{R}} f_{\tilde{\delta}}(r) \mu_N(dr) - \mathtt{N} \cdot \int_{\mathbb{R}} f_{\tilde{\delta}}(r) \mu(r) dr \right| > \eta/2 \right),$$

where we used that $N = \mathtt{N}\epsilon_k^{-1} + o(1)$, $f_{\tilde{\delta}}(x)$, $\mu(x)$ are both in $[0,1]$ and we performed a change of variables for the integral involving μ. From Theorem 7.4.4 we know that the RHS above converges to 0 as $k \to \infty$, which proves the theorem.

8.3.2. Proof of Theorem 7.2.4

We suppose that we have a sequence ε_k, which converges to 0^+ and also sequences $q(\varepsilon_k), N(\varepsilon_k), T(\varepsilon_k), S(\varepsilon_k)$ and $\kappa(\varepsilon_k)$ as in Definition 7.2.1. In addition, we assume that R_i are real polynomials such that $R_i'(x) = f_i(x)$ for $i = 1, \ldots, m$.

Observe that for all large k we have

$$\int_{\mathbb{R}} \left(\mathcal{H}(q^{-t}, v) - \mathbb{E}_{\mathbb{P}_{\varepsilon_k}} \left[\mathcal{H}(q^{-t}, v) \right] \right) f_i(v) dv = \int_1^R \left(\mathcal{H}(q^{-t}, v) - \mathbb{E}_{\mathbb{P}_{\varepsilon_k}} \left[\mathcal{H}(q^{-t}, v) \right] \right) f_i(v) dv, \tag{8.13}$$

where $R = \mathtt{q}^{-\mathtt{S}-\mathtt{N}} + \mathtt{k}^2\mathtt{q}^{-\mathtt{S}-\mathtt{T}} + 1$. The latter truncation is allowed since a.s. all particles will have v-coordinate in $[1, R]$, which makes the height function $\mathcal{H}$ deterministic outside this interval and the above integrand zero there. In addition, we observe that

$$\int_1^R \mathcal{H}(q^{-t}, v) f_i(v) dv = \sum_{j=1}^N \int_{V_j}^{V_{j+1}} j \cdot f_i(v) dv = -\sum_{j=1}^N R_i(V_j) + N R_i(R), \tag{8.14}$$

where $V_j = V(t, j)$ for $j = 1, \ldots, N$ (see (7.6)) and $V_{N+1} = R$. Combining (8.13) and (8.14) we conclude that for all large k we have

$$\begin{aligned} &\int_{\mathbb{R}} \left(\mathcal{H}(q^{-t}, v) - \mathbb{E}_{\mathbb{P}_{\varepsilon_k}} \left[\mathcal{H}(q^{-t}, v) \right] \right) f_i(v) dv = \\ &- N \int_{\mathbb{R}} R_j(x) \rho^t(dx) + N \mathbb{E}_{\mathbb{P}_{\varepsilon_k}} \left[\int_{\mathbb{R}} R_j(x) \rho^t(dx) \right], \end{aligned} \tag{8.15}$$

where $\rho^t = \frac{1}{N}\sum_{j=1}^{N}\delta(V_j)$. By possibly passing to a subsequence we may assume that the parameters t, S, T fall into one of the four cases in Theorem 7.3.5. These cases need to be handled separately, but as the arguments are analogous we will assume that we are in the first case $t < \min(S, T-S)$.

It follows from Theorem 7.3.5 that ρ^t under law $\mathbb{P}_{\varepsilon_k}$ has the same distribution as ρ_N under law $\mathbb{P}_N$, where ρ_N is as in (7.18) and $\mathbb{P}_N$ is as in Definition 7.4.1 for the parameters $q_N = q$, $M_N = t+N-1$, $\alpha_N = q_N^{-S-N}$, $\beta_N = q_N^{S-T-N}$, $\gamma_N = q_N^{-t-N}$ and $\delta_N = \kappa^2 q_N^{-S+N}$. If we denote by $X_i^{\varepsilon_k}$ the RHS of (8.15) for $i = 1, \dots, m$ we conclude from Theorem 7.4.5 that $X_i^{\varepsilon_k}$ converge as $k \to \infty$ to a Gaussian vector $(X_1, \dots, X_m)$ which has zero mean and covariance

$$\mathbb{E}[X_i X_j] = \frac{1}{(2\pi i)^2}\oint_\gamma\oint_\gamma \frac{R_i(v_1)R_j(v_2)}{2(v_1-v_2)^2}\left(-1+\frac{(v_1-a)(v_2-b)+(v_2-a)(v_1-b)}{2\sqrt{v_1-a}\sqrt{v_1-b}\sqrt{v_2-a}\sqrt{v_2-b}}\right)dv_1dv_2,$$

where γ is a positively oriented contour, which encloses the interval $[1+u\mathtt{k}^2\mathtt{q}^{-\mathtt{S}}, u\mathtt{q}^{-\mathtt{N}}+\mathtt{k}^2\mathtt{q}^{-\mathtt{S}+\mathtt{N}}]$ and the square roots are defined with respect to the principal branch of the logarithm. In deriving the above we implicitly used Lemma 7.4.3 and (8.5). To complete the proof it suffices to show

$$\begin{aligned}&\frac{\pi}{(2\pi i)^2}\oint_\gamma\oint_\gamma \frac{R_i(v_1)R_j(v_2)}{2(v_1-v_2)^2}\left(-1+\frac{(v_1-a)(v_2-b)+(v_2-a)(v_1-b)}{2\sqrt{v_1-a}\sqrt{v_1-b}\sqrt{v_2-a}\sqrt{v_2-b}}\right)dv_1dv_2 =\\ &\int_a^b\int_a^b R_i'(x)R_j'(y)\left(-\frac{1}{2\pi}\log\left|\frac{\Omega(u,x)-\Omega(u,y)}{\Omega(u,x)-\overline{\Omega}(u,y)}\right|\right)dxdy, \text{where } \Omega \text{ is as in (8.7).}\end{aligned} \tag{8.16}$$

We start with the LHS of (8.16) and deform the v_2 contour so that it traverses the segment $[a,b]$ once in the positive and once in the negative direction. Observe the square roots are purely imaginary and come with opposite sign when we approach $[a,b]$ from the upper and lower half-planes. On the other hand, the term $\frac{1}{2(v_1-v_2)^2}$ cancels when we integrate over $[a,b]$ in the positive and negative direction. By Cauchy's theorem we do not change the value of the integral during the deformation and so from the Bounded convergence theorem we see that the LHS of (8.16) equals

$$\frac{\pi i}{(2\pi i)^2}\oint_\gamma\int_a^b \frac{R_i(v_1)R_j(v_2)}{(v_1-v_2)^2}\cdot\frac{(v_1-a)(v_2-b)+(v_2-a)(v_1-b)}{2\sqrt{v_1-a}\sqrt{v_1-b}\sqrt{v_2-a}\sqrt{b-v_2}}dv_2dv_1.$$

We integrate by parts in the v_2 variable and change the order of the integrals, which leads to the following expression for the LHS in (8.16)

$$\frac{\pi i}{(2\pi i)^2}\int_a^b \oint_\gamma \frac{R_i(v_1)R_j'(v_2)}{(v_1-v_2)}\cdot\frac{\sqrt{v_2-a}\sqrt{b-v_2}}{\sqrt{v_1-a}\sqrt{v_1-b}}dv_1dv_2. \tag{8.17}$$

At this time we claim that for each $v_2\in(a,b)$ we have

$$\oint_\gamma \frac{R_i(v_1)}{(v_1-v_2)}\cdot\frac{\sqrt{v_2-a}\sqrt{b-v_2}}{\sqrt{v_1-a}\sqrt{v_1-b}}dv_1 = 2i\int_a^b R_i'(v_1)\times$$
$$\left[2\log\left(\sqrt{(v_1-a)(b-v_2)}+\sqrt{(v_2-a)(b-v_1)}\right)-\log|v_1-v_2|-\log(b-a)\right]dv_1. \tag{8.18}$$

We will prove (8.18) below. For now we assume its validity and finish the proof of (8.16).

From (8.17) and (8.18) we see that to show (8.16) it suffices to have

$$-\log\left|\frac{\Omega(u,v_1)-\Omega(u,v_2)}{\Omega(u,v_1)-\overline{\Omega}(u,v_2)}\right| = 2\log\left(\sqrt{(v_1-a)(b-v_2)}+\sqrt{(v_2-a)(b-v_1)}\right)- \\ -\log|v_1-v_2|-\log(b-a). \tag{8.19}$$

From (8.9) we know that

$$-\log\left|\frac{\Omega(u,v_1)-\Omega(u,v_2)}{\Omega(u,v_1)-\overline{\Omega}(u,v_2)}\right| = \log\left|\frac{\sqrt{(v_2-a)(b-v_1)}+\sqrt{(v_1-a)(b-v_2)}}{\sqrt{(v_2-a)(b-v_1)}-\sqrt{(v_1-a)(b-v_2)}}\right|.$$

In addition, one readily checks that $\log|v_1-v_2|+\log(b-a)$ is equal to

$$\log\left|\sqrt{(v_2-a)(b-v_1)}+\sqrt{(v_1-a)(b-v_2)}\right|+$$
$$\log\left|\sqrt{(v_2-a)(b-v_1)}-\sqrt{(v_1-a)(b-v_2)}\right|.$$

The last two statements imply (8.19), which concludes the proof of (8.16).

In the remainder of the section we establish (8.18). Fix $v_2\in(a,b)$ and let $\varepsilon>0$ be such that $(v_2-\varepsilon,v_2+\varepsilon)\subset[a,b]$. For $\delta\in(0,\varepsilon)$ we define the contour $\Gamma_{\delta,\varepsilon}$ as follows. $\Gamma_{\delta,\varepsilon}$ starts from the point $b-i\delta$ and follows the circle centered at b with radius δ counterclockwise until the point $b+i\delta$, afterwards it goes to the left along the segment connecting the points $b+i\delta$ and $v_2+\varepsilon+i\delta$; it follows the circle centered at $v_2+i\delta$ and radius ε counterclockwise until the point $v_2-\varepsilon+i\delta$ and goes to the left along the segment connecting $v_2-\varepsilon+i\delta$ and $a+i\delta$; it then follows the circle centered at a with radius δ counterclockwise until the point $a-i\delta$ and then goes to the right along the segment connecting $a-i\delta$ and $v_2-\varepsilon-i\delta$; finally, it follows the circle centered at $v_2-i\delta$ and radius ε counterclockwise until the point $v_2-\varepsilon-i\delta$ and goes to the right along the segment connecting $v_2-\varepsilon-i\delta$ and $b-i\delta$, see Fig. 8.

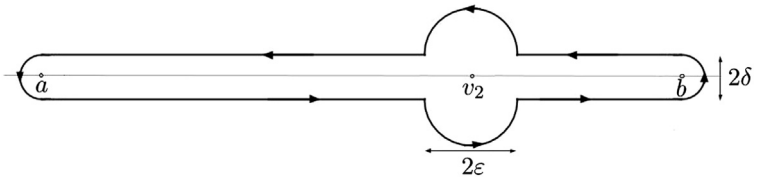

Fig. 8. The contour $\Gamma_{\delta,\varepsilon}$.

By Cauchy's theorem we see that

$$\oint_{\gamma} \frac{R_i(v_1)}{(v_1-v_2)} \cdot \frac{\sqrt{v_2-a}\sqrt{b-v_2}}{\sqrt{v_1-a}\sqrt{v_1-b}} dv_1 = \oint_{\Gamma_{\delta,\varepsilon}} \frac{R_i(v_1)}{(v_1-v_2)} \cdot \frac{\sqrt{v_2-a}\sqrt{b-v_2}}{\sqrt{v_1-a}\sqrt{v_1-b}} dv_1 \tag{8.20}$$

We next let δ go to 0^+ and see that

$$\lim_{\delta\to 0^+} \oint_{\Gamma_{\delta,\varepsilon}} \frac{R_i(v_1)}{(v_1-v_2)} \cdot \frac{\sqrt{v_2-a}\sqrt{b-v_2}}{\sqrt{v_1-a}\sqrt{v_1-b}} dv_1 = T_1(\varepsilon) + T_2(\varepsilon) + T_3(\varepsilon), \text{ where}$$

$$T_1(\varepsilon) = 2i \int_a^{v_2-\varepsilon} \frac{R_i(v_1)}{(v_1-v_2)} \cdot \frac{\sqrt{v_2-a}\sqrt{b-v_2}}{\sqrt{v_1-a}\sqrt{b-v_1}} dv_1,$$

$$T_2(\varepsilon) = 2i \int_{v_2+\varepsilon}^{b} \frac{R_i(v_1)}{(v_1-v_2)} \cdot \frac{\sqrt{v_2-a}\sqrt{b-v_2}}{\sqrt{v_1-a}\sqrt{b-v_1}} dv_1,$$

$$T_3(\varepsilon) = \int_{C_\varepsilon^+(v_2)} \frac{R_i(v_1)}{(v_1-v_2)} \cdot \frac{\sqrt{v_2-a}\sqrt{b-v_2}}{\sqrt{v_1-a}\sqrt{v_1-b}} dv_1 + \int_{C_\varepsilon^-(v_2)} \frac{R_i(v_1)}{(v_1-v_2)} \cdot \frac{\sqrt{v_2-a}\sqrt{b-v_2}}{\sqrt{v_1-a}\sqrt{v_1-b}} dv_1 \tag{8.21}$$

with $C_\varepsilon^+(v_2), C_\varepsilon^-(v_2)$ being positively oriented half-circles of radius ε around v_2 in the upper and lower half-planes respectively. In deriving the above expression we used the Bounded convergence theorem and the fact that the square roots are purely imaginary and come with opposite sign when we approach $[a,b]$ from the upper and lower half-planes.

We next integrate by parts the integrals in $T_1(\varepsilon)$ and $T_2(\varepsilon)$ to get

$$T_1(\varepsilon) = -2i \int_a^{v_2-\varepsilon} R_i'(u) G_v(v_1) dv_1 + 2i \cdot [R_i(v_2-\varepsilon) G_{v_2}(v_2-\varepsilon) - R_i(a) G_{v_2}(a)],$$

$$T_2(\varepsilon) = -2i \int_{v_2+\varepsilon}^{b} R_i'(v_1) G_{v_2}(v_1) dv_1 + 2i \cdot [R_i(b) G_{v_2}(b) - R_i(v_2+\varepsilon) G_{v_2}(v_2+\varepsilon)], \text{ where}$$

$$G_{v_2}(v_1) = -2\log\left(\sqrt{(v_1-a)(b-v_2)} + \sqrt{(v_2-a)(b-v_1)}\right) + \log|v_1-v_2|. \tag{8.22}$$

We observe that $G_{v_2}(a) = -\log(b-a) = G_{v_2}(b)$ and $R_i(v_2-\varepsilon)G_{v_2}(v_2-\varepsilon) - R_i(v_2+\varepsilon)G_{v_2}(v_2+\varepsilon) = O\left(\varepsilon\log\varepsilon^{-1}\right)$. The latter statements together with the Dominated convergence theorem imply that

$$\lim_{\varepsilon\to 0^+} T_1(\varepsilon) + T_2(\varepsilon) = -2i\int_a^b R_i'(v_1)G_{v_2}(v_1)dv_1 - 2i\log(b-a)\cdot[R(b)-R(a)]. \tag{8.23}$$

We next turn to $T_3(\varepsilon)$ and parametrize $C_\varepsilon^+(v)$ through $v_1 = v_2 + \varepsilon e^{i\theta}$ with $\theta\in(0,\pi)$ and $C_\varepsilon^-(v)$ through $v_1 = v_2+\varepsilon e^{i\theta}$ with $\theta\in(-\pi,0)$. This leads to

$$T_3(\varepsilon) = i\sqrt{v_2-a}\sqrt{b-v_2}\cdot\left[\int_0^\pi \frac{R_i(v_2+\varepsilon e^{i\theta})d\theta}{\sqrt{\varepsilon e^{i\theta}+v_2-a}\sqrt{\varepsilon e^{i\theta}+v_2-b}} + \int_{-\pi}^0 \frac{R_i(v_2+\varepsilon e^{i\theta})d\theta}{\sqrt{\varepsilon e^{i\theta}+v_2-a}\sqrt{\varepsilon e^{i\theta}+v_2-b}}\right].$$

We can let ε converge to 0^+ above, which by the Bounded convergence theorem implies

$$\lim_{\varepsilon\to 0^+} T_3(\varepsilon) = \sqrt{v_2-a}\sqrt{b-v_2}\cdot\left[\int_0^\pi \frac{R_i(v_2)d\theta}{\sqrt{v_2-a}\sqrt{b-v_2}} - \int_{-\pi}^0 \frac{R_i(v_2)d\theta}{\sqrt{v_2-a}\sqrt{b-v_2}}\right] = 0, \tag{8.24}$$

where the sign change came from the fact that we are approaching the real line from the upper and lower half-planes in the two cases. Combining (8.20), (8.21), (8.23), (8.24) we conclude that

$$\oint_\gamma \frac{R_i(v_1)}{(v_1-v_2)}\cdot\frac{\sqrt{v_2-a}\sqrt{b-v_2}}{\sqrt{v_1-a}\sqrt{v_1-b}}dv_1 = -2i\int_a^b R_i'(v_1)G_{v_2}(v_1)dv_1 - 2i\log(b-a)\cdot[R(b)-R(a)].$$

The latter is equivalent to (8.18) once we use that $R(b)-R(a) = \int_a^b R'(v_1)dv_1$.

8.4. Conjectural 2d fluctuations

In this section we isolate the following two-dimensional (conjectural) extension to Theorem 7.2.4.

Conjecture 8.4.1. *Assume the same notation as in Theorem 7.2.4. Then the centered random height function*

$$\sqrt{\pi}\left(\mathcal{H}(q^{-t},v)-\mathbb{E}_{\mathbb{P}_\varepsilon}\left[\mathcal{H}(q^{-t},v)\right]\right)$$

converges to the pullback of the Gaussian free field with Dirichlet boundary conditions on the upper half-plane $\mathbb{H}$ with respect to the map Ω in the following sense: For any set of polynomials $f_i \in \mathbb{R}[x]$, numbers $u_i \in (1, \mathfrak{q}^{-\mathfrak{r}})$ and sequences $t_i(\epsilon)$ such that $q^{-t_i(\epsilon)} = u_i + O(\epsilon)$ for $i = 1, \dots, m$ the joint distribution of

$$\int_{\mathbb{R}} \sqrt{\pi}\left(\mathcal{H}(q^{-t_i},v)-\mathbb{E}_{\mathbb{P}_\varepsilon}\left[\mathcal{H}(q^{-t_i},v)\right]\right) f_i(v)dv, \qquad i=1,\dots,m, \tag{8.25}$$

converges to the joint distribution of the similar averages

$$\int_{a(u_i)}^{b(u_i)} \mathcal{F}(\Omega(u,y))f_i(y)dy, \qquad i=1,\dots,m$$

of the pullback of the GFF. In the above formula $a(u), b(u)$ are the v-coordinates of the two points where the vertical line through u intersects the ellipse $\tilde{Q}(u,v)=0$.

Equivalently, the variables in (8.25) converge jointly to a Gaussian vector $(X_1, \dots, X_m)$ with mean zero and covariance

$$\mathbb{E}[X_iX_j] = \int_{a(u_i)}^{b(u_i)}\int_{a(u_j)}^{b(u_j)} f_i(x)f_j(y)\left(-\frac{1}{2\pi}\log\left|\frac{\Omega(u_i,x)-\Omega(u_j,y)}{\Omega(u_i,x)-\overline{\Omega}(u_j,y)}\right|\right)dxdy. \tag{8.26}$$

Remark 8.4.2. We emphasize that our methods only allow us to study the global fluctuations of the tiling model for a *single* vertical section, and in order to establish the above statement one needs to be able to study the *joint* distribution of the particles on *several* vertical slices.

9. Connection to log-gases on a quadratic lattice

In this section we explain how our model fits into the framework of a discrete log-gas on a quadratic lattice as in Section 2. The latter will allow us to deduce Theorems 7.4.4 and 7.4.5 as consequences of Theorems 3.1.1 and 5.2.7 respectively.

9.1. Asymptotics of the weight function

We first consider with the weight function $w^{qR}(x)$ of the q-Racah ensemble defined by (7.11). We are interested in understanding the asymptotic behavior of $w^{qR}(x)$ when the parameters $\alpha, \beta, \gamma, \delta, q$ scale as in Definition 7.4.1. In order to do this we will need the following technical lemma.

Lemma 9.1.1. *Let $a, b \in (0,1)$ and $\mathtt{q} \in [a,b]$ be given. Suppose that x_N, q_N for $N \in \mathbb{N}$ are sequences such that $\left|q_N - \mathtt{q}^{1/N}\right| \leq AN^{-2}$ and $x_N \in [0, q_N]$, where A is a positive constant. Then we have*

$$(x_N; q_N)_\infty = \exp\left(N \cdot \frac{\mathrm{Li}(x_N)}{\log(\mathtt{q})} + O\left(\log(N)\right)\right), \tag{9.1}$$

where $\mathrm{Li}(x) = \sum_{k=1}^{\infty} \frac{x^k}{k^2}$ is the dilogarithm function and the constant in the big O notation depends on a, b and A.

Proof. Taking logarithm of $\prod_{i=1}^{\infty}(1 - x_N q_N^{i-1})$ and power expanding $\log(1-y)$ for $0 < y < 1$ gives

$$\log\left[(x_N; q_N)_\infty\right] = -\sum_{r=1}^{\infty}\sum_{k=1}^{\infty} \frac{x_N^k}{k} \cdot q_N^{k(r-1)}.$$

We change the order of the sums and use the geometric series formula to get

$$\log\left[(x_N; q_N)_\infty\right] = -\sum_{k=1}^{\infty} \frac{x_N^k}{k} \cdot \frac{1}{1-q_N^k} = -\frac{1}{1-q_N}\sum_{k=1}^{\infty} \frac{x_N^k}{k} \cdot \frac{1-q_N}{1-q_N^k} = A_N + B_N + C_N,$$

where

$$A_N = -\frac{1}{1-q_N}\sum_{k=1}^{\infty} \frac{x_N^k}{k}\left[\frac{1-q_N}{1-q_N^k} - \frac{1}{k}\right],\ B_N = -\frac{1}{1-q_N}\sum_{k=1}^{\infty} \frac{x_N^k}{k^2} - N\frac{\mathrm{Li}(x_N)}{\log(\mathtt{q})},$$
$$C_N = N\frac{\mathrm{Li}(x_N)}{\log(\mathtt{q})}.$$

What we need to show is that A_N and B_N are both $O\left(\log(N)\right)$.

Notice that

$$|B_N| = \mathrm{Li}(x_N)\left|-\frac{1}{1-q_N} - \frac{N}{-\log(\mathtt{q})}\right| \leq \mathrm{Li}(1) + O(1),$$

where we used that $x_N \in [0,1]$ and $1 - q_N = -\log(\mathtt{q})/N + O(N^{-2})$. This proves that $B_N = O(1)$ and we focus on A_N for the remainder.

Combining $\frac{1-q_N}{1-q_N^k} - \frac{1}{k} \geq 0$ with $x_N \in [0, q_N]$ we conclude

$$-\frac{1}{1-q_N}\sum_{k=1}^{\infty} \frac{q_N^k}{k}\left[\frac{1-q_N}{1-q_N^k} - \frac{1}{k}\right] \geq A_N \geq 0.$$

Since $1 - q_N = -\log(\mathtt{q})/N + O(N^{-2})$, we see that what remains to be shown is that

$$\sum_{k=1}^{\infty} \frac{q_N^k}{k} \left[\frac{1-q_N}{1-q_N^k} - \frac{1}{k} \right] = O\left(\frac{\log(N)}{N} \right). \tag{9.2}$$

Suppose that $-\frac{1}{4\log(\mathfrak{q})} \geq \varepsilon_0 > 0$ is sufficiently small so that when $\varepsilon \in [0, -\log(\mathfrak{q})\varepsilon_0]$ we have $1-e^{-\varepsilon} \geq \varepsilon - 2\varepsilon^2 \geq \varepsilon/2$. Using the latter together with the inequality $\frac{1-q_N}{1-q_N^k} - \frac{1}{k} \geq 0$ we see that

$$\sum_{k \leq \varepsilon_0 N} \frac{q_N^k}{k} \left[\frac{1-q_N}{-(k/N)\log(\mathfrak{q}) - 2\log(\mathfrak{q})^2 (k/N)^2} - \frac{1}{k} \right] \geq \sum_{k \leq \varepsilon_0 N} \frac{q_N^k}{k} \left[\frac{1-q_N}{1-q_N^k} - \frac{1}{k} \right] \geq 0.$$

As $1-q_N = -\log(\mathfrak{q})/N + O(N^{-2})$ the above statement implies that for sufficiently large C we have

$$\frac{C}{N} \sum_{k \leq \varepsilon_0 N} \frac{q_N^k}{k} + \sum_{k \leq \varepsilon_0 N} \frac{q_N^k}{k} \left[\frac{-\log(\mathfrak{q})}{-k\log(\mathfrak{q}) - 2\log(\mathfrak{q})^2 k^2 N^{-1}} - \frac{1}{k} \right] \geq$$
$$\sum_{k \leq \varepsilon_0 N} \frac{q_N^k}{k} \left[\frac{1-q_N}{1-q_N^k} - \frac{1}{k} \right] \geq 0.$$

Notice that

$$\begin{aligned} \frac{-\log(\mathfrak{q})}{k\log(\mathfrak{q}) + 2\log(\mathfrak{q})^2 k^2 N^{-1}} - \frac{1}{k} &= \frac{1}{k + 2\log(\mathfrak{q}) k^2 N^{-1}} - \frac{1}{k} \\ &= -\frac{2\log(\mathfrak{q}) N^{-1}}{1 + 2\log(\mathfrak{q}) k N^{-1}} \leq -4\log(\mathfrak{q}) N^{-1}, \end{aligned}$$

where the last inequality holds since $kN^{-1} \leq \varepsilon_0 \leq -\frac{1}{4\log(q)}$. The latter estimates show that for some (possibly different than before) constant $C > 0$ we have

$$\frac{C}{N} \sum_{k \leq \varepsilon_0 N} \frac{q_N^k}{k} \geq \sum_{k \leq \varepsilon_0 N} \frac{q_N^k}{k} \left[\frac{1-q_N}{1-q_N^k} - \frac{1}{k} \right] \geq 0.$$

Since $\sum_{k \leq \varepsilon_0 N} \frac{q_N^k}{k} \leq -\log(1-q_N) = O(\log(N))$ we conclude that

$$\sum_{k \leq \varepsilon_0 N} \frac{q_N^k}{k} \left[\frac{1-q_N}{1-q_N^k} - \frac{1}{k} \right] = O\left(\frac{\log(N)}{N} \right). \tag{9.3}$$

We next have that

$$\frac{1-q_N}{1-\mathfrak{q}^{\varepsilon_0}} \sum_{k \geq \varepsilon_0 N} \frac{q_N^k}{k} \geq \sum_{k \geq \varepsilon_0 N} \frac{q_N^k}{k} \cdot \frac{1-q_N}{1-q_N^k} \geq \sum_{k \geq \varepsilon_0 N} \frac{q_N^k}{k} \left[\frac{1-q_N}{1-q_N^k} - \frac{1}{k} \right] \geq 0.$$

Since $1-q_N = O(N^{-1})$ and $\sum_{k \geq \varepsilon_0 N} \frac{q_N^k}{k} \leq -\log(1-q_N) = O(\log(N))$ we conclude that

$$\sum_{k\geq \varepsilon_0 N}^{\infty} \frac{q_N^k}{k}\left[\frac{1-q_N}{1-q_N^k}-\frac{1}{k}\right] = O\left(\frac{\log(N)}{N}\right). \tag{9.4}$$

Combining (9.3) and (9.4) we conclude (9.2), which proves the lemma. □

In addition, we require the following alternative formula for the weight $w^{qR}(x)$.

Lemma 9.1.2. *Suppose that we have parameters $\alpha, \beta, \gamma, \delta, q$ and M as in Definition 7.3.4. Then we have*

$$\begin{aligned} w^{qR}(x) &= \frac{(\beta\delta q, \gamma\delta q; q)_\infty}{(q, \alpha^{-1}\gamma\delta q, \beta^{-1}\gamma q, \delta q, \alpha^{-1}, \gamma^{-1}; q)_\infty}\tilde{w}^{qR}(x), \text{ where} \\ \tilde{w}^{qR}(x) &= \\ (\gamma/\beta)^x q^{x^2} &\frac{1-\gamma\delta q^{2x+1}}{1-\gamma\delta q}\frac{(q^{x+1}, \alpha^{-1}\gamma\delta q^{x+1}, \beta^{-1}\gamma q^{x+1}, \delta q^{x+1}, \alpha^{-1}q^{-x}, \gamma^{-1}q^{-x}; q)_\infty}{(\beta\delta q^{x+1}, \gamma\delta q^{x+1}; q)_\infty}. \end{aligned} \tag{9.5}$$

Proof. We recall the definition of $w^{qR}(x)$ from Definition 7.3.1 for the reader's convenience.

$$w^{qR}(x) = \frac{(\alpha q, \beta\delta q, \gamma q, \gamma\delta q; q)_x}{(q, \alpha^{-1}\gamma\delta q, \beta^{-1}\gamma q, \delta q; q)_x}\cdot\frac{(1-\gamma\delta q^{2x+1})}{(\alpha\beta q)^x(1-\gamma\delta q)}.$$

Observe that

$$\begin{aligned} (\alpha q; q)_x &= \prod_{i=1}^{x}(1-\alpha q^i) = q^{x(x+1)/2}\alpha^x(-1)^x\cdot\prod_{i=1}^{x}(1-\alpha^{-1}q^{-i}) \\ &= q^{x(x+1)/2}\alpha^x(-1)^x\cdot(\alpha^{-1}q^{-x}; q)_x. \end{aligned}$$

Similarly, we have $(\gamma q; q)_x = q^{x(x+1)/2}\gamma^x(-1)^x\cdot(\gamma^{-1}q^{-x}; q)_x$. Substituting the latter identities and performing a bit of cancellation we arrive at

$$w^{qR}(x) = \frac{(\alpha^{-1}q^{-x}, \beta\delta q, \gamma^{-1}q^{-x}, \gamma\delta q; q)_x}{(q, \alpha^{-1}\gamma\delta q, \beta^{-1}\gamma q, \delta q; q)_x}\frac{(1-\gamma\delta q^{2x+1})}{(1-\gamma\delta q)}\cdot(\gamma/\beta)^x q^{x^2}.$$

Observe that for $a\in[0,1)$ we have $(a;q)_x = \frac{(a;q)_\infty}{(aq^x;q)_\infty}$. Substituting the latter identity in the above expression we see that $w^{qR}(x)$ equals

$$\begin{aligned} &\frac{(\alpha^{-1}q^{-x}, \beta\delta q, \gamma^{-1}q^{-x}, \gamma\delta q; q)_\infty}{(q, \alpha^{-1}\gamma\delta q, \beta^{-1}\gamma q, \delta q; q)_\infty}\frac{(q^{x+1}, \alpha^{-1}\gamma\delta q^{x+1}, \beta^{-1}\gamma q^{x+1}, \delta q^{x+1}; q)_\infty}{(\alpha^{-1}, \beta\delta q^{x+1}, \gamma^{-1}, \gamma\delta q^{x+1}; q)_\infty} \\ &\times\frac{(1-\gamma\delta q^{2x+1})}{(1-\gamma\delta q)}\cdot(\gamma/\beta)^x q^{x^2}. \end{aligned}$$

From here (9.5) is immediate. □

The main statement of this section is the following.

Lemma 9.1.3. *Assume that we have the same notation as in Definition 7.4.1 and let* $\tilde{w}_N^{qR}(x)$ *be as in (9.5) with parameters* $\alpha_N, \beta_N, \gamma_N, \delta_N, q_N$. *Then, for* $x \in \{0, \dots, M_N\}$ *we have the following asymptotic expansion of* $\tilde{w}_N^{qR}(x)$

$$\begin{aligned}
&\tilde{w}_N^{qR}(x) = \exp\left(-NV\left(\frac{x}{N}; \mathtt{a}, \mathtt{b}, \mathtt{c}, \mathtt{d}\right) + O\left(\log(N)\right)\right), \text{ where}\\
&V(s; \mathtt{a}, \mathtt{b}, \mathtt{c}, \mathtt{d}) = \frac{1}{-\log \mathtt{q}}\Big[\mathrm{Li}(\mathtt{q}^s) + \mathrm{Li}(\mathtt{a}^{-1}\mathtt{c}\mathtt{d}\mathtt{q}^s) + \mathrm{Li}(\mathtt{b}^{-1}\mathtt{c}\mathtt{q}^s) + \mathrm{Li}(\mathtt{d}\mathtt{q}^s) +\\
&+ \mathrm{Li}(\mathtt{a}^{-1}\mathtt{q}^{-s}) + \mathrm{Li}(\mathtt{c}^{-1}\mathtt{q}^{-s}) - \mathrm{Li}(\mathtt{b}\mathtt{d}\mathtt{q}^s) - \mathrm{Li}(\mathtt{c}\mathtt{d}\mathtt{q}^s) + \log(\mathtt{q})^2 s^2 + s\log(\mathtt{q})\log(\mathtt{c}/\mathtt{b})\Big]
\end{aligned} \tag{9.6}$$

and the constant in the big O notation depends on the parameters A *and* $\mathtt{a}, \mathtt{b}, \mathtt{c}, \mathtt{d}, \mathtt{q}, \mathtt{M}$ *and is uniform as the latter vary over compact subsets of* $\mathtt{P}$ *(recall that* A *and* $\mathtt{P}$ *were given in Definition 7.4.1).*

Proof. Using Lemma 9.1.1 we have that

$$\begin{aligned}
\tilde{w}^{qR}(x) = (\gamma/\beta)^x q^{x^2} \frac{1-\gamma\delta q^{2x+1}}{1-\gamma\delta q} \cdot \exp\Big(\frac{N}{\log q}\Big[&\mathrm{Li}(q^{x+1}) + \mathrm{Li}(\alpha^{-1}\gamma\delta q^{x+1}) +\\
&+ \mathrm{Li}(\beta^{-1}\gamma q^{x+1}) + \mathrm{Li}(\delta q^{x+1}) + \mathrm{Li}(\alpha^{-1}q^{-x}) + \mathrm{Li}(\gamma^{-1}q^{-x}) - \mathrm{Li}(\beta\delta q^{x+1}) -\\
&- \mathrm{Li}(\gamma\delta q^{x+1})\Big] + O(\log(N))\Big),
\end{aligned}$$

where for brevity we suppressed the dependence of the parameters on N. Since $\mathtt{c}/\mathtt{b} = \gamma/\beta + O(N^{-1})$, $q = \mathtt{q}^{1/N} + O(N^{-2})$ and $\gamma\delta q$ is bounded away from 1 we see that

$$(\gamma/\beta)^x q^{x^2} \frac{1-\gamma\delta q^{2x+1}}{1-\gamma\delta q} = \exp\left(N \cdot \frac{x}{N} \cdot \log(\mathtt{c}/\mathtt{b}) + N \cdot \frac{x^2}{N^2} \cdot \log(\mathtt{q}) + O(1)\right).$$

This handles the first factor in $\tilde{w}^{qR}(x)$ and we only need to match the dilogarithms with (9.6).

We claim that if $x, y \in [0,1]$ and $C > 0$ are such that $|x-y| \le CN^{-1}$ then

$$|\mathrm{Li}(x) - \mathrm{Li}(y)| \le C \cdot \frac{\log(N)+1}{N} + \frac{1}{N}. \tag{9.7}$$

It is clear that applying (9.7) to each of the dilogarithms in $\tilde{w}^{qR}(x)$ we can match the corresponding ones in (9.6) upto an error of order $\log(N)$. Thus, to prove the lemma it suffices to show (9.7).

Without loss of generality suppose that $x \le y$ and set $\varepsilon = y - x$. Then we have

$$0 \le \mathrm{Li}(y) - \mathrm{Li}(x) = \sum_{k=1}^{\infty} \frac{y^k - (y-\varepsilon)^k}{k^2} \le \sum_{k=1}^{\infty} \frac{1-(1-\varepsilon)^k}{k^2},$$

where we used that the functions $y^k-(y-\varepsilon)^k$ are increasing on $[\varepsilon,1]$ for $k\geq 1$. In addition,

$$\sum_{k=1}^{N}\frac{1-(1-\varepsilon)^k}{k^2}\leq\sum_{k=1}^{N}\frac{k\varepsilon}{k^2}\leq C\cdot\frac{\log(N)+1}{N}\text{ and }\sum_{k=N+1}^{\infty}\frac{1-(1-\varepsilon)^k}{k^2}$$
$$\leq\sum_{k=N+1}^{\infty}\frac{1}{k^2}\leq\int_{N}^{\infty}\frac{dz}{z^2}=\frac{1}{N}.$$

Combining the above inequalities we conclude (9.7) and hence the lemma. □

9.2. Proof of Theorems 7.4.4 and 7.4.5

Our first task is to verify that $\mathbb{P}_N(\lambda_1,\ldots,\lambda_N)$ from Definition 7.4.1 satisfy Assumptions 1–4 and 6–7 in Section 2.2. In view of (9.5) we have the following alternative representation for $\mathbb{P}_N(\lambda_1,\ldots,\lambda_N)$

$$\mathbb{P}_N(\lambda_1,\ldots,\lambda_N)=$$
$$\frac{\mathbf{1}_{(0\leq\lambda_1<\lambda_2<\cdots<\lambda_N\leq M_N)}}{\tilde{Z}(N,M_N,\alpha_N,\beta_N,\gamma_N,\delta_N)}\prod_{1\leq i<j\leq N}\left(\sigma_N(q_N^{-\lambda_i})-\sigma_N(q_N^{-\lambda_j})\right)^2\prod_{i=1}^{N}\tilde{w}_N^{qR}(\lambda_i),\quad(9.8)$$

where $\tilde{Z}$ is a new normalization constant, $\sigma_N(z)=z+u_Nz^{-1}$ with $u_N=\gamma_N\delta_Nq_N$ and $\tilde{w}_N^{qR}$ is as in (9.5) for the parameters $\alpha_N,\beta_N,\gamma_N,\delta_N,q_N$ and M_N.

If we set $\ell_i=q_N^{-\lambda_i}+u_Nq_N^{\lambda_i}$ for $i=1,\ldots,N$ then we see that the induced law on particles ℓ_i for $i=1,\ldots,N$ from (9.8) agrees with (2.3) for $\theta=1$. Specifically, we are in the single-cut case with $a_1(N)=0$, $b_1(N)=M_N-N+2$ and setting $w(\ell_i;N):=\tilde{w}_N^{qR}(\lambda_i)$ we have

$$\mathbb{P}_N(\ell_1,\ldots,\ell_N)=Z_N^{-1}\cdot\mathbf{1}_{\{(\ell_1,\ldots,\ell_N)\in\mathfrak{X}_N^1\}}\cdot\prod_{1\leq i<j\leq N}(\ell_i-\ell_j)^2\prod_{i=1}^{N}w(\ell_i;N).\quad(9.9)$$

It is clear that Assumptions 1 and 3 in Section 2.2 are satisfied in this case. In addition, in view of Lemma 9.1.3, we know that Assumption 2 holds for the function $V(x)=V(\sigma_{\mathtt{q}}^{-1}(x);\mathtt{a},\mathtt{b},\mathtt{c},\mathtt{d})$, where $V(\cdot;\mathtt{a},\mathtt{b},\mathtt{c},\mathtt{d})$ is as in (9.6) and $\sigma_{\mathtt{q}}(x)=\mathtt{q}^{-x}+\mathtt{u}\mathtt{q}^x$ with $\mathtt{u}=\mathtt{cd}$.

Let Φ_N^+,Φ_N^- be as in Definition 7.4.2 with parameters $a=\alpha_N,b=\beta_N,c=\gamma_N,d=\delta_N$ and $q=q_N^N$. Observe that $\Phi_N^{\pm}$ satisfy Assumptions 4 and 6 in Section 2.2. In particular, we have that $\Phi_N^{\pm}$ converge to $\Phi_\infty^{\pm}$, where the latter are as in Definition 7.4.2 for the parameters $a=\mathtt{a}$, $b=\mathtt{b}$, $c=\mathtt{c}$, $d=\mathtt{d}$ and $q=\mathtt{q}^{\mathtt{N}}$. With the same choice of parameters we also define R_∞ and Q_∞ as in that definition and (7.16). Finally, one checks that Assumption 7 holds from the definition of V and $\Phi_\infty^{\pm}$.

Since Assumptions 1–4 and 6 hold we can apply Theorem 4.1.1 and obtain that the function $\tilde{R}_N$ on the right side of (4.3) is a degree 4 polynomial. We isolate the following asymptotic statement about $\tilde{R}_N$, which will be used here and whose proof is the focus of Section 10.

Fact 9.2.1. (see (10.30)) If R_∞ and $\tilde{R}_N$ are as above then for all $z \in \mathbb{C}$

$$\lim_{N\to\infty} R_\infty(z) - \tilde{R}_N(z) = 0. \tag{9.10}$$

Proof. (Theorem 7.4.4) As discussed earlier we know that $\mathbb{P}_N(\ell_1, \dots, \ell_N)$ in (9.9) satisfies Assumptions 1–4 and 6–7 in Section 2.2. We conclude from Theorem 3.1.1 that the empirical measures

$$\rho_N = \frac{1}{N}\sum_{i=1}^{N} \delta\left(\sigma_N(q_N^{-\lambda_i})\right)$$

converge to a limiting measure ρ, which by Lemma 4.2.2 has density

$$\rho(y_0 + \mathtt{u}y_0^{-1}) = \frac{1}{\log(\mathtt{q})\pi(y_0 - \mathtt{u}y_0^{-1})} \cdot \arccos\left(\frac{R_\rho(y_0)}{2\sqrt{\Phi_\infty^-(y_0)\Phi_\infty^+(y_0)}}\right) \text{ for } y_0 \in [1, \mathtt{c}]. \tag{9.11}$$

Combining (4.10) and (9.10) we conclude that $R_\rho = R_\infty$. The latter implies that the measures μ_N in Theorem 7.4.4 satisfy the conditions in that theorem for the measure $\mu := \rho \circ \sigma_{\mathtt{q}}^{-1}$, which in view of (9.11) agrees with Definition 7.4.2. □

Proof. (Theorem 7.4.5) From the proof of Theorem 7.4.4 we know that $R_\rho = R_\infty$, and so $Q_\rho^2 = R_\rho^2 - 4\Phi_\infty^+\Phi_\infty^- = Q_\infty^2$. The formula for Q_∞ in (7.16) implies that Q_ρ satisfies Assumption 5 for $r_1 = a_-$, $s_1 = a_+$ and $H(z) = z(z^2 - \mathtt{cd})(\mathtt{abq}^{\mathtt{N}} - \mathtt{q}^{-\mathtt{N}})$. Overall, the measures (9.9) satisfy Assumptions 1–6 in Section 2.2 and so Theorem 7.4.5 follows from Theorem 5.2.7 and Remark 5.2.5. □

10. Proof of Fact 9.2.1

The goal of this section is to prove Fact 9.2.1, which is the missing ingredient necessary to complete the proofs of Theorems 7.4.4 and 7.4.5. We summarize some basic facts about discrete Riemann–Hilbert problems and q-Racah orthogonal polynomials in Section 10.1. In Section 10.2 we introduce a matrix-valued function $A_N(z)$ and derive some of its properties. Section 10.3 contains some asymptotic results about $A_N(z)$, which suffice to show Fact 9.2.1.

10.1. Discrete Riemann–Hilbert problems and orthogonal polynomials

10.1.1. Discrete Riemann–Hilbert problems

In this section we relate solutions of discrete Riemann–Hilbert problems (DRHP) for jump matrices of a special type to orthogonal polynomials. Our exposition closely follows that in [12, Section 2], which in turn dates back to [8,9].

Let $\mathfrak{X}$ be a finite subset of $\mathbb{C}$ such that $\mathrm{card}(\mathfrak{X}) = M+1 < \infty$ and $w : \mathfrak{X} \to \mathrm{Mat}(2,\mathbb{C})$ be any function. We say that an analytic function

$$m\colon \mathbb{C} \setminus \mathfrak{X} \to \mathrm{Mat}(2,\mathbb{C})$$

solves the DRHP$(\mathfrak{X}, w)$ if m has simple poles at the points of $\mathfrak{X}$ and its residues at these points are given by the *jump* (or *residue*) condition

$$\operatorname*{Res}_{\zeta=x} m(\zeta) = \lim_{\zeta\to x} m(\zeta)w(\zeta),\ x \in \mathfrak{X}. \tag{10.1}$$

We will assume that the matrix $w(x)$ depends on a function $\omega : \mathfrak{X} \to \mathbb{C}$ and has the form

$$w(x) = \begin{bmatrix} 0 & \omega(x) \\ 0 & 0 \end{bmatrix}. \tag{10.2}$$

Recall that a collection $\{P_n(\zeta)\}_{n=0}^{N}$ of complex polynomials is called the collection of *orthogonal polynomials associated to the weight function* ω if

- P_n is a polynomial of degree n for all $n = 1,\dots,M$ and $P_0 \equiv const$;
- if $m \neq n$ then $\sum_{x\in\mathfrak{X}} P_m(x)P_n(x)\omega(x) = 0$.

We will always take P_n to be monic, i.e. $P_n(x) = x^n +$ lower terms.

We consider the following inner product on the space $\mathbb{C}[\zeta]$ of all complex polynomials:

$$(f(\zeta), g(\zeta))_\omega = \sum_{x\in\mathfrak{X}} f(x)g(x)\omega(x).$$

It is clear that there exists a collection of orthogonal polynomials $\{P_n(\zeta)\}_{n=0}^{M}$ associated to ω such that $(P_n, P_n)_\omega \neq 0$ for all $n = 0,\dots,M$ if and only if the restriction of $(\cdot,\cdot)_\omega$ to the space $\mathbb{C}[\zeta]^{\leq d}$ of polynomials of degree at most d is non-degenerate for all $d = 0,\dots,M$. If this condition holds we say that the function ω is *nondegenerate*, and then it is clear that the collection $\{P_n(\zeta)\}_{n=0}^{M}$ is unique. For convenience we isolate the following notation

$$c_n := (P_n, P_n)_\omega\,, \qquad H_n(\zeta) := \sum_{x\in\mathfrak{X}} \frac{P_n(x)\omega(x)}{\zeta - x}, \qquad n = 0,\dots,M. \tag{10.3}$$

The connection between the orthogonal polynomials $\{P_n(\zeta)\}_{n=0}^M$ and solutions to DRHP$(\mathfrak{X}, w)$ is detailed in the following statement.

Theorem 10.1. *[12, Lemma 2.1 and Theorem 2.4] Let $\mathfrak{X}$ be a finite subset of $\mathbb{C}$ such that card$(\mathfrak{X}) = M + 1 < \infty$ and w be as in (10.2) with $\omega\colon \mathfrak{X} \to \mathbb{C}$ a nondegenerate weight function. Then for any $k = 1, 2, \dots, M$ the DRHP$(\mathfrak{X}, w)$ has a unique solution $m_{\mathfrak{X}}(\zeta)$, satisfying an asymptotic condition*

$$m_{\mathfrak{X}}(\zeta) \cdot \begin{bmatrix} \zeta^{-k} & 0 \\ 0 & \zeta^k \end{bmatrix} = I + O\left(\zeta^{-1}\right) \text{ as } \zeta \to \infty, \tag{10.4}$$

where I is the identity matrix. This solution is explicitly given by

$$m_{\mathfrak{X}}(\zeta) = \begin{bmatrix} P_k(\zeta) & H_k(\zeta) \\ c_{k-1}^{-1} P_{k-1}(\zeta) & c_{k-1}^{-1} H_{k-1}(\zeta) \end{bmatrix}, \text{ with } c_n, H_n \text{ as in (10.3)}$$

and satisfies $\det m_{\mathfrak{X}}(\zeta) \equiv 1$.

10.1.2. q-Racah polynomials

In this section we recall and establish some basic properties of the q-Racah orthogonal polynomials, cf. [47, Section 3.2]. Recall from Definition 7.3.1 that the q-Racah weight function is defined on $\mathfrak{X} = \{q^{-x} + \gamma\delta q^{x+1} : x = 0, \dots, M\}$ as

$$\omega^{qR}(q^{-x} + \gamma\delta q^{x+1}) = \frac{(\alpha q, \beta\delta q, \gamma q, \gamma\delta q; q)_x}{(q, \alpha^{-1}\gamma\delta q, \beta^{-1}\gamma q, \delta q; q)_x} \frac{(1 - \gamma\delta q^{2x+1})}{(\alpha\beta q)^x (1 - \gamma\delta q)}. \tag{10.5}$$

We assume the parameters are as in Definition 7.3.4. It is well known that ω^{qR} is a nondegenerate weight function and the orthogonal polynomials $\{P_n(\zeta)\}_{n=0}^M$ associated to it are the q-Racah orthogonal polynomials. Explicitly, they are given by

$$\begin{aligned} P_n(q^{-x} + \gamma\delta q^{x+1}) &= {}_4\tilde{\phi}_3\left(\begin{matrix} q^{-n}, \alpha\beta q^{n+1}, q^{-x}, \gamma\delta q^{x+1} \\ \alpha q, \beta\delta q, \gamma q \end{matrix} \middle| q; q \right), \text{ where} \\ {}_4\tilde{\phi}_3\left(\begin{matrix} a_1, a_2, a_3, a_4 \\ b_1, b_2, b_3 \end{matrix} \middle| q; z \right) &= \sum_{n=0}^{M} \frac{(a_1, a_2, a_3, a_4; q)_k}{(b_1, b_2, b_3; q)_k} \frac{z^k}{(q; q)_k}. \end{aligned} \tag{10.6}$$

The q-Racah polynomials P_n satisfy the following orthogonality relation.

$$\sum_{x=0}^{M} \omega^{qR}(q^{-x} + \gamma\delta q^{x+1}) P_m(q^{-x} + \gamma\delta q^{x+1}) P_n(q^{-x} + \gamma\delta q^{x+1}) = c_n \cdot \delta_{mn}, \text{ where}$$

$$c_n = \frac{(\gamma\delta q^2, \alpha^{-1}\beta^{-1}\gamma, \alpha^{-1}\delta, \beta^{-1}; q)_\infty}{(\alpha^{-1}\gamma\delta q, \beta^{-1}\gamma q, \delta q, \alpha^{-1}\beta^{-1}q^{-1}; q)_\infty} \frac{(1 - \alpha\beta q)(\gamma\delta q)^n}{(1 - \alpha\beta q^{2n+1})} \frac{(q, \beta q, \alpha\delta^{-1}q, \alpha\beta\gamma^{-1}q; q)_n}{(\alpha\beta q, \alpha q, \beta\delta q, \gamma q; q)_n}. \tag{10.7}$$

In addition, setting $y(x) = P_n(q^{-x}+\gamma\delta q^{x+1})$ we have the following q-Difference equations

$$\begin{aligned}
&q^{-n}(1-q^n)(1-\alpha\beta q^{n+1})y(x) = \\
&\hat{B}(x)y(x+1) - [\hat{B}(x)+\hat{D}(x)]y(x) + \hat{D}(x)y(x-1) \text{ where} \\
&\hat{B}(x) = \frac{(1-\alpha q^{x+1})(1-\beta\delta q^{x+1})(1-\gamma q^{x+1})(1-\gamma\delta q^{x+1})}{(1-\gamma\delta q^{2x+1})(1-\gamma\delta q^{2x+2})}, \text{ and} \\
&\hat{D}(x) = \frac{q(1-q^x)(1-\delta q^x)(\beta-\gamma q^x)(\alpha-\gamma\delta q^x)}{(1-\gamma\delta q^{2x})(1-\gamma\delta q^{2x+1})}.
\end{aligned} \tag{10.8}$$

We next prove a couple of easy facts about the q-Racah polynomials. For convenience we set $\sigma(z) = z + uz^{-1}$ with $u = \gamma\delta q$.

Proposition 10.2. *Each $P_n(\zeta)$ with $M \geq n \geq 0$ has n roots in the interval $(1+u, q^{-M}+uq^M)$.*

Proof. Let m be the number of roots of $P_n(\zeta)$ in the interval $[1+u, q^{-M}+uq^M]$ counted with multiplicities. Since $\deg(P_n) = n$, we know that $m \leq n$. If $m < n$ let $x_1, \ldots, x_m$ be an enumeration of these roots in some order. By (10.7) we know that

$$\sum_{x=0}^{M} P_n(q^{-x}+uq^x)\omega^{qR}(q^{-x}+uq^x)\cdot\prod_{i=1}^{m}(q^{-x}+uq^x-x_i) = 0. \tag{10.9}$$

Note that the polynomial $P_n(\zeta)\cdot\prod_{i=1}^m(\zeta-x_i)$ does not change its sign on $[1+u, q^{-M}+uq^M]$ and so all of the above summands must be zero. But then $P_n(q^{-x}+uq^x) = 0$ for $x = 0, \ldots, M$ and so $M+1 \leq m < n$, contradicting $M \geq n$. We conclude that $m = n$.

Let $x_1, \ldots, x_m$ be the roots of $P_n(\zeta)$ that are not equal to $1+u$. By our work above we know that $x_i \in (1+u, q^{-M}+uq^M]$ for $i = 1, \ldots, m$. If again we suppose that $m < n$ then $P_n(\zeta)\cdot\prod_{i=1}^m(\zeta-x_i)$ does not change its sign on $[1+u, q^{-M}+uq^M]$ and so all the terms in the analogous sum (10.9) must be zero, implying $P_n(q^{-x}+uq^x) = 0$ for $x = 0, \ldots, M$. We conclude that P_n has roots at $q^{-x}+uq^x$ for $x = 0, \ldots, M$. But this is impossible, since $M \geq n$ leading to too many roots of P_n. We reach a contradiction, which arose from our assumption that $1+u$ is a root of $P_n(\zeta)$. A similar argument shows $q^{-M}+uq^M$ is also not a root of $P_n(\zeta)$. □

Lemma 10.3. *Let $M \geq n \geq 1$ be given. If $t > 1$ is such that $P_n(\sigma(t)) = 0$ then $P_n(\sigma(q^{-1}t)) \neq 0$.*

Proof. Suppose that $\alpha_1, \ldots, \alpha_n$ are the roots of $P_n(z)$. Then (10.8) can be written as

$$[A(z)+B(z)+C_n(z)]\,Q_n(z) = q^n A(z)Q_n(q^{-1}z) + q^{-n}B(z)Q_n(qz), \text{ where}$$

$$Q_n(z) = \prod_{i=1}^{n}(z^2-\alpha_i z+u), \; A(z) = (z-\alpha q)(z-\beta\delta q)(z-\gamma q)(z-\gamma\delta q)(z^2-\gamma\delta),$$

$$B(z) = q(z-1)(z-\delta)(z\beta-\gamma)(z\alpha-\gamma\delta)(z^2-\gamma\delta q^2),$$
$$C_n(z) = (z^2-\gamma\delta)(z^2-\gamma\delta q)(z^2-\gamma\delta q^2)(q^{-n}-1)(1-\alpha\beta q^{n+1}). \tag{10.10}$$

From Proposition 10.2 we know that $\alpha_i \in (1+u, q^{-M}+uq^M)$ for $i = 1, \dots, n$. Consequently, $z^2 - \alpha_i z + u$ has two real roots $q^{-M} > \beta_i > 1$ and $1 > u\beta_i^{-1} > 0$.

Suppose that $P_n(\sigma(t)) = 0$ for some $t > 1$. Let $k \geq 1$ be the maximal integer such that $P_n(\sigma(tq^{-i})) = 0$ for $i = 0, \dots, k-1$ and assume for the sake of contradiction that $k \geq 2$. Then we know that tq^{-k+1} is a root of $Q_n(z)$ and $Q_n(qz)$. From the top line in (10.10) we conclude that $A(tq^{-k+1})Q_n(tq^{-k}) = 0$. By the maximality of k we must have $A(tq^{-k+1}) = 0$. Since $tq^{-k+1} > 1$ we conclude that $tq^{-k+1} = \alpha q$ and then as $tq^{-k+1} = \beta_j$ for some $n \geq j \geq 1$ we conclude that $q^{-M} > \alpha q$ or $\gamma > \alpha$. This contradicts Definition 7.3.4 and so $k = 1$ as desired. □

We end this section with a lemma, which classifies the bounded degree polynomials $A(z), B(z), C(z)$ that satisfy the q-Difference equation (10.10).

Lemma 10.4. *Fix $M \geq n > 7$ and suppose that $\tilde{A}(z), \tilde{B}(z), \tilde{C}(z)$ are polynomials, each of degree at most $n-7$ or zero and such that*

$$\left[\tilde{A}(z) + \tilde{B}(z) + \tilde{C}(z)\right] P_n(\sigma(z)) = \tilde{A}(z)P_n(\sigma(q^{-1}z)) + \tilde{B}(z)P_n(\sigma(qz)). \tag{10.11}$$

Then

$$\tilde{B}(z)A(z) = B(z)\tilde{A}(z) \text{ and } \tilde{B}(z)C_n(z) = B(z)\tilde{C}(z), \tag{10.12}$$

where $A(z), B(z), C_n(z)$ are as in (10.10).

Proof. Using the notation from the proof of Lemma 10.3 we can alternatively rewrite (10.11) as

$$\left[\tilde{A}(z) + \tilde{B}(z) + \tilde{C}(z)\right] Q_n(z) = q^n\tilde{A}(z)Q_n(q^{-1}z) + q^{-n}\tilde{B}(z)Q_n(qz).$$

We multiply the above by $B(z)$ and subtract it from the top line of (10.10) multiplied by $\tilde{B}(z)$

$$\left[\tilde{B}(z)(A(z)+C_n(z)) - B(z)(\tilde{A}(z)+\tilde{C}(z))\right] Q_n(z) = q^n[\tilde{B}(z)A(z) - B(z)\tilde{A}(z)]Q_n(q^{-1}z).$$

As discussed in the proof of Lemma 10.3 we know that $Q_n(z)$ has n roots $\beta_1, \dots, \beta_n > 1$ and neither is a root of $Q_n(q^{-1}z)$. This implies that $\prod_{i=1}^n (z-\beta_i)$ divides $\tilde{B}(z)A(z) - B(z)\tilde{A}(z)$ and as the latter is of degree at most $n-1$ or zero, we conclude that it must be zero. This suffices for the proof. □

10.2. The matrix $A_N(z)$

We go back to the notation of Section 10.1.1. Let w be as in (10.2) with $\omega \equiv \omega^{qR}$ as in (10.5). From Theorem 10.1 we know DRHP$(\mathfrak{X}, w)$ has a unique solution $m_N(\zeta)$ with

$$m_N(\zeta) \cdot \begin{bmatrix} \zeta^{-N} & 0 \\ 0 & \zeta^N \end{bmatrix} = I + O\left(\zeta^{-1}\right) \text{ as } \zeta \to \infty.$$

For the sequel we let $\sigma(z) = z + uz^{-1}$ with $u = \gamma\delta q$ and define the following matrix-valued function that will play a central role in the arguments that follow

$$A_N(z) := m_N\left(\sigma(qz)\right) \cdot D(z) \cdot m_N^{-1}\left(\sigma(z)\right), \text{ where } D(z) = \begin{bmatrix} \Phi^-(z) & 0 \\ 0 & \Phi^+(z) \end{bmatrix}, \text{ where} \tag{10.13}$$

$$\Phi^+(z) = (z-\alpha)(z-\beta\delta)(z-\gamma)(z-\gamma\delta) \text{ and } \Phi^-(z) = (z-1)(\alpha z - \gamma\delta)(\beta z - \gamma)(z-\delta). \tag{10.14}$$

The significance of the functions $\Phi^{\pm}$ is as follows. If $\mathbb{P}_N$ denotes the measure $\mathbb{P}^{qR}$ as in Definition 7.3.3 then $\mathbb{P}_N$ satisfies the Nekrasov's equation, Theorem 4.1.1, for the functions $\Phi^{\pm}$, $\theta = 1$, $k = 1$ and $a_1 = 0$, $b_1 = M - N + 2$. In particular, if $\tilde{R}$ is given by

$$\tilde{R}(z) = \Phi^-(z) \cdot \mathbb{E}_{\mathbb{P}_N}\left[\prod_{i=1}^{N} \frac{\sigma(qz) - \ell_i}{\sigma(z) - \ell_i}\right] + \Phi^+(z) \cdot \mathbb{E}_{\mathbb{P}_N}\left[\prod_{i=1}^{N} \frac{\sigma(z) - \ell_i}{\sigma(qz) - \ell_i}\right], \tag{10.15}$$

then $\tilde{R}$ is a degree four polynomial. Fact 9.2.1 concerns the asymptotic behavior of $\tilde{R}$ when the parameters $\alpha, \beta, \gamma, \delta, q$ and M scale as in Definition 7.4.1 and we can relate it to the matrix A_N through the following result.

Lemma 10.5. *If $A_N(z)$ and $\tilde{R}(z)$ are as in (10.13) and (10.15) respectively then*

$$Tr(A_N(z)) = \tilde{R}(z). \tag{10.16}$$

Proof. We first recall from [3, Theorem 2.13] that

$$\mathbb{E}_{\mathbb{P}_N}\left[\prod_{i=1}^{N} \frac{x - \ell_i}{y - \ell_i}\right] = c_{N-1}^{-1} \det \begin{bmatrix} H_{N-1}(y) & H_N(y) \\ P_{N-1}(x) & P_N(x) \end{bmatrix},$$

with c_n, H_n, P_n as in (10.3), (10.6) and (10.7).

Combining the above with (10.15) we obtain

$$\begin{aligned}\tilde{R}(z) =& c_{N-1}^{-1} \cdot \Phi^{-}(z) \cdot [H_{N-1}(\sigma(z))P_N(\sigma(qz)) - H_N(\sigma(z))P_{N-1}(\sigma(qz))] + \\ & c_{N-1}^{-1} \cdot \Phi^{+}(z) \cdot [H_{N-1}(\sigma(qz))P_N(\sigma(z)) - H_N(\sigma(qz))P_{N-1}(\sigma(z))]\,.\end{aligned} \tag{10.17}$$

On the other hand, from (10.13) and Theorem 10.1 we know

$$\begin{aligned}A_N(z) = &\begin{bmatrix} P_N(\sigma(qz)) & H_N(\sigma(qz)) \\ c_{N-1}^{-1}P_{N-1}(\sigma(qz)) & c_{N-1}^{-1}H_{N-1}(\sigma(qz)) \end{bmatrix} \cdot \begin{bmatrix} \Phi^{-}(z) & 0 \\ 0 & \Phi^{+}(z) \end{bmatrix} \cdot \\ &\begin{bmatrix} c_{N-1}^{-1}H_{N-1}(\sigma(z)) & -H_N(\sigma(z)) \\ -c_{N-1}^{-1}P_{N-1}(\sigma(z)) & P_N(\sigma(z)) \end{bmatrix},\end{aligned} \tag{10.18}$$

where we also used that $\det m_N(\zeta) \equiv 1$. The trace in (10.18) matches the rights side in (10.17). □

In the remainder of this section we establish several properties about the matrix A_N.

Proposition 10.6. *The matrix $A_N(z)$ is entire.*

Proof. The result and its proof are analogous to [12, Proposition 3.3]. Appealing to (10.18) it is clear that the only possible singularities of $A_N(z)$ are simple poles at $z = q^{-x}$ for $x = 0, 1, \dots, M+1$ and $z = uq^y$ for $y = -1, 0, \dots, M$.

From (10.1) we have for z near q^{-x} and $x = 1, \dots, M+1$

$$m_N(\sigma(qz)) = F_1(z)\left(I + \frac{q^{-x}}{(q^{-(x-1)} - uq^{(x-1)})(z - q^{-x})} w\left(q^{-x+1} + uq^{x-1}\right)\right), \tag{10.19}$$

where $F_1(z)$ is an analytic, invertible matrix-valued function defined in a neighborhood of q^{-x}. By definition of w we have $w(q^1 + uq^{-1}) = 0$ and so (10.19) holds near $z = 1$ as well.

For z near q^{-x} and $x = 0, \dots, M$ we can similarly write

$$m_N^{-1}(\sigma(z)) = \left(I - \frac{q^{-x}}{(q^{-x} - uq^{x})(z - q^{-x})} w(q^{-x} + uq^{x})\right) F_2(z), \tag{10.20}$$

where $F_2(z)$ is an analytic, invertible matrix-valued function defined in a neighborhood of q^{-x}. By definition of w we have $w(q^{-M-1} + uq^{M+1}) = 0$ and so (10.20) holds near $z = q^{-M-1}$ as well.

Overall, to show that the matrix $A_N(z)$ is analytic at q^{-x} for $x = 0, 1, \dots, M+1$ it suffices for

$$\begin{aligned}X(z) := &\left(I + \frac{q^{-x}}{(q^{-(x-1)} - uq^{(x-1)})(z - q^{-x})} w(\sigma(qx))\right) \\ &\times D(x)\left(I - \frac{q^{-x}}{(q^{-x} - uq^{x})(z - q^{-x})} w(x)\right),\end{aligned}$$

to be analytic near those points. Since

$$X(z) = \begin{bmatrix} \Phi^-(x) & \frac{1}{(z-q^{-x})}\left(\Phi^+(x)\cdot\frac{q^{-x}}{(q^{-(x-1)}-uq^{(x-1)})}\omega^{qR}(\sigma(qx)) - \frac{q^{-x}}{(q^{-x}-uq^x)}\omega^{qR}(\sigma(x))\Phi^-(x)\right) \\ 0 & \Phi^+(x) \end{bmatrix}$$

$$\text{and } \frac{\Phi^-(x)}{\Phi^+(x)} = \frac{(q^{-x}-uq^x)\omega^{qR}(\sigma(qx))}{\omega^{qR}(\sigma(x))(q^{-(x-1)}-uq^{(x-1)})},$$

we conclude that $X(z)$ is indeed analytic near q^{-x} for $x = 0, 1, \dots, M$.

One verifies the analyticity of $A_N(z)$ at the points uq^y for $y = -1, 0, \dots, M$ analogously. □

Proposition 10.7. *Let $A_N^{ij}(z)$ denote the entry of $A_N(z)$ at the i-th row and j-th column. Then $A_N^{11}(z), A_N^{22}(z)$ are degree four polynomials and $A_N^{12}(z), A_N^{21}(z)$ are degree three polynomials. If $A_N^{ij}(z) = \sum_k a_{N,k}^{ij} z^k$ then*

$$a_{N,4}^{11} = \alpha\beta q^N, a_{N,4}^{22} = \alpha\beta q^{-N}, a_{N,3}^{12} = c_{N-1}^{-1}c_N[-q^N\alpha\beta + q^{-N}], a_{N,3}^{21} = [q^{-N+1}\alpha\beta - q^{N-1}] \tag{10.21}$$

Proof. From equation (10.18) we have

$$\begin{aligned} A_N^{11}(z) =& c_{N-1}^{-1}\left[P_N(\sigma(qz))H_{N-1}(\sigma(z))\Phi^-(z) - H_N(\sigma(qz))P_{N-1}(\sigma(z))\Phi^+(z)\right], \\ A_N^{22}(z) =& c_{N-1}^{-1}\left[-P_{N-1}(\sigma(qz))H_N(\sigma(z))\Phi^-(z) + H_{N-1}(\sigma(qz))P_N(\sigma(z))\Phi^+(z)\right], \\ A_N^{12}(z) =& c_{N-1}^{-1}\left[-P_N(\sigma(qz))H_N(\sigma(z))\Phi^-(z) + H_N(\sigma(qz))P_N(\sigma(z))\Phi^+(z)\right], \\ A_N^{21}(z) =& c_{N-1}^{-1}\left[P_{N-1}(\sigma(qz))H_{N-1}(\sigma(z))\Phi^-(z) - H_{N-1}(\sigma(qz))P_{N-1}(\sigma(z))\Phi^+(z)\right]. \end{aligned} \tag{10.22}$$

We know from Proposition 10.6 that $A_N^{ij}(z)$ are all entire functions. In addition, by Theorem 10.1 we know that $c_{N-1}^{-1}P_N(\sigma(qz))H_{N-1}(\sigma(z)) = \left[(\sigma(qz))^{-N}P_N(\sigma(qz))\right]\cdot\left[(\sigma(z))^N c_{N-1}^{-1}H_{N-1}(\sigma(z))\right]\cdot\left[(\sigma(z))^{-N}(\sigma(qz))^N\right] \sim q^N$ as $|z| \to \infty$. Analogous arguments show that

$$\begin{aligned} &c_{N-1}^{-1}H_{N-1}(\sigma(qz))P_N(\sigma(z)) \sim q^{-N}, \\ &H_N(\sigma(qz))P_{N-1}(\sigma(z)) = O\left(|z|^{-2}\right) = P_{N-1}(\sigma(qz))H_N(\sigma(z)). \end{aligned}$$

The above show that $A_N^{11}(z) = O(|z|^4) = A_N^{22}(z)$ and by Liouville's theorem we conclude that they are at most degree 4 polynomials. Our work above also shows that as $|z| \to \infty$ we have

$$A_N^{11}(z) \sim \alpha\beta q^N z^4 \text{ and } A_N^{22}(z) \sim \alpha\beta q^{-N} z^4,$$

which establishes the first two equations in (10.21).

Also as a consequence of Theorem 10.1 we have that $A_N^{12}(z) = O(|z|^3) = A_N^{21}(z)$, which by Liouville's theorem implies that they are at most degree three polynomials. What remains is to show that their leading coefficients are as in (10.21).

We first note by definition

$$\begin{aligned}\zeta^N H_N(\zeta) &= \zeta^N \sum_{x=0}^{M} \frac{P_N(q^{-x}+uq^x)\omega^{qR}(q^{-x}+uq^x)}{\zeta - q^{-x} - uq^x} = \\ &= \zeta^N \sum_{x=0}^{M} P_N(q^{-x}+uq^x)\omega^{qR}(q^{-x}+uq^x) \\ &\quad \cdot \left[\frac{1}{\zeta} + \frac{[q^{-x}+uq^x]}{\zeta^2} + \cdots + \frac{[q^{-x}+uq^x]^N}{\zeta^{N+1}} + O(\zeta^{-N-2})\right].\end{aligned}$$

Using (10.7) we get

$$\sum_{x=0}^{M} P_N(q^{-x}+uq^x)w^{qR}(q^{-x}+uq^x)[q^{-x}+uq^x]^i = 0 \text{ for } i = 0,\dots,N-1 \text{ and}$$

$$\sum_{x=0}^{M} P_N(q^{-x}+uq^x)w^{qR}(q^{-x}+uq^x)[q^{-x}+uq^x]^N =$$

$$\sum_{x=0}^{N} P_N(q^{-x}+uq^x)^2 w^{qR}(q^{-x}+uq^x) = c_N.$$

The above implies that as $|z| \to \infty$ we have

$$\begin{aligned}P_N(\sigma(qz))H_N(\sigma(z)) &= \left[(\sigma(qz))^{-N}P_N(\sigma(qz))\right] \cdot \left[(\sigma(z))^N H_N(\sigma(z))\right] \\ &\quad \cdot \left[(\sigma(z))^{-N}(\sigma(qz))^N\right] \sim z^{-1}c_N q^N.\end{aligned}$$

Analogously, $H_N(\sigma(qz))P_N(\sigma(z)) \sim z^{-1}c_N q^{-N}$, $P_{N-1}(\sigma(qz))H_{N-1}(\sigma(z)) \sim z^{-1}c_{N-1} \times q^{-N+1}$, $H_{N-1}(\sigma(qz))P_{N-1}(\sigma(z)) \sim z^{-1}c_{N-1}q^{N-1}$. The latter identities and (10.22) imply

$$a_{N,3}^{12} = c_{N-1}^{-1}[-c_N q^N \alpha\beta + c_N q^{-N}] \text{ and } a_{N,3}^{21} = c_{N-1}^{-1}[c_{N-1}q^{-N+1}\alpha\beta - c_{N-1}q^{N-1}],$$

which establishes the last two equations in (10.21). □

We end this section by relating $A_N(z)$ and the polynomials $A(z), B(z), C_n(z)$ in the q-Difference equation (10.10). From (10.13) we know that $A_N(z) \cdot m_N(\sigma(z)) = m_N(\sigma(qz)) \cdot D(z)$, and identifying the left two entries of the 2×2 matrices on both sides we arrive at

$$\begin{aligned} A_N^{11}(z)P_N(\sigma(z)) + c_{N-1}^{-1}A_N^{12}(z)P_{N-1}(\sigma(z)) &= \Phi^-(z)P_N(\sigma(qz)), \\ A_N^{21}(z)P_N(\sigma(z)) + c_{N-1}^{-1}A_N^{22}(z)P_{N-1}(\sigma(z)) &= c_{N-1}^{-1}\Phi^-(z)P_{N-1}(\sigma(qz)). \end{aligned} \tag{10.23}$$

Expressing $c_{N-1}^{-1}P_{N-1}(\cdot)$ from the first equation and substituting it in the second we obtain

$$\begin{aligned} &A_N^{21}(z)P_N(\sigma(z)) + A_N^{22}(z)\left[\frac{\Phi^-(z)P_N(\sigma(qz)) - A_N^{11}(z)P_N(\sigma(z))}{A_N^{12}(z)}\right] = \\ &\Phi^-(z)\cdot\left[\frac{\Phi^-(qz)P_N(\sigma(q^2z)) - A_N^{11}(qz)P_N(\sigma(qz))}{A_N^{12}(qz)}\right]. \end{aligned} \tag{10.24}$$

Replacing z with $q^{-1}z$, and reorganizing terms we obtain

$$\begin{aligned} &P_N(\sigma(z))\left[\frac{\Phi^-(q^{-1}z)A_N^{22}(q^{-1}z)}{A_N^{12}(q^{-1}z)} + \frac{\Phi^-(q^{-1}z)A_N^{11}(z)}{A_N^{12}(z)}\right] = \\ &P_N(\sigma(q^{-1}z))\left[\frac{A_N^{22}(q^{-1}z)A_N^{11}(q^{-1}z) - A_N^{12}(q^{-1}z)A_N^{21}(q^{-1}z)}{A_N^{12}(q^{-1}z)}\right] + \\ &P_N(\sigma(qz))\left[\frac{\Phi^-(q^{-1}z)\Phi^-(z)}{A_N^{12}(z)}\right]. \end{aligned}$$

Note that the first numerator of the second line above is $\det A_N(q^{-1}z)$, which we know to equal $\Phi^+(q^{-1}z)\Phi^-(q^{-1}z)$ (recall that $\det m_N = 1$ by Theorem 10.1). Making the substitution, and multiplying both sides by $A_N^{12}(z)A_N^{12}(q^{-1}z)\cdot\Phi^-(q^{-1}z)^{-1}$ we arrive at

$$\begin{aligned} &P_N(\sigma(z))\left[A_N^{22}(q^{-1}z)A_N^{12}(z) + A_N^{12}(q^{-1}z)A_N^{11}(z)\right] = \\ &P_N(\sigma(q^{-1}z))A_N^{12}(z)\Phi^+(q^{-1}z) + P_N(\sigma(qz))A_N^{12}(q^{-1}z)\Phi^-(z). \end{aligned} \tag{10.25}$$

If we alternatively express $c_{N-1}^{-1}P_N(\cdot)$ from the second equation in (10.23), substitute it into the first and perform the same steps we will arrive at

$$\begin{aligned} &P_{N-1}(\sigma(z))\left[A_N^{22}(z)A_N^{21}(q^{-1}z) + A_N^{21}(z)A_N^{11}(q^{-1}z)\right] = \\ &P_{N-1}(\sigma(q^{-1}z))A_N^{21}(z)\Phi^+(q^{-1}z) + P_{N-1}(\sigma(qz))A_N^{21}(q^{-1}z)\Phi^-(z). \end{aligned} \tag{10.26}$$

In the remainder we assume that $N \geq 14$. Then applying Lemma 10.4 to (10.25) we conclude

$$\begin{aligned} &q(z-1)(z-\delta)(z\beta-\gamma)(z\alpha-\gamma\delta)(z^2-\gamma\delta q^2)\cdot A_N^{12}(z)\Phi^+(q^{-1}z) = \\ &(z-\alpha q)(z-\beta\delta q)(z-\gamma q)(z-\gamma\delta q)(z^2-\gamma\delta)\cdot A_N^{12}(q^{-1}z)\Phi^-(z). \end{aligned}$$

Replacing the formulas for $\Phi^\pm$ from (10.14) and canceling common terms we arrive at

$$q^{-3}(z^2-\gamma\delta q^2)\cdot A_N^{12}(z) = (z^2-\gamma\delta)\cdot A_N^{12}(q^{-1}z).$$

The above and Proposition 10.7 imply

$$A_N^{12}(z) = c_{N-1}^{-1} c_N [-q^N \alpha\beta + q^{-N}] \cdot z(z^2 - \gamma\delta). \tag{10.27}$$

If we alternatively apply Lemma 10.4 to (10.26) and repeat the same argument we arrive at

$$A_N^{21}(z) = [q^{-N+1}\alpha\beta - q^{N-1}] \cdot z(z^2 - \gamma\delta). \tag{10.28}$$

10.3. Asymptotics of $A_N(z)$

We now assume that $\alpha, \beta, \gamma, \delta, q$ and M all depend on N and scale as in Definition 7.4.1. For this choice of parameters we denote $\tilde{R}$ in (10.15) by $\tilde{R}_N$ and $\Phi^{\pm}$ by $\Phi_N^{\pm}$.

Our first goal is to show that under the above parameter scaling $A_N(z)$ converges, to a fixed matrix-valued function as $N \to \infty$. Let us first consider the off-diagonal entries $A_N^{12}(z), A_N^{21}(z)$. In view of (10.27) and (10.28) we know that

$$\begin{aligned} A_N^{12}(z) &= c_{N-1}^{-1} c_N [-q_N^N \alpha_N \beta_N + q_N^{-N}] \cdot z(z^2 - \gamma_N \delta_N) \text{ and} \\ A_N^{21}(z) &= [q_N^{-N+1} \alpha_N \beta_N - q_N^{N-1}] \cdot z(z^2 - \gamma_N \delta_N) \end{aligned}$$

Using (10.7) we have

$$\begin{aligned} c_N c_{N-1}^{-1} = \frac{q^{-M}(1 - \alpha_N \beta_N q_N^{2N-1})}{1 - \alpha_N \beta_N q_N^{2N+1}} \\ \cdot \frac{(1 - q_N^N)(1 - \beta_N q_N^N)(\delta_N - \alpha_N q_N^N)(1 - \alpha_N \beta_N q^{M+N+1})}{(1 - \alpha_N \beta_N q_N^N)(1 - \alpha_N q_N^N)(1 - \beta_N \delta_N q_N^N)(1 - q_N^{-M_N+N})}. \end{aligned}$$

Consequently, we see that

$$\lim_{N\to\infty} c_N c_{N-1}^{-1} = \frac{\mathtt{c}(1-\mathtt{q})(1-\mathtt{bq})(\mathtt{d}-\mathtt{aq})(1-\mathtt{abqc}^{-1})}{(1-\mathtt{abq})(1-\mathtt{aq})(1-\mathtt{bdq})(1-\mathtt{cq})} =: \lambda. \tag{10.29}$$

From the above work we conclude that

$$\lim_{N\to\infty} A_N^{12}(z) = \lambda[\mathtt{q}^{-1} - \mathtt{qab}] \cdot z(z^2 - \mathtt{cd}) \text{ and } \lim_{N\to\infty} A_N^{21}(z) = [\mathtt{q}^{-1}\mathtt{ab} - \mathtt{q}] \cdot z(z^2 - \mathtt{cd}).$$

Next, we know that $\det A_N(z) = \Phi_N^+(z)\Phi_N^-(z)$ converges to $(z-\mathtt{a})(z-\mathtt{bd})(z-\mathtt{c})(z-\mathtt{cd})(z-1)(\mathtt{a}z-\mathtt{cd})(\mathtt{b}z-\mathtt{c})(z-\mathtt{d})$ as $N \to \infty$, while from (10.16) and (4.10) we know that $\mathrm{Tr} A_N(z)$ converges to some degree four polynomial $\tilde{R}_\infty(z)$. This implies that $A_N^{11}(z)$ and $A_N^{22}(z)$ converge to some degree four polynomials $A_\infty^{11}(z)$ and $A_\infty^{22}(z)$.

We end this section by proving Fact 9.2.1. From Lemma 10.4 applied to (10.25) we have that

$$B(z)\cdot\left[A_N^{22}(q_N^{-1}z)A_N^{12}(z)+A_N^{12}(q_N^{-1}z)A_N^{11}(z)\right]=A_N^{12}(q_N^{-1}z)\Phi^-(z)\cdot[A(z)+B(z)+C_N(z)]\,.$$

Taking the limit as $N\to\infty$ on both sides we get

$$(z-1)(z-\mathtt{d})(z\mathtt{b}-\mathtt{c})(z\mathtt{a}-\mathtt{cd})(z^2-\mathtt{cd})\lambda[\mathtt{q}^{-1}-\mathtt{qab}]\cdot z(z^2-\mathtt{cd})\cdot\left[A_\infty^{22}(z)+A_\infty^{11}(z)\right]=$$
$$\lambda[\mathtt{q}^{-1}-\mathtt{qab}]z(z^2-\mathtt{cd})\cdot(z-1)(\mathtt{a}z-\mathtt{cd})(\mathtt{b}z-\mathtt{c})(z-\mathtt{d})\times$$
$$[(z-\mathtt{a})(z-\mathtt{bd})(z-\mathtt{c})(z-\mathtt{cd})(z^2-\mathtt{cd})+(z-1)(z-\mathtt{d})(z\mathtt{b}-\mathtt{c})(z\mathtt{a}-\mathtt{cd})(z^2-\mathtt{cd})+$$
$$(z^2-\mathtt{cd})^3(\mathtt{q}^{-1}-1)(1-\mathtt{abq})].$$

Canceling common factors and utilizing (10.16) we see that

$$\lim_{N\to\infty}\tilde{R}_N(z)=\lim_{N\to\infty}\mathrm{Tr}A_N(z)=A_\infty^{22}(z)+A_\infty^{11}(z)=(z-\mathtt{a})(z-\mathtt{bd})(z-\mathtt{c})(z-\mathtt{cd})+$$
$$(z-1)(z-\mathtt{d})(z\mathtt{b}-\mathtt{c})(z\mathtt{a}-\mathtt{cd})+(z^2-\mathtt{cd})^2(\mathtt{q}^{-1}-1)(1-\mathtt{abq}),\tag{10.30}$$

which concludes the proof of Fact 9.2.1.

Remark 10.3.1. We presented the computation in the q-Racah case but the same ideas would work for other families of classical orthogonal polynomials.

References

[1] G. Anderson, A. Guionnet, O. Zeitouni, Introduction to Random Matrices, Cambridge Studies in Advanced Mathematics, 2009.
[2] G. Andrews, R. Askey, R. Roy, Special Functions, Cambridge Univ. Press, New York, 1999.
[3] J. Baik, P. Deift, E. Strahov, Products and ratios of characteristic polynomials of random Hermitian matrices, J. Math. Phys. 44 (2003) 3657–3670.
[4] J. Baik, T. Kriecherbauer, K.D.T.-R. McLaughlin, P.D. Miller, Uniform asymptotics for polynomials orthogonal with respect to a general class of discrete weights and universality results for associated ensembles: announcement of results, Int. Math. Res. Not. 3 (2003) 821–858.
[5] F. Bekerman, A. Figalli, A. Guionnet, Transport maps for β-matrix models and universality, Comm. Math. Phys. 338 (2015) 589–619.
[6] G. Ben Arous, A. Guionnet, Large deviations for Wigner's law and Voiculescu's non-commutative entropy, Probab. Theory Related Fields 108 (1997) 517–542.
[7] F. Bonnet, F. David, B. Eynard, Breakdown of universality in multi-cut models, J. Phys. A: Math. Gen. 33 (2000) 6739–6768.
[8] A. Borodin, Riemann–Hilbert problem and the discrete Bessel kernel, Int. Math. Res. Not. 9 (2000) 467–494.
[9] A. Borodin, Discrete gap probabilities and discrete Painlevé equations, Duke Math. J. 117 (2003) 489–542.
[10] A. Borodin, CLT for spectra of submatrices of Wigner random matrices, Mosc. Math. J. 14 (2014) 29–38.
[11] A. Borodin, CLT for spectra of submatrices of Wigner random matrices II. Stochastic evolution, in: Random Matrix Theory, Interacting Particle Systems, and Integrable Systems, in: Math. Sci. Res. Inst. Publ., vol. 65, Cambridge Univ. Press, New York, 2014, pp. 57–69.
[12] A. Borodin, D. Boyarchenko, Distribution of the first particle in discrete orthogonal polynomial ensembles, Comm. Math. Phys. 234 (2003) 287–338.
[13] A. Borodin, P. Ferrari, Anisotropic growth of random surfaces in $2+1$ dimensions, Comm. Math. Phys. 325 (2014) 603–684.

[14] A. Borodin, V. Gorin, General β-Jacobi corners process and the Gaussian free field, Comm. Pure Appl. Math. 68 (2015) 1774–1844.
[15] A. Borodin, V. Gorin, A. Guionnet, Gaussian asymptotics of discrete β-ensembles, Publ. Math. Inst. Hautes Études Sci. 125 (2017) 1–78.
[16] A. Borodin, V. Gorin, E. Rains, q-distributions on boxed plane partitions, Selecta Math. 16 (2010) 731–789.
[17] G. Borot, A. Guionnet, Asymptotic expansion of β matrix models in the multi-cut regime, Preprint, arXiv:1303.1045, 2013.
[18] G. Borot, A. Guionnet, Asymptotic expansion of β matrix models in the one-cut regime, Comm. Math. Phys. 317 (2013) 447–483.
[19] P. Bourgade, E. László, H.T. Yau, Edge universality of beta ensembles, Comm. Math. Phys. 332 (2014) 261–353.
[20] A. Boutet de Monvel, L. Pastur, M. Shcherbina, On the statistical mechanics approach in the random matrix theory: integrated density of states, J. Stat. Phys. 79 (1995) 585–611.
[21] J. Breuer, M. Duits, Central limit theorems for biorthogonal ensembles and asymptotics of recurrence coefficients, J. Amer. Math. Soc. 30 (2017) 27–66.
[22] A. Bufetov, V. Gorin, Fluctuations of particle systems determined by Schur generating functions, arXiv:1604.01110, 2016.
[23] A. Bufetov, A. Knizel, Asymptotics of random domino tilings of rectangular aztec diamonds, Preprint, arXiv:1604.01491 [math], 2016.
[24] H. Cohn, R. Kenyon, J. Propp, A variational principle for domino tilings, J. Amer. Math. Soc. 14 (2001) 297–346.
[25] H. Cohn, M. Larsen, J. Propp, The shape of a typical boxed plane partition, New York J. Math. 4 (1998) 137–165.
[26] N. Destainville, Entropy and boundary conditions in random rhombus tilings, J. Phys. A 31 (1998) 6123–6139.
[27] N. Destainville, R. Mosseri, F. Bailly, Configurational entropy of codimension-one tilings and directed membranes, J. Stat. Phys. 87 (1997) 697–754.
[28] P.D. Dragnev, E.B. Saff, Constrained energy problems with applications to orthogonal polynomials of a discrete variable, J. Anal. Math. 72 (1997) 223–259.
[29] J. Dubédat, SLE and free field: partition functions and couplings, J. Amer. Math. Soc. 22 (2009) 995–1054.
[30] B. Eynard, All genus correlation functions for the hermitian 1-matrix model, J. High Energy Phys. 04111 (2004) 031.
[31] K. Féral, On large deviations for the spectral measure of discrete Coulomb gas, Sémin. Probab. XLI Lecture Notes in Math. 1934 (2008) 19–49.
[32] P.J. Forrester, Log-Gases and Random Matrices, Princeton University Press, Princeton, 2010.
[33] A. Guionnet, J. Huang, Rigidity and edge universality of discrete β-ensembles, Preprint, arXiv: 1705.05527, 2017.
[34] E. Hille, J. Tamarkin, On the absolute integrability of Fourier transforms, Fund. Math. 25 (1935) 329–352.
[35] L. Hörmander, The Analysis of Linear Partial Differential Operators I, Springer-Verlag, Berlin, 2003.
[36] X. Hu, J. Miller, Y. Peres, Thick points of the Gaussian free field, Ann. Probab. 38 (2010) 896–926.
[37] W. Jockusch, J. Propp, P. Shor, Random domino tilings and the arctic circle theorem, Preprint, arXiv:math/9801068, 1998.
[38] K. Johansson, On fluctuations of eigenvalues of random Hermitian matrices, Duke Math. J. 91 (1998) 151–204.
[39] K. Johansson, Shape fluctuations and random matrices, Comm. Math. Phys. 209 (2000) 437–476.
[40] K. Johansson, Non-intersecting paths, random tilings and random matrices, Probab. Theory Related Fields 123 (2002) 225–280.
[41] R. Kenyon, Dominos and the Gaussian free field, Ann. Probab. 29 (2001) 1128–1137.
[42] R. Kenyon, Height fluctuations in the honeycomb dimer model, Comm. Math. Phys. 281 (2008) 675–709.
[43] R. Kenyon, Lectures on dimers, in: Statistical Mechanics, in: IAS/ Park City Math. Ser., vol. 16, Amer. Math. Soc., Providence, RI, 2009, pp. 191–230.
[44] R. Kenyon, A. Okounkov, Limit shapes and the complex Burgers equation, Acta Math. 199 (2007) 263–302.
[45] R. Kenyon, A. Okounkov, S. Sheffield, Dimers and Amoebae, Ann. Math. 3 (2006) 1019–1056.

[46] R. Koekoek, P.A. Lesky, R.F. Swarttouw, Hypergeometric Orthogonal Polynomials and Their q-analogues, Springer Monographs in Mathematics, Springer-Verlag, New York, 2010.
[47] R. Koekoek, R.F. Swarttouw, The Askey-scheme of hypergeometric orthogonal polynomials and its q-analogue, Preprint, https://arxiv.org/abs/math/9602214, 1996.
[48] R. Koekoek, R.F. Swarttouw, The Askey-Scheme of Hypergeometric Orthogonal Polynomials and Its q-analogue, Report no. 98–17, Delft University of Technology, 1998, http://homepage.tudelft.nl/11r49/askey/.
[49] T.H. Koorwinder, Askey–Wilson polynomials for root systems of type BC, Contemp. Math. 47 (1992) 189–204.
[50] T. Kriecherbauer, M. Shcherbina, Fluctuations of eigenvalues of matrix models and their applications, Preprint, arXiv:1003.6121, 2010.
[51] A.B.J. Kuijlaars, On the finite-gap ansatz in the continuum limit of the Toda lattice, Duke Math. J. 104 (2000) 433–462.
[52] M. Maïda, É. Maurel-Segala, Free transport-entropy inequalities for non-convex potentials and application to concentration for random matrices, Probab. Theory Related Fields 159 (2014) 329–356.
[53] M.L. Mehta, Random Matrices, 3rd edition, Elsevier/Academic Press, Amsterdam, 2004.
[54] A.A. Migdal, Loop equations and $1/n$ expansion, Phys. Rep. 102 (1983) 199–290.
[55] K. Mimachi, A duality of Macdonald–Koorwinder polynomials and its application to integral representations, Duke Math. J. 2 (2001) 265–281.
[56] N.I. Muskhelishvili, Singular Integral Equations, Wolters-Noordhoff Publishing, Groningen, the Netherlands, 1958.
[57] N. Nekrasov, BPS/CFT correspondence: non-perturbative Dyson–Schwinger equations and qq-characters, J. High Energy Phys. 3 (2016) 181.
[58] A.F. Nikiforov, S.K. Suslov, V.B. Uvarov, Classical Orthogonal Polynomials of a Discrete Variable, Springer Series in Computational Physics, Springer Series in Computational Physics, Springer-Verlag, New York, 1991.
[59] A. Okounkov, N. Reshetikhin, Correlation functions of Schur process with application to local geometry of a random 3-dimensional Young diagram, J. Amer. Math. Soc. 16 (2003) 581–603.
[60] L. Pastur, M. Shcherbina, Eigenvalue Distribution of Large Random Matrices, AMS, Providence, 2011.
[61] L. Petrov, Asymptotics of random lozenge tilings via Gelfand–Tsetlin schemes, Probab. Theory Related Fields 160 (2014) 429–487.
[62] L. Petrov, Asymptotics of uniformly random lozenge tilings of polygons. Gaussian free field, Ann. Probab. 43 (2015) 1–43.
[63] E.M. Rains, BC_n-symmetric polynomials, Transform. Groups 10 (2005) 63–132.
[64] M. Shcherbina, Fluctuations of linear eigenvalue statistics of β matrix models in the multi-cut regime, J. Stat. Phys. 151 (2013) 1004–1034.
[65] S. Sheffield, Random surfaces, Astérisque, vol. 304, 2006.
[66] S. Sheffield, Gaussian free fields for mathematicians, Probab. Theory Related Fields 139 (2007) 521–541.
[67] R.P. Stanley, Enumerative Combinatorics, vol. 2, Cambridge University Press, 1999.
[68] E. Stein, R. Shakarchi, Complex Analysis, Princeton University Press, Princeton, 2003.
[69] E.C. Titchmarsh, Introduction to the Theory of Fourier Integrals, Oxford University Press, London, 1937.
[70] E.P. Wigner, On the distribution of the roots of certain symmetric matrices, Ann. Math. 67 (1958) 325–327.