Uniform Asymptotics for Polynomials Orthogonal with Respect to a General Class of Discrete Weights and Universality Results for Associated Ensembles: Announcement of Results

J. Baik, **T. Kriecherbauer**,
K. D. T.-R. McLaughlin, **and P. D. Miller**

1 Introduction

This announcement concerns asymptotic properties of polynomials that are orthogonal with respect to pure point measures supported on finite sets. Let $N \in \mathbb{N}$ be fixed, and consider N distinct real nodes $x_{N,0} < x_{N,1} < \cdots < x_{N,N-1}$ to be given; together the nodes make up the support of the pure point measures we consider. We use the notation $X_N := \{x_{N,n}\}_{n=0}^{N-1}$ for the support set. Along with nodes we are given positive weights $w_{N,0}, w_{N,1}, \ldots, w_{N,N-1}$, which are the magnitudes of the point masses located at the corresponding nodes. The *discrete orthogonal polynomials* associated with this data are polynomials $\{p_{N,k}(z)\}_{k=0}^{N-1}$ where $p_{N,k}(z)$ is of degree exactly k with a positive leading coefficient and where

$$\sum_{j=0}^{N-1} p_{N,k}(x_{N,j}) p_{N,l}(x_{N,j}) w_{N,j} = \delta_{kl}. \tag{1.1}$$

If $p_{N,k}(z) = c_{N,k}^{(k)} z^k + \cdots + c_{N,k}^{(0)}$, then we denote by $\pi_{N,k}(z)$ the associated monic polynomial $p_{N,k}(z)/c_{N,k}^{(k)}$. These polynomials exist and are uniquely determined by the orthogonality conditions because the inner product associated with (1.1) is positive definite on

$\text{span}(1, z, z^2, \ldots, z^{N-1})$ but is degenerate on larger spaces of polynomials. The polynomials $p_{N,k}(z)$ may be constructed from the monomials by a Gram-Schmidt process. A general reference for properties of orthogonal polynomials specific to the discrete case is the book of Nikiforov, Suslov, and Uvarov [25]. In contrast to the discrete orthogonal polynomials, we refer to the polynomials orthogonal with respect to an absolutely continuous measure as the *continuous orthogonal polynomials*.

We use the notation $\mathbb{Z}_N$ for the set $\{0, 1, 2, \ldots, N-1\}$. Examples of classical discrete weights are (see [1])

(i) Krawtchouk weight: on the nodes $x_{N,j} := (j + 1/2)/N$ for $j \in \mathbb{Z}_N$ in the interval $(0, 1)$, we define the weight

$$w_{N,j}^{\text{Kraw}}(p, q) := \frac{N^{N-1}\sqrt{pq}}{q^N \Gamma(N)} \binom{N-1}{j} p^j q^{N-1-j}. \tag{1.2}$$

(ii) Hahn weight: on the infinite set of nodes $x_{N,j} := (j + 1/2)/N$ for $j = 0, 1, 2, \ldots,$ we define the weight

$$w_{N,j}(b, c, d) := \frac{N^{N-1}}{\Gamma(N)} \cdot \frac{\Gamma(b)\Gamma(c+j)\Gamma(d+j)}{\Gamma(j+1)\Gamma(b+j)\Gamma(c)\Gamma(d)}, \tag{1.3}$$

where b, c, and d are real parameters. Special cases are the following. First, taking $d = \alpha$ and $b = 2 - N - \beta$, with $\alpha, \beta > 0$, and taking the limit $c \to 1 - N$, we obtain the weight

$$w_{N,j}^{\text{Hahn}}(\alpha, \beta) := \frac{N^{N-1}}{\Gamma(N)} \cdot \frac{\binom{j+\alpha-1}{j}\binom{N+\beta-2-j}{N-1-j}}{\binom{N+\beta-2}{\beta-1}}, \quad \text{for } j \in \mathbb{Z}_N, \tag{1.4}$$

which is defined on the *finite* set of nodes $X_N := (\mathbb{Z}_N + 1/2)/N$. Second, taking $d = 2 - N - \beta$ and $b = \alpha$, with $\alpha, \beta > 0$, and taking the limit $c \to 1 - N$, we obtain the weight

$$w_{N,j}^{\text{Assoc}}(\alpha, \beta) := \frac{N^{N-1}}{\Gamma(N)} \cdot \frac{\Gamma(N)\Gamma(N+\beta-1)\Gamma(\alpha)}{\Gamma(j+1)\Gamma(\alpha+j)\Gamma(N-j)\Gamma(N+\beta-1-j)}, \quad \text{for } j \in \mathbb{Z}_N, \tag{1.5}$$

which is again defined on the nodes $X_N := (\mathbb{Z}_N + 1/2)/N$.

Remark 1.1. In [1], the polynomials orthogonal with respect to the weight (1.3) are called the Hahn polynomials, and in [5], the same weight (on the whole integers) is called the Askey-Lesky weight. For the two special cases where the weight is supported on a finite set of nodes, we now adopt the terminology used by Johansson [18], and thus refer to the

weight (1.4) simply as the *Hahn weight* (corresponding to the Hahn polynomials) and we refer to the weight (1.5) as the *associated Hahn weight* (corresponding to the associated Hahn polynomials).

Our goal is to establish the asymptotic behavior of the polynomials $p_{N,k}(z)$ or their monic counterparts $\pi_{N,k}(z)$ in the limit of large degree, assuming certain asymptotic properties of the nodes and the weights. In particular, the number of nodes must necessarily increase to admit polynomials with arbitrarily large degree, and the weights we consider involve an exponential factor with exponent proportional to the number of nodes (such weights are sometimes called *varying weights*). We will obtain pointwise asymptotics with precise error bound uniformly valid in the whole complex plane with the exception of certain arbitrarily small open discs. Our assumptions on the nodes and weights include as special cases all relevant classical discrete orthogonal polynomials, but are significantly more general; in particular, we will consider nodes that are not necessarily equally spaced.

1.1   Motivation and applications

In the context of approximation theory, there has been recent activity [10, 11] in the study of polynomials orthogonal on the real axis with respect to general continuous varying weights and the corresponding large degree pointwise asymptotics. One important application of the results of [10, 11] is the proof of several universality conjectures of random matrix theory. Thus, a natural question to ask is whether it is possible to extend the results of [10, 11] to handle discrete weights, and obtain similar universality results for the so-called *discrete orthogonal polynomial ensembles* (see Section 3.1).

Indeed, it has turned out recently that various problems of percolation models, random tiling, queueing theory, nonintersecting paths, and representation theory can be reformulated as asymptotic questions of discrete orthogonal polynomial ensembles with very concrete weights (see, e.g., [5, 16, 17]). In these ensembles, the weights are all classical (Meixner, Charlier, Krawtchouk, or Hahn). Using integral formulae for the corresponding orthogonal polynomials, the relevant asymptotics have been analyzed except for the Hahn weight case. The weights handled in this paper include the Hahn weight (and also Krawtchouk), and hence as a corollary, we obtain new asymptotic results for the Hahn polynomials. As the discrete orthogonal polynomial ensemble for the Hahn weight arises in the statistical analysis of random rhombus tilings of a hexagon, our asymptotic results for Hahn polynomials yield new results on this problem (see Section 3.2).

1.2   Methodology

The method we use is the Riemann-Hilbert characterization of discrete orthogonal poly-nomials, and an adaption of the Deift-Zhou method for the steepest-descent analysis of Riemann-Hilbert problems. The Riemann-Hilbert problem for discrete orthogonal poly-nomials has poles instead of usual jump conditions on a continuous contour, and the poles are accumulating in the limit of interest to form a continuum.

There has been some recent progress [19, 24] in the integrable systems literature concerning the problem of computing asymptotics for solutions of integrable nonlin-ear partial differential equations (e.g., the nonlinear Schrödinger equation) in the limit where the spectral data associated with the solution via the inverse-scattering transform is made up of a large number of discrete eigenvalues. Significantly, inverse-scattering theory also exploits much of the theory of matrix Riemann-Hilbert problems, and it turns out that the discrete eigenvalues appear as poles in the corresponding matrix-valued un-known. So, the methods recently developed in the context of inverse-scattering actually suggest a general scheme by means of which an accumulation of poles in the matrix un-known can be analyzed.

In this paper, we extend the method of [19, 24] and suggest a general and unifying approach to handle Riemann-Hilbert problems for the situation when poles are accumu-lating. Especially, we overcome the following two issues.

(a) How to transform a Riemann-Hilbert problem with pole conditions to a Riemann-Hilbert problem with an analytic jump condition on a continuous contour so that a formal continuum limit of poles can be rigorously justified and the Deift-Zhou method can be applied.

(b) How to handle the upper constraint of the so-called equilibrium measure, and thus to correctly formulate an appropriate g-function.

See Section 4 for more information about these ideas. Full details will be given in the paper corresponding to this announcement.

1.3   Basic assumptions

We state here precise assumptions on the nodes X_N and weights $\{w_{N,j}\}$.

1.3.1   Conditions on the nodes.   (1) The nodes lie in a bounded open interval (a, b) and are distributed with a density $\rho^0(x)$.

(2) The density function $\rho^0(x)$ is real analytic in a complex neighborhood of the closed interval $[a, b]$, and satisfies

$$\int_a^b \rho^0(x)dx = 1, \tag{1.6}$$

$$\rho^0(x) > 0 \quad \text{strictly, for all } x \in [a,b]. \tag{1.7}$$

(3) The nodes are defined precisely in terms of the density function $\rho^0(x)$ by the quantization rule

$$\int_a^{x_{N,j}} \rho^0(x)dx = \frac{2j+1}{2N} \tag{1.8}$$

for $N \in \mathbb{N}$ and $j \in \mathbb{Z}_N$.

1.3.2   Conditions on the weights. (1) Without loss of generality, we write the weights in the form

$$
\begin{aligned}
w_{N,j} &= (-1)^{N-1-j} e^{-NV_N(x_{N,j})} \prod_{\substack{n=0 \\ n \neq j}}^{N-1} \left(x_{N,j} - x_{N,n}\right)^{-1} \\
&= e^{-NV_N(x_{N,j})} \prod_{\substack{n=0 \\ n \neq j}}^{N-1} \left|x_{N,j} - x_{N,n}\right|^{-1},
\end{aligned}
\tag{1.9}
$$

where the family of functions $\{V_N(x)\}$ is a priori specified only at the nodes.

(2) We assume that for each sufficiently large N, $V_N(x)$ may be taken to be a real analytic function defined in a neighborhood G of the closed interval $[a,b]$, and that

$$V_N(x) = V(x) + \frac{\gamma}{N} + \frac{\eta_N(x)}{N^2}, \tag{1.10}$$

where $V(x)$ is a fixed real analytic function defined in G, γ is a constant, and

$$\limsup_{N \to \infty} \sup_{z \in G} \left|\eta_N(z)\right| < \infty. \tag{1.11}$$

Remark 1.2. In some applications it is desirable to generalize further by allowing γ in (1.10) to be a real analytic function in G with $\gamma'(z)$ not identically zero. It is possible to take into account such variation, but for simplicity we take γ to be constant in this paper. For classical cases of Hahn and Krawtchouk weights, γ is indeed constant in scalings of interest.

The familiar examples of classical discrete orthogonal polynomials correspond to nodes that are equally spaced, say on $(a,b) = (0,1)$ (in which case we have $\rho^0(x) \equiv 1$).

In this special case, the product factor on the right-hand side of (1.9) becomes simply

$$\prod_{\substack{n=0 \\ n \neq j}}^{N-1} |x_{N,j} - x_{N,n}|^{-1} = \frac{N^{N-1}}{j!(N-j-1)!}. \tag{1.12}$$

By Stirling's formula, taking the continuum limit of this factor (i.e., considering $N \to \infty$ with $j/N \to x$) shows that in these cases formula (1.9) leads to a continuous weight on $(0, 1)$ of the form

$$w(x) = \left(\frac{e^{-V(x)}}{x^x(1-x)^{1-x}} \right)^N \tag{1.13}$$

up to an overall multiplicative constant.

Our choice of the form (1.9) for the weights is motivated by several specific examples of classical discrete orthogonal polynomials. The form (1.9) is sufficiently general for us to carry out useful calculations related to proofs of universality conjectures arising in statistical problems like the random rhombus tiling of a hexagon.

1.3.3 Conditions on the equilibrium measure. There is an additional assumption, which is not as explicit as the previous two conditions. This assumption will be explained in the next subsection.

1.4   The equilibrium energy problem and third assumption on the weights

It has been recognized for some time (see [21] and the references therein) that, as in the continuous orthogonal polynomial cases, the asymptotic behavior of discrete orthogonal polynomials, in particular the distribution of zeros in (a, b), is related to a constrained equilibrium problem for logarithmic potentials in a field $\varphi(x)$ given by the formula

$$\varphi(x) := V(x) + \int_a^b \log|x - y| \rho^0(y) dy \tag{1.14}$$

for $x \in (a, b)$. We can also view $\varphi(x)$ as being defined via a continuum limit

$$\varphi(x) = -\lim_{N \to \infty} \frac{\log(w_{N,j})}{N}, \tag{1.15}$$

where $w_{N,j}$ is expressed in terms of $x_{N,j}$ which in turn is identified with x. Thus at the moment we are working with the *formal continuum limit* of the weight $w_{N,j}$.

In the specific context of this paper, the field $\varphi(x)$ is a real analytic function in the open interval (a, b) because $V(x)$ and $\rho^0(x)$ are real analytic functions in a neighborhood

of $[a, b]$. Unlike $V(x)$ and $\rho^0(x)$, however, the field $\varphi(x)$ does not extend analytically beyond the endpoints of (a, b) due to condition (1.7).

Given the parameter $c \in (0, 1)$, which has the interpretation of the ratio of the degree k of the polynomial of interest to the number N of nodes, and the field $\varphi(x)$ as above, consider the quadratic functional

$$E_c[\mu] := c \int_a^b \int_a^b \log \frac{1}{|x - y|} d\mu(x) d\mu(y) + \int_a^b \varphi(x) d\mu(x) \tag{1.16}$$

of Borel measures μ on $[a, b]$. Let $\mu_{\min}^c$ be the measure that minimizes $E_c[\mu]$ over the class of measures satisfying the upper and lower constraints

$$0 \leq \int_{x \in \mathcal{B}} d\mu(x) \leq \frac{1}{c} \int_{x \in \mathcal{B}} \rho^0(x) dx \tag{1.17}$$

for all Borel sets $\mathcal{B} \subset [a, b]$, and the normalization condition

$$\int_a^b d\mu(x) = 1. \tag{1.18}$$

The existence of a unique minimizer under the conditions enumerated in Sections 1.3.1 and 1.3.2 follows from the Gauss-Frostman Theorem; see [26] for details. We will often refer to the minimizer as the *equilibrium measure*. It has been shown [21] that the equilibrium measure is the weak limit of the normalized counting measure of the zeros of $p_{N,k}(z)$ in the limit $N \to \infty$ with $c = k/N$ fixed.

That a variational problem plays a central role in asymptotic behavior is a familiar theme in the theory of orthogonal polynomials. The key new feature contributed by discreteness is the appearance of the upper constraint on the equilibrium measure (i.e., the upper bound in (1.17)). The upper constraint can be traced to the following well-known fact.

Proposition 1.3. Each discrete orthogonal polynomial $p_{N,k}(z)$ has k simple real zeros. All zeros lie in the range $x_{N,0} < z < x_{N,N-1}$ and no more than one zero lies in the closed interval $[x_{N,n}, x_{N,n+1}]$ between any two consecutive nodes.   □

Thus, the presence of the upper constraint proportional to the local density of nodes is necessary for the interpretation of the equilibrium measure as the weak limit of the normalized counting measure of zeros.

The theory of the "doubly constrained" variational problem we are considering is well established. In particular, the analytic properties we assume of $V(x)$ and $\rho^0(x)$ turn out to be unnecessary for the mere existence of the minimizer. However, it has been shown in [20] that analyticity of $V(x)$ and $\rho^0(x)$ in a neighborhood of $[a, b]$ guarantees that $\mu_{\min}^c$ is continuously differentiable with respect to $x \in (a, b)$. Moreover, the derivative $d\mu_{\min}^c/dx$ is piecewise analytic, with a finite number of points of nonanalyticity that may not occur at any x where both (strict) inequalities $d\mu_{\min}^c/dx(x) > 0$ and $d\mu_{\min}^c/dx(x) < \rho^0(x)/c$ hold. We want to exploit these facts, which is why we have chosen to restrict attention to analytic functions $V(x)$ and $\rho^0(x)$.

For a method of computing the equilibrium measure from the coefficients in the three-term recurrence relation for a special class of discrete weights, see [21]. See also [9] for continuous weights.

For simplicity of exposition we want to exclude certain nongeneric phenomena that may occur even under conditions of analyticity of $V(x)$ and $\rho^0(x)$. Therefore, we introduce the following assumptions.

1.4.1   Third assumption; *conditions on the equilibrium measure.* Let $\underline{\mathcal{F}} \subset [a, b]$ denote the closed set of x-values where $d\mu_{\min}^c/dx(x) = 0$. Let $\overline{\mathcal{F}} \subset [a, b]$ denote the closed set of x-values where $d\mu_{\min}^c/dx(x) = \rho^0(x)/c$.

(1) Each connected component of $\underline{\mathcal{F}}$ and $\overline{\mathcal{F}}$ has a nonempty interior. Therefore, $\underline{\mathcal{F}}$ and $\overline{\mathcal{F}}$ are both finite unions of closed intervals, where each closed interval that is part of the union contains more than one point. Note that this does not exclude the possibility that either $\underline{\mathcal{F}}$ or $\overline{\mathcal{F}}$ might be empty.

(2) For each open subinterval I of $(a, b) \setminus \underline{\mathcal{F}} \cup \overline{\mathcal{F}}$ and each limit point $z_0 \in \underline{\mathcal{F}}$ of I, we have

$$\lim_{x \to z_0, x \in I} \frac{1}{\sqrt{|x - z_0|}} \frac{d\mu_{\min}^c}{dx}(x) = K \quad \text{with } 0 < K < \infty, \tag{1.19}$$

and for each limit point $z_0 \in \overline{\mathcal{F}}$ of I, we have

$$\lim_{x \to z_0, x \in I} \frac{1}{\sqrt{|x - z_0|}} \left[\frac{1}{c}\rho^0(x) - \frac{d\mu_{\min}^c}{dx}(x) \right] = K \quad \text{with } 0 < K < \infty. \tag{1.20}$$

Therefore, the derivative of the minimizing measure meets each constraint exactly like a square root.

(3) A constraint is active at each endpoint: $\{a, b\} \subset \underline{\mathcal{F}} \cup \overline{\mathcal{F}}$.

It is difficult to translate these conditions on $\mu_{\min}^c$ into sufficient conditions on c, $V(x)$, and $\rho^0(x)$. However, there is a sense in which they are satisfied generically.

Remark 1.4. Relaxing the condition that a constraint should be active at each endpoint requires specific local analysis near these two points. We expect that a constraint being active at each endpoint is a generic phenomenon in the sense that the opposite situation occurs only for isolated values of c. We know this statement to be true in all relevant classical cases. For the Krawtchouk polynomials only the values $c = p$ or $c = q = 1 - p$ correspond to an equilibrium measure that is not constrained at both endpoints (see [13]). The situation is similar for the Hahn polynomials.

1.4.2   Voids, bands, and saturated regions.   Under the conditions enumerated in Sections 1.3.1, 1.3.2, and 1.4.1, the minimizer μ_{min}^c partitions (a, b) into three kinds of subintervals, a finite number of each, and each having a nonempty interior. There is a real constant ℓ_c, the Lagrange multiplier associated with the condition (1.18), so that with the variational derivative defined as

$$\frac{\delta E_c}{\delta \mu}(x) := -2c \int_a^b \log|x - y| d\mu(y) + \varphi(x), \tag{1.21}$$

we have, when $\mu = \mu_{min}^c$, the following types of subintervals.

Definition 1.5 (voids). A void Γ is an open subinterval of $[a, b]$ of maximal length in which $\mu_{min}^c(x) \equiv 0$, and thus the minimizer realizes the lower constraint. For $x \in \Gamma$, we have the strict inequality

$$\frac{\delta E_c}{\delta \mu}(x) > \ell_c. \tag{1.22}$$

Definition 1.6 (bands). A band I is an open subinterval of $[a, b]$ of maximal length where $\mu_{min}^c(x)$ is a measure with a real analytic density satisfying $0 < d\mu_{min}^c/dx < \rho^0(x)/c$, and thus variations of the minimizer are free. For $x \in I$, the equilibrium condition is

$$\frac{\delta E_c}{\delta \mu}(x) \equiv \ell_c. \tag{1.23}$$

Definition 1.7 (saturated regions). A saturated region Γ is an open subinterval of $[a, b]$ of maximal length in which $d\mu_{min}^c/dx \equiv \rho^0(x)/c$, and thus the minimizer realizes the upper constraint. For $x \in \Gamma$, the strict inequality is

$$\frac{\delta E_c}{\delta \mu}(x) < \ell_c. \tag{1.24}$$

See Figure 1.1 for an example of the voids, bands, and saturated regions associated with a hypothetical equilibrium measure.
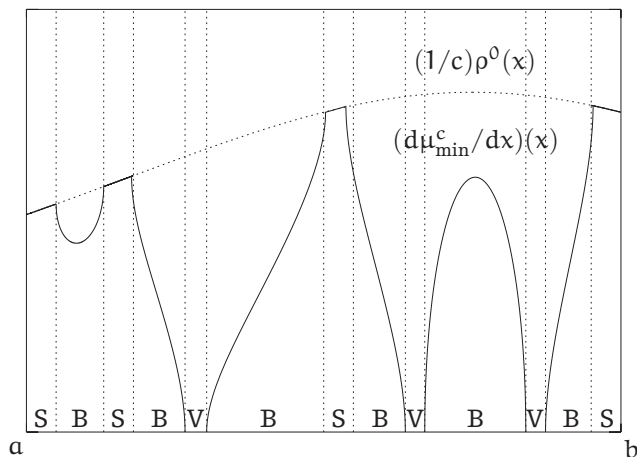
Figure 1.1   The hypothetical minimizer illustrated here partitions the interval (a, b) into voids (denoted V), bands (denoted B), and saturated regions (denoted S).

Voids and saturated regions will also be called *gaps* when it is not necessary to distinguish between these two types of intervals. The closure of the union of all subintervals of the three types defined above is the interval $[a, b]$. From condition (1) in Section 1.4.1, bands cannot be adjacent to each other; a band that is not adjacent to an endpoint of $[a, b]$ has on each side either a void or a saturated region.

Some of our asymptotic results for the discrete orthogonal polynomials under the above assumptions are stated in Section 2. In Section 3, the asymptotics of discrete orthogonal polynomials are applied to discrete orthogonal polynomial ensembles, and asymptotics of corresponding correlation functions are thus obtained. By specializing to the Hahn ensemble, we arrive at specific results relevant in the problem of random rhombus tiling of a hexagon. Our new methods for the asymptotic analysis of general discrete orthogonal polynomials are discussed in Section 4.

2   Results: pointwise asymptotics of orthogonal polynomials

As stated in the introduction, we obtain pointwise asymptotics of the orthogonal polynomials $p_{N,k}(z)$ for z in the complex plane except for a finite union of discs of arbitrarily small but fixed radii in terms of a log transform of the equilibrium measure. These discs are centered at the edges of the bands. The difficulty at the edge of the bands will be explained below (see Section 4). We hope to be able to handle the band edge problem in our future publication.

We actually present here formulae for the monic polynomials $\pi_{N,k}(z)$, rather than the normalized polynomials $p_{N,k}(z)$. The asymptotics are given by different formulae in five different regions of the complex plane excluding small discs around the edges of the bands. These regions are (recall that $[a, b]$ is the interval where the nodes are accumulating)

   (a) outside the interval $[a, b]$;
   (b) voids in $[a, b]$;
   (c) bands in $[a, b]$;
   (d) saturated regions in $[a, b]$;
   (e) near the endpoints of $[a, b]$ adjacent to a saturated region.

The results for the first three cases are analogous to the corresponding results for the continuous orthogonal polynomials analyzed in [11]. Also the asymptotic formula for $\pi_{N,k}$ near the endpoints of $[a, b]$ adjacent to a void is analogous to the corresponding asymptotics of continuous orthogonal polynomials whose weight is supported on a finite interval. The new cases that did not occur in the continuous orthogonal polynomial theory are the regions (d) and (e). In these regions, the discrete nature of the support of the weights is strongly present. Here we present only the three regions (c), (d), and (e). Asymptotic formulae for z in regions (a) and (b) will appear in the full version of this paper, together with the asymptotics of the leading coefficient $c_{N,k}^{(k)}$ which completes the connection with the polynomials $p_{N,k}(z)$.

For simplicity, we take the limit $k, N \to \infty$ while $c = k/N$ is a fixed rational number. (Recall that k is the degree of the orthogonal polynomial, and N is the number of nodes.) This means that if the fixed rational constant c is represented in lowest terms as $c = p/q$, then we are taking $k = Mp$ and $N = Mq$ for $M \in \mathbb{N}$. The more general case of $k/N = c + O(N^{-1})$ will be considered in a future publication.

In Theorems 2.1 and 2.3 below, the error bounds are different depending on the following two situations of the support of the equilibrium measure.

Case I. There is at least one void interval and at least one saturated region.

Case II. There are only voids and bands, or only saturated regions and bands.

We have obtained a better error bound for Case II than for Case I. Whether this is only a technical point, or whether this is in the very nature of discrete orthogonal polynomials is not clear yet.

2.1   Bands

Theorem 2.1 (asymptotics of $\pi_{N,k}(z)$ in bands)**.** Assume the conditions enumerated in Sections 1.3.1, 1.3.2, and 1.4.1. Let $c = k/N$ be fixed. Then, uniformly for z in any fixed

compact subinterval in the interior of a band I,

$$\pi_{N,k}(z) = \exp\left(k\int_a^b \log|z-x|d\mu_{\min}^c(x)\right)$$
$$\cdot\left[2A_I(z)\cos\left(k\pi\int_z^b d\mu_{\min}^c(x) + \Phi_I(z)\right) + \varepsilon_N(z)\right],$$

(2.1)

where

$$\varepsilon_N(z) = \begin{cases} O\left(\dfrac{\log(N)}{N^{1/3}}\right), & \text{Case I,} \\[3mm] O\left(\dfrac{\log(N)}{N^{2/3}}\right), & \text{Case II,} \end{cases}$$

(2.2)

and $A_I(z) > 0$ and $\Phi_I(z)$ are real functions defined in terms of a Riemann theta function associated to the hyperelliptic surface with cuts given by the bands. The functions A_I and Φ_I are uniformly bounded along with all derivatives. □

Remark 2.2. The error estimates quoted above are derived from the asymptotic procedure we have devised for the case of node distributions associated with a general analytic density function $\rho^0(x)$. In the special case of equally spaced nodes (when $\rho^0(x)$ is a constant function), we have recently found, using a different procedure, that the logarithmic term $\log(N)$ in the error bound $\varepsilon_N(z)$ can apparently be removed. At this time we do not know whether this alternate procedure can be modified for nonconstant $\rho^0(x)$ so as to remove the logarithm from the estimates in all cases.

2.2   Saturated regions

Theorem 2.3 (asymptotics of $\pi_{N,k}(z)$ in saturated regions). Assume the conditions enumerated in Sections 1.3.1, 1.3.2, and 1.4.1. Let $c = k/N$ be fixed. Then, uniformly for z in a saturated region Γ, but bounded away from any band edge points where the equilibrium measure becomes unconstrained by a fixed distance, and from the endpoints a and b by a distance of size $N^{-2/3}$, we have

$$\pi_{N,k}(z) = \exp\left(k\int_a^b \log|z-x|d\mu_{\min}^c(x)\right)$$
$$\cdot\left[(\phi_\Gamma(z) + \varepsilon_N(z))\cdot 2\cos\left(\pi N\int_z^b \rho^0(x)dx\right) + \text{ exponentially small}\right],$$

(2.3)

with $\varepsilon_N(z)$ having the same error bound as in (2.2), where $\phi_\Gamma(z)$ is a real function defined in terms of a Riemann theta function, uniformly bounded along with all derivatives

and having at most one zero in Γ. If the saturated region Γ is adjacent to an endpoint of $[a, b]$, then $\phi_\Gamma(z)$ has no zeros in Γ. The exponentially small term is proportional to $\exp(N[\delta E_c/d\mu - \ell_c])$ where the variational derivative is evaluated on the equilibrium measure. $\qquad\square$

Since the zeros of the cosine function in (2.3) are exactly the nodes of orthogonalization making up the set X_N, and since the slope of the cosine is proportional to N at the nodes, we have the following.

Corollary 2.4 (exponential confinement of zeros)**.** Let J be a closed subinterval of a saturated region where the equilibrium measure achieves the upper constraint. Then the monic discrete orthogonal polynomial $\pi_{N,k}(z)$ has a zero uniformly close to each of the nodes $x_{N,n} \in X_N \cap J$, with the possible exception of one node. $\qquad\square$

Remark 2.5. The factor $\phi_\Gamma(z)+\varepsilon_N(z)$ has at most one zero in Γ, and the zero can be present for some N and not others. Also, the zero generally moves about in a quasiperiodic manner as N is varied. So it seems that one should regard the situation in which this zero is exponentially close to one of the nodes (which form a set of measure zero in J) as being anomalous and quite rare. Therefore, one should generally expect to see a zero exponentially close to *each* node in $X_N \cap J$.

Let K be a subinterval of the interval J that is the subject of Corollary 2.4 such that there is a zero of $\pi_{N,k}(z)$ exponentially close to each node in $X_N \cap K$ (so according to Remark 2.5, one expects that it is typically consistent to take $K = J$). The exponential confinement of the zeros in K has further consequences due to the rigidity of the zeros for general discrete orthogonal polynomials described in Proposition 1.3. A particular zero $z_0 \in K$ of $\pi_{N,k}(z)$, asymptotically exponentially localized near a node $x_{N,n}$, can lie on one side or the other of the node. But if z_0 lies to the right of $x_{N,n}$, then it follows from Proposition 1.3 that the smallest zero greater than z_0 must also lie to the right of $x_{N,n+1}$ and so on, all the way to the right endpoint of K. Likewise, if z_0 lies to the left of $x_{N,n}$, then all zeros in K less than z_0 also lie to the left of the nodes to which they are exponentially attracted.

When we consider those zeros of $\pi_{N,k}(z)$ that converge exponentially fast to the nodes (these are analogous to the *Hurwitz zeros* of the approximation theory literature, whereas the possible lone zero of $\phi_\Gamma(z)$ would be called a *spurious zero*), we therefore see that there can be at most one "dislocation" (i.e., a closed interval of the form $[x_{N,n}, x_{N,n+1}]$ containing no Hurwitz zeros) in the pattern of zeros lying to one side or the other of the nodes. See Figure 2.1.

···ㅇ···ㅇ···ㅇ···ㅇ···ㅇ···ㅇ···ㅇ···ㅇ···ㅇ···ㅇ···ㅇ···ㅇ···ㅇ···ㅇ···ㅇ··

··ㅇ···ㅇ···ㅇ···ㅇ···ㅇ···ㅇ···ㅇ···ㅇ···ㅇ···ㅇ···ㅇ···ㅇ···ㅇ···ㅇ···ㅇ··

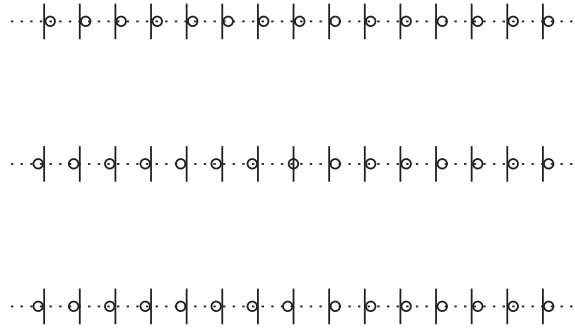··ㅇ···ㅇ···ㅇ···ㅇ···ㅇ···ㅇ···ㅇ···ㅇ···ㅇ···ㅇ···ㅇ···ㅇ···ㅇ···ㅇ···ㅇ··

Figure 2.1   The Hurwitz zeros are exponentially close to the nodes of orthogonalization in saturated regions where the equilibrium measure achieves its upper constraint. Top: a pattern without any dislocation, where all Hurwitz zeros (pictured as circles) lie to the right of the nodes (vertical line segments) to which they are exponentially attracted. Bottom: a pattern with a dislocation. Middle: there may only be one dislocation, but it can move as parameters (e.g., c) are continuously varied and a Hurwitz zero passes through one of the nodes.

Remark 2.6. It should perhaps be mentioned that there is nothing that prevents a zero of $\pi_{N,k}(z)$ from coinciding *exactly* with one of the nodes $x_{N,j} \in X_N$.

Furthermore, due to Proposition 1.3, a spurious zero of $\pi_{N,k}(z)$ in the subinterval K of the saturated region Γ can only occur if the pattern of Hurwitz zeros in K has a dislocation as in the bottom picture in Figure 2.1, in which case the spurious zero must lie in the closed interval $[x_{N,n}, x_{N,n+1}]$ associated with the dislocation. Equivalently, the presence of a spurious zero in $K \subset J \subset \Gamma$ indicates a dislocation in the pattern of Hurwitz zeros.

From the analysis we have presented, it is not clear whether the presence of a dislocation in the pattern of Hurwitz zeros implies that the function $\phi_\Gamma(z)$ has a (spurious) zero in the corresponding closed interval $[x_{N,n}, x_{N,n+1}]$, or equivalently whether the absence of any zeros of $\phi_\Gamma(z)$ in a saturated region Γ means that the lone possible dislocation in the pattern of Hurwitz zeros is indeed absent.

2.3   Near hard edges adjacent to a saturated region

Theorem 2.7 (asymptotics of $\pi_{N,k}(z)$ near hard edges)**.** Assume the conditions enumerated in Sections 1.3.1, 1.3.2, and 1.4.1. Let $c = k/N$ be fixed. If the upper constraint is

achieved at the endpoint $z = b$, then uniformly for $b - CN^{-2/3} < z < b$,

$$\pi_{N,k}(z) = \exp\left(k\int_a^b \log|z - x| d\mu_{min}^c(x)\right)$$
$$\cdot \left[\left(\phi_\Gamma(z) + O\left(\frac{\log(N)}{N^{1/3}}\right)\right)\frac{\Gamma(1/2 - \zeta_b)}{\sqrt{2\pi}e^{\zeta_b}(-\zeta_b)^{-\zeta_b}}2\cos\left(N\pi\int_z^b \rho^0(x)dx\right) \quad (2.4)\right.$$
$$\left. + \text{ exponentially small}\right],$$

as $N \to \infty$, and uniformly for $b < z < b + CN^{-2/3}$,

$$\pi_{N,k}(z) = \exp\left(k\int_a^b \log|z - x| d\mu_{min}^c(x)\right)\left[\phi^{outside}(z)\frac{\sqrt{2\pi}e^{-\zeta_b}\zeta_b^{\zeta_b}}{\Gamma(1/2 + \zeta_b)} + O\left(\frac{\log(N)}{N^{1/3}}\right)\right],$$
$$(2.5)$$

as $N \to \infty$, where $\zeta_b := N\rho^0(b)(z - b)$ and the function $\phi_\Gamma(z)$ is the same function that appears in (2.3). The function $\phi^{outside}(z)$ is real valued, and nonvanishing, and like $\phi_\Gamma(z)$ is constructed from Riemann theta functions and is along with all derivatives uniformly bounded in z as $N \to \infty$. There are similar formulae near the endpoint $z = a$ when the upper constraint is active there. The quantity $\Gamma(1/2 - \zeta_b)$ is the Euler gamma function, while the subscript Γ refers to the saturated region adjacent to the hard edge. The exponentially small term is proportional to $\exp(N[\delta E_c/\delta\mu - \ell_c])$ evaluated on the equilibrium measure. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

It is particularly interesting that the exponential attraction of the zeros to the nodes of orthogonalization in X_N, that we have seen is a feature of the asymptotics in subintervals of $[a, b]$ where the upper constraint is achieved by the equilibrium measure, persists right up to the first and last nodes; in other words if the upper constraint is achieved at $z = a$, then there is a zero exponentially close to $x_{N,0}$ and if the upper constraint is achieved at $z = b$, then there is a zero exponentially close to $x_{N,N-1}$.

More is true, however. From Proposition 1.3, we know that a zero z_0 of $\pi_{N,k}(z)$ that is exponentially close to the first node $x_{N,0}$ must in fact satisfy the strict inequality $z_0 > x_{N,0}$. Similarly, if a zero z_0 is exponentially close to the last node $x_{N,N-1}$, then it must satisfy the strict inequality $z_0 < x_{N,N-1}$. Going back to the discussion in Section 2.2, we see that if there is a hard edge at an endpoint of $[a, b]$, then in the saturated region adjacent to the hard edge there can be no dislocations in the pattern of Hurwitz zeros. This is consistent with the fact (see Theorem 2.3) that the function $\phi_\Gamma(z)$ does not vanish in any saturated region adjacent to endpoints of $[a, b]$ so that there is no spurious zero.

Remark 2.8. The fact that the asymptotic formulae presented in Theorem 2.7 are in terms of the Euler gamma function is directly related to the discrete nature of the weights. In a

sense, the poles of the functions $\Gamma(1/2 \pm \zeta_a)$ and $\Gamma(1/2 \pm \zeta_b)$ are "shadows" of the poles of the Riemann-Hilbert problem that we will discuss in Section 4.

3   Applications

3.1   Discrete orthogonal polynomial ensembles

Recall that $X_N = \{x_{N,n}\}_{n=0}^{N-1}$ is the set of nodes in (a, b). In this section, we use the notation $w_N(x)$ for a weight on X_N; to connect with our previous notation, note simply that for a node $x = x_{N,j} \in X_N$,

$$w_N(x) = w_{N,j}. \tag{3.1}$$

Consider the joint probability distribution of finding k particles, say $P_1, \ldots, P_k$, at respective positions $x_1, \ldots, x_k$ in X_N, to be given by the following expression:

$$
\begin{aligned}
&\mathbb{P}\,(\text{particle } P_j \text{ lies at the site } x_j, \text{ for } j = 1, \ldots, k) \\
&\quad = p^{(N,k)}(x_1, \ldots, x_k) \\
&\quad := \frac{1}{Z_{N,k}} \prod_{1 \le i < j \le k} (x_i - x_j)^2 \prod_{j=1}^{k} w_N(x_j),
\end{aligned} \tag{3.2}
$$

(we are using the symbol $\mathbb{P}\,(\text{event})$ to denote the probability of an event) where $Z_{N,k}$ is a normalization constant (or partition function) chosen so that

$$\sum_{\text{admissable configurations of } P_1, \ldots, P_k} p^{(N,k)}(x_1, \ldots, x_k) = 1. \tag{3.3}$$

Since the distribution function is symmetric in all x_j, we can consider the particles P_j to be either distinguishable or indistinguishable, and only the normalization constant will depend on this choice (the meaning of "admissable configurations" in (3.3) is different in the two cases). The statistical ensemble associated with the density function (3.2) is called a *discrete orthogonal polynomial ensemble*.

Discrete orthogonal polynomial ensembles arise in a number of specific contexts (see, e.g., [5, 16, 17, 18]), with particular choices of the weight function $w_N(\cdot)$ related (in cases we are aware of) to classical discrete orthogonal polynomials. It is of some theoretical interest to determine properties of the ensembles that are more or less independent of the particular choice of weight function, at least within some class. Such properties are said to support the conjecture of *universality* within the class of weight functions under consideration.

Some common properties of discrete orthogonal polynomial ensembles can be read off immediately from formula (3.2). For example, the presence of the Vandermonde factor means that the probability of finding two particles at the same site in X_N is zero. Thus, a discrete orthogonal polynomial ensemble always describes an exclusion process. This phenomenon is the discrete analogue of the familiar *level repulsion* phenomenon in random matrix theory. Also, since the weights are associated with nodes, the interpretation is that configurations where particles are concentrated in sets of nodes where the weight is larger are more likely.

The goal of this section is to establish asymptotic formulae for various statistics associated with the ensemble (3.2) for a general class of weights in the continuum limit $N \to \infty$ with the number of particles k chosen so that for some fixed rational $c \in (0, 1)$, we have $k = cN$. Note that the number of particles k will have the same role as the degree of orthogonal polynomials k. We use the same assumptions on the nodes and weights as in the rest of the paper (see Sections 1.3.1, 1.3.2, and 1.4.1). The main idea is that, as is well known, the formulae for all relevant statistics of ensembles of the form (3.2) can be written explicitly in terms of the discrete orthogonal polynomials associated with the nodes X_N and the weights $w_{N,j} = w_N(x_{N,j})$.

To relate the statistics of interest to the discrete orthogonal polynomials, we first define the so-called reproducing kernel (Christoffel-Darboux kernel)

$$K_{N,k}(x, y) := \sqrt{w_N(x)w_N(y)} \sum_{n=0}^{k-1} p_{N,n}(x)p_{N,n}(y), \tag{3.4}$$

for x, y in the nodes. Using the Christoffel-Darboux formula [27], which holds for all orthogonal polynomials, even in the discrete case, the sum on the right telescopes:

$$\begin{aligned}
K_{N,k}(x, y) &= \sqrt{w_N(x)w_N(y)} \frac{c_{N,k-1}^{(k-1)}}{c_{N,k}^{(k)}} \cdot \frac{p_{N,k}(x)p_{N,k-1}(y) - p_{N,k-1}(x)p_{N,k}(y)}{x - y} \\
&= \sqrt{w_N(x)w_N(y)} \frac{\pi_{N,k}(x) \cdot c_{N,k-1}^{(k-1)} p_{N,k-1}(y) - c_{N,k-1}^{(k-1)} p_{N,k-1}(x) \cdot \pi_{N,k}(y)}{x - y}.
\end{aligned} \tag{3.5}$$

Standard calculations (see, e.g., [23, 29]) of random matrix theory, in the case of so-called $\beta = 2$ ensembles, yield the following exact formulae. The m-*point correlation function* defined for $m \leq k$ by

$$R_m^{(N,k)}(x_1, \ldots, x_m) := \frac{k!}{(k-m)!} \sum_{(x_{m+1}, \ldots, x_k) \in X_N^{k-m}} p^{(N,k)}(x_1, \ldots, x_k), \tag{3.6}$$

where X_N^p denotes the pth Cartesian power of X_N, can be expressed in terms of the discrete orthogonal polynomials by the formula

$$R_m^{(N,k)}(x_1, \ldots, x_m) = \det \left(K_{N,k}(x_i, x_j) \right)_{1 \le i,j \le m}. \tag{3.7}$$

$B \subset X_N$, the one-point correlation function has the following interpretation:

$$\sum_{x \in B} R_1^{(N,k)}(x) = \mathbb{E} \text{ (number of particles in } B), \tag{3.8}$$

where $\mathbb{E}$ denotes the expected value. Similarly, the two-point correlation function has the following interpretation:

$$\sum_{x,y \in B} R_2^{(N,k)}(x,y) = \mathbb{E} \text{ (number of (ordered) pairs of particles in } B). \tag{3.9}$$

Furthermore, the statistics defined for a set $B \subset X_N$ and $m \le \min(\#B, k)$:

$$A_m^{(N,k)}(B) := \mathbb{P} \text{ (there are precisely } m \text{ particles in the set } B) \tag{3.10}$$

(this probability is automatically zero if $m > \#B$ by exclusion) is well known to be expressible by the exact formula

$$A_m^{(N,k)}(B) = \frac{1}{m!} \left(-\frac{d^m}{dt^m} \right) \Big|_{t=1} \det \left(1 - tK_{N,k} \big|_B \right), \tag{3.11}$$

where $K_{N,k}$ is the operator (in this case a finite matrix, since B is contained in the finite set X_N) acting in $\ell^2(X_N)$ given by the kernel $K_{N,k}(x,y)$, and $K_{N,k}|_B$ denotes the restriction of $K_{N,k}$ to $\ell^2(B)$.

This is by no means an exhaustive list of statistics that can be directly expressed in terms of the orthogonal polynomials associated with the (discrete) weight $w_N(\cdot)$. For example, one may consider the fluctuations and in particular the variance of the number of particles in an interval $B \subset X_N$. The continuum limit asymptotics for this statistics were computed in [18] for the Krawtchouk ensemble (see Proposition 2.5 of that paper) with the result that the fluctuations are Gaussian; it would be of some interest to determine whether this is a special property of the Krawtchouk ensemble or a universal property of a large class of ensembles. Also, there are convenient formulae for statistics associated with the spacings between particles; the reader can find such formulae in [8, Section 5.6].

Depending on the location of interest, we have different results. We distinguish again three regions: bands, voids, and saturated regions.

3.1.1   In a band

Theorem 3.1 (universality of the discrete sine kernel). For a node $x \in X_N$ lying in a band I,

$$K_{N,k}(x,x) = \frac{c}{\rho^0(x)} \frac{d\mu_{min}^c}{dx}(x) \left(1 + O\left(\frac{\log(N)}{N}\right)\right), \tag{3.12}$$

where the error is uniform in compact subsets. For distinct nodes x and y in I,

$$K_{N,k}(x,y) = \frac{O(1)}{N \cdot (x-y)}, \tag{3.13}$$

where $O(1)$ is uniform for x and y in a compact subsets. Also with a given node $x \in I$, and for ξ and η such that

$$x + \frac{\xi}{N\rho^0(x)K_{N,k}(x,x)} \in X_N, \qquad x + \frac{\eta}{N\rho^0(x)K_{N,k}(x,x)} \in X_N, \tag{3.14}$$

we have

$$\frac{1}{K_{N,k}(x,x)} K_{N,k}\left(x + \frac{\xi}{N\rho^0(x)K_{N,k}(x,x)}, x + \frac{\eta}{N\rho^0(x)K_{N,k}(x,x)}\right)$$
$$= \frac{\sin(\pi \cdot (\xi - \eta))}{\pi \cdot (\xi - \eta)} + O\left(\frac{\log(N)}{N}\right), \tag{3.15}$$

where the error is uniform for x in a compact subset of the band I, and ξ and η in a compact set of $\mathbb{R}$. $\qquad\square$

Remark 3.2. Let $\psi_{min}^c = d\mu_{min}^c/dx$. By the same analysis, we have the same limit for

$$\frac{\rho^0(x)}{c\psi_{min}^c(x)} K_{N,k}\left(x + \frac{\xi}{Nc\psi_{min}^c(x)}, x + \frac{\eta}{Nc\psi_{min}^c(x)}\right) \tag{3.16}$$

with the same error bound.

Remark 3.3. We believe that the logarithmic term $\log(N)$ in the error can be replaced by 1 whenever the nodes are equally spaced (see Remark 2.2), and it may be the case that such an improved estimate holds more generally. In any case, the improved factor of $1/N$ compared to either $1/N^{1/3}$ or $1/N^{2/3}$, as one might expect from the form of the error term $\varepsilon_N(z)$ in Theorems 2.1 and 2.3, is due to the particular structure of the kernel $K_{N,k}$. Operators with this special type of kernel are called *integrable operators* (see [7, 15]).

Let the operator $\mathcal{S}_x$ act on $\ell^2(\mathbb{Z})$ with the kernel (see [4, 23])

$$\mathcal{S}_x(i,j) = \frac{\sin\left(\frac{\pi c \psi_{\min}^c(x)}{\rho^0(x)} \cdot (i-j)\right)}{\pi \cdot (i-j)}, \quad i,j \in \mathbb{Z}. \tag{3.17}$$

Recall formula (3.11) for $A_m^{(N,k)}(B)$ and its interpretation (3.10) as a probability.

Theorem 3.4 (asymptotics of local occupation probabilities). Let $B_N \subset X_N$ be a set of M nodes of the form

$$B_N = \left\{ x_{N,j}, x_{N,j+k_1}, x_{N,j+k_2}, \ldots, x_{N,j+k_{M-1}} \right\}, \tag{3.18}$$

where $\#B_N = M$ is independent of N, and where

$$0 < k_1 < k_2 < \cdots < k_{M-1} \quad \text{all in } \mathbb{Z} \tag{3.19}$$

are also all independent of N. Set $\mathbb{B}_N := \{0, k_1, k_2, \ldots, k_{M-1}\} \subset \mathbb{Z}$. Suppose also that as $N \to \infty$, $x_{N,j} = \min B_N \to x$ with x lying in a band (and hence the same holds for $x_{N,j+k_{M-1}} = \max B_N$). Then, as $N \to \infty$,

$$\det\left(1 - tK_{N,k}\big|_{B_N}\right) = \det\left(1 - t\mathcal{S}_x\big|_{\mathbb{B}_N}\right) + O\left(\frac{\log(N)}{N}\right), \tag{3.20}$$

for t in a compact set in $\mathbb{C}$, and

$$A_m^{(N,k)}(B_N) = \frac{1}{m!}\left(-\frac{d}{dt}\right)^m\bigg|_{t=1} \det\left(1 - t\mathcal{S}_x\big|_{\mathbb{B}_N}\right) + O\left(\frac{\log(N)}{N}\right). \tag{3.21}$$
$$\square$$

3.1.2 In voids and saturated regions

Theorem 3.5 (exponential asymptotics of the one-point function in voids). Let Γ be a void interval. For each compact subset F of Γ, there is a constant $K_F > 0$ such that

$$K_{N,k}(x,x) = O\left(e^{-K_F N}\right), \quad \text{as } N \longrightarrow \infty, \tag{3.22}$$

holds for all nodes $x \in X_N \cap F$. Also, for distinct nodes x, y in $X_N \cap F$, we have

$$K_{N,k}(x,y) = \frac{O\left(e^{-K_F N}\right)}{x - y}. \tag{3.23}$$
$$\square$$

Thus, the one-point function is exponentially small in void intervals as $N \to \infty$, going to zero with a decay rate that is determined by the size of $\delta E_c/\delta\mu - \ell_c$ at the node x.

Theorem 3.6 (exponential asymptotics of the one-point function in saturated regions)**.**
Let Γ be a saturated region. For each compact subset F of Γ, there is a constant $K_F > 0$ such
that

$$K_{N,k}(x,x) = 1 + O\left(e^{-K_F N}\right), \quad \text{as } N \longrightarrow \infty, \tag{3.24}$$

holds for all nodes $x \in X_N \cap F$. Also, for distinct nodes x, y in $X_N \cap F$, we have

$$K_{N,k}(x,y) = \frac{O(e^{-K_F N})}{x-y}. \tag{3.25}$$

$\square$

Therefore, the one-point function is exponentially close to one in saturated re-
gions.

3.2 Random rhombus tiling of a hexagon

Let a, b, c be positive integers, and consider a hexagon with sides of lengths that proceed
in counterclockwise order, b, a, c, b, a, c. All interior angles of this hexagon are equal and
measure $2\pi/3$ radians. We call this an abc-*hexagon*. See Figure 3.1 for an example of an
abc-hexagon. We denote by $\mathcal{L}$ the lattice points indicated in Figure 3.1. By definition, $\mathcal{L}$
includes the points on the sides (P_6, P_1), (P_1, P_2), (P_2, P_3), and (P_3, P_4), but excludes the
points on the sides (P_4, P_5) and (P_5, P_6).

Consider tiling the abc-hexagon with rhombi having sides of unit length. Such
rhombi come in three different types (orientations) that we refer to as type I, type II,
and type III; see Figure 3.2. Rhombi of types I and II are sometimes collectively called
horizontal rhombi, while rhombi of type III are sometimes called *vertical rhombi*. The
"position" of each rhombus tile in the hexagon is a specific lattice point in $\mathcal{L}$ defined as
indicated in Figure 3.2. See Figure 3.3 for an example of a rhombus tiling.

MacMahon's formula [22] gives the total number of all possible rhombus tilings
of the abc-hexagon as the expression

$$\prod_{i=1}^{a}\prod_{j=1}^{b}\prod_{k=1}^{c} \frac{i+j+k-1}{i+j+k-2}. \tag{3.26}$$

Consider the set of all rhombus tilings equipped with uniform probability. It is of some
current interest to determine the behavior of various corresponding statistics of this en-
semble in the limit as $a, b, c \to \infty$.

In the scaling limit of $n \to \infty$ where

$$a = \alpha n, \qquad b = \beta n, \qquad c = \gamma n, \tag{3.27}$$

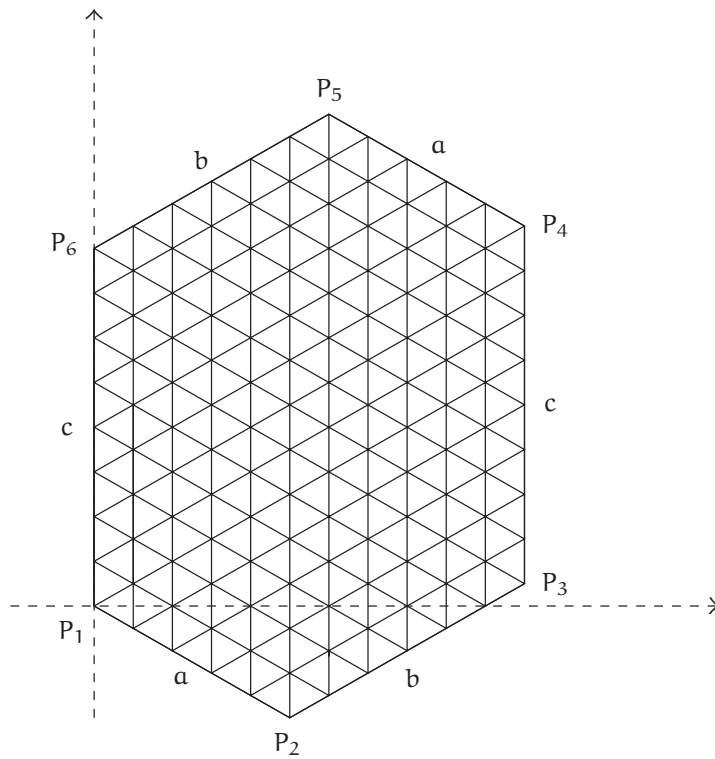**Figure 3.1**   The abc-hexagon with vertices $P_1, \ldots, P_6$, and the lattice $\mathcal{L}$.

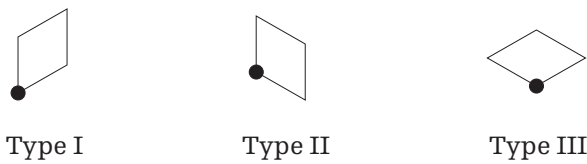Type I   Type II   Type III

Figure 3.2   The three types of rhombi; the position of each rhombus is indicated with a dot.

with fixed $\alpha, \beta, \gamma > 0$, the regions near the six corners are "frozen" or "polar zones," while the inside of the hexagon is "temperate." Cohn, Larsen, and Propp [6] showed that in such a limit, the expected shape of the boundary of the frozen regions is given by the inscribed ellipse. Moreover, the same authors also computed the expected number of vertical rhombi in an arbitrary set $U \in \mathbb{R}^2$. However, this calculation was provided without specific error bounds. In [18], Johansson proved a large deviation result for the boundary shape, and also proved weak convergence of the marginal probability of finding, say,
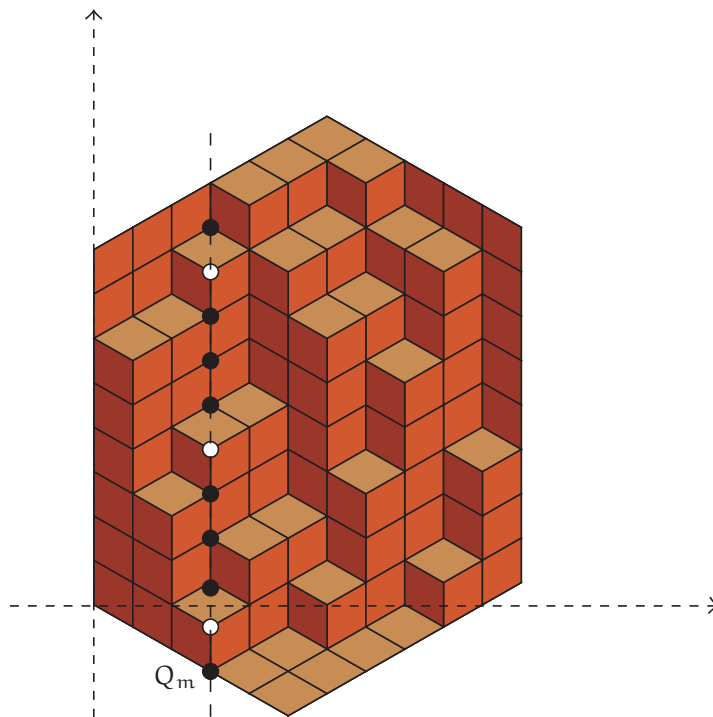
Figure 3.3   A rhombus tiling of the abc-hexagon, and the lattice $\mathcal{L}_m$ when $m = 3$; holes are represented by white dots and particles are represented by black dots.

a vertical tile near a given location in a temperate region. The same paper also contains an investigation of the related Aztec diamond tiling model and a proof that the fluctuation of the boundary in this model is governed (in a proper scaling limit) by the so-called Tracy-Widom law for large random matrices from the Gaussian unitary ensemble [28]. The same is expected to be true for rhombus tilings of hexagons, but this is still open.

In [18], Johansson expresses the induced probability for a given configuration of vertical or horizontal rhombi on a given sublattice in terms of discrete orthogonal polynomial ensembles with Hahn or associated Hahn weights. Even though the Hahn weight is a classical weight, the relevant asymptotics for Hahn polynomials have not been previously established. However, the asymptotics of the previous sections may now be applied to the special case of the Hahn polynomials, and this yields new results for the asymptotic properties of the hexagon tiling problem (see Theorems 3.8 and 3.10 below).

We first state the relation between hexagon tiling and discrete orthogonal polynomial ensembles. We will assume, without loss of generality that $a \geq b$ (by the symmetry of the hexagon, the case when $a \leq b$ is completely analogous). Consider the mth

vertical line of the lattice $\mathcal{L}$ counted from the left. We denote by $\mathcal{L}_m$ the intersection of this line and $\mathcal{L}$. In a given tiling, the points in $\mathcal{L}_m$ correspond to positions (in the sense defined above) of a number of rhombi of types I, II, and III. We call the positions of horizontal rhombi the *particles*, and the positions of vertical rhombi the *holes*. See Figure 3.3 for an example of $\mathcal{L}_m$ when $m = 3$, illustrating the corresponding particles and holes.

The uniform probability distribution on the ensemble of tilings induces the probability distribution for finding particles and holes at particular locations in the one-dimensional finite lattice $\mathcal{L}_m$. A surprising result is due to Johansson [18] which states that the induced probability distribution functions for holes and particles are both discrete orthogonal polynomial ensembles with Hahn and associated Hahn weight functions, respectively (see (1.4) and (1.5)).

Let Q_m be the lowest point in the sublattice $\mathcal{L}_m$. On the sublattice $\mathcal{L}_m$, there are always c particles and L_m holes. We set $\gamma_m = c + L_m - 1$. Now, let $x_1 < \cdots < x_c$, where $x_j \in \{0, 1, 2, \ldots, \gamma_m\}$, denote the (ordered) distances of the particles in $\mathcal{L}_m$ from Q_m, and let $\xi_1 < \cdots < \xi_{L_m}$, where $\xi_j \in \{0, 1, 2, \ldots, \gamma_m\}$, denote the distances of the holes in $\mathcal{L}_m$ from Q_m. In particular, we then have $\{x_1, \ldots, x_c\} \cup \{\xi_1, \ldots, \xi_{L_m}\} = \{0, 1, 2, \ldots, \gamma_m\}$. Let $\tilde{P}_m(x_1, \ldots, x_c)$ denote the probability of finding the particle configuration $x_1, \ldots, x_c$, and let $P_m(\xi_1, \ldots, \xi_{L_m})$ denote the probability of finding the hole configuration $\xi_1, \ldots, \xi_{L_m}$.

Proposition 3.7 (see [18, Theorem 4.1]). Let $a, b, c \geq 1$ be given integers with $a \geq b$. Set $a_m := |a - m|$ and $b_m := |b - m|$. Then

$$\tilde{P}_m(x_1, \ldots, x_c) = \frac{1}{\tilde{Z}_m} \prod_{1 \leq j < k \leq c} (x_j - x_k)^2 \prod_{j=1}^{c} \tilde{w}(x_j), \tag{3.28}$$

where $\tilde{Z}_m$ is the normalization constant (partition function), and where the weight function is the associated Hahn weight

$$\tilde{w}(n) := w_{N,n}^{\text{Assoc}}(a_m + 1, b_m + 1) = \frac{\tilde{C}}{n!(a_m + n)!(N - n - 1)!(N - n - 1 + b_m)!}, \tag{3.29}$$

for a certain constant $\tilde{C}$. Also,

$$P_m(\xi_1, \ldots, \xi_{L_m}) = \frac{1}{Z_m} \prod_{1 \leq j < k \leq L_m} (\xi_j - \xi_k)^2 \prod_{j=1}^{L_m} w(\xi_j), \tag{3.30}$$

where Z_m is the normalization constant, and where the weight function is the Hahn weight

$$w(n) := w_{N,n}^{\text{Hahn}}(a_m + 1, b_m + 1) = C\frac{(n + a_m)!(N - n - 1 + b_m)!}{n!(N - n - 1)!}, \tag{3.31}$$

for a certain constant C. □

With

$$m = \tau n, \tag{3.32}$$

for $\gamma > 0$, the scaling (3.27) is precisely the same scaling that we analyzed in Section 3.1. Also we can explicitly compute the equilibrium measure for Hahn and associated Hahn using either the result of [21], or solving the variational problem as in [9], which will appear in the full version of this paper. The calculations of the equilibrium measure and the one-point correlation function imply that as $n \to \infty$, the one-dimensional lattice $\mathcal{L}_m$, after rescaling to finite size independent of n, consists of three disjoint intervals: one band, surrounded by two gaps (either saturated regions or voids, depending on parameters). The saturated regions and voids correspond to the frozen regions or polar zones, while the central band is a temperate region. Hence, in particular, the endpoints of the band when considered as functions of τ determine the typical shape of the boundary between the polar and temperate zones of the rescaled abc-hexagon. Also the one-point function converges pointwise except at the band edges, or at the boundary, to the equilibrium measure. This was conjectured in [18] (including the edges), in which weak convergence was obtained. Moreover, our computation of the one-point correlation function provides the relevant error bounds, when we consider sets U contained in a single line $\mathcal{L}_m$. One expects that, with additional analysis of the same formulae, it should be possible to show that the error is locally uniform with respect to τ, in which case the same bounds should hold for more general regions $U \in \mathbb{R}^2$. We state our result in this direction as follows.

Theorem 3.8 (strong asymptotics with explicit error bounds)**.** On the line $\mathcal{L}_m$, where $m = \tau n$ and τ is fixed as $n \to \infty$, the scaled holes ξ_j/n lying in the polar zones, uniformly bounded away from the rescaled expected boundary between the polar and temperate zones, have a one-point function asymptotically convergent to either 1 (in the polar zones near the vertices P_2 and P_5), or to 0 (in the polar zones near the vertices P_1, P_3, P_4, and P_6), with an exponential rate of convergence of the order $O(e^{-Kn})$ for some constant $K > 0$. The one-point correlation function for the scaled holes ξ/n in the temperate zone converges to the corresponding equilibrium measure with an error of the order

$O(1/n)$, which is uniformly valid away from the rescaled expected boundary between the polar and temperate zones.   □

Remark 3.9. We give the above error estimate as $O(1/n)$ rather than $O(\log(n)/n)$ because the Hahn and associated Hahn polynomials are orthogonal on a set of nodes X_n that are equally spaced.

In the temperate zone, in addition to the one-point function, which is the marginal distribution, we can control all k-point correlation functions under proper scaling. One such consequence is the following theorem on the scaling limit for the locations of the holes.

Theorem 3.10 (discrete sine kernel correlations)**.** Let $x > 0$ be rational such that $nx \in \mathbb{Z}_N$ and such that nx is in the temperate zone away from the expected boundary between the polar and temperate zones with uniform order in n. Let $B_m = \{nx, nx + j_1, nx + j_2, \ldots, nx + j_M\}$, and set $\mathbb{B} = \{0, j_1, j_2, \ldots, j_M\}$. Then

$$\lim_{n \to \infty} \mathbb{P}\left(\text{there are precisely } p \text{ holes in the set } B_m\right) = \frac{1}{p!}\left(-\frac{d}{dt}\right)^p\Bigg|_{t=1} \det\left(1 - t\mathcal{S}|_{\mathbb{B}}\right),$$

(3.33)

where $\mathcal{S}$ acts on $\ell^2(\mathbb{Z})$ with the kernel

$$\mathcal{S}(i, j) = \frac{\sin(c(x)(i - j))}{\pi(i - j)},$$

(3.34)

for some constant $c(x)$.   □

Remark 3.11. All of the results we have written down for holes have analogous statements in terms of particles using the duality relation between the Hahn and associated Hahn weights that will be explained in Section 4.2.

Remark 3.12. Once one obtains the asymptotics near the band edge of the equilibrium measure for discrete orthogonal polynomial ensembles, fluctuation statistics of the boundary curve will be computable. It is conjectured in [18] that the limiting law at the band edge is the Tracy-Widom distribution known from the Gaussian unitary ensemble of random matrix theory.

4   Riemann-Hilbert problems for discrete orthogonal polynomials

In this section, we discuss the main ideas of asymptotic analysis of discrete orthogonal polynomials via a Riemann-Hilbert problem.

4.1 The fundamental Riemann-Hilbert problem

We first introduce the Riemann-Hilbert problem characterization of discrete orthogonal polynomials. For $k \in \mathbb{Z}$, consider the matrix $\mathbf{P}(z; N, k)$ solving the following problem, which is a discrete version of the analogous problem for continuous weights first used in [14]. This Riemann-Hilbert problem with pole conditions has been known for a while to be solved in terms of discrete orthogonal polynomials. An example is [3] in which the authors used the Riemann-Hilbert problem to derive difference equations for a class of discrete orthogonal polynomials.

Riemann-Hilbert Problem 1. Find a 2×2 matrix $\mathbf{P}(z; N, k)$ with the following properties:

(1) *Analyticity.* $\mathbf{P}(z; N, k)$ is an analytic function of z for $z \in \mathbb{C} \setminus X_N$.

(2) *Normalization.* As $z \to \infty$,

$$\mathbf{P}(z; N, k) \begin{pmatrix} z^{-k} & 0 \\ 0 & z^k \end{pmatrix} = \mathbb{I} + O\left(\frac{1}{z}\right). \tag{4.1}$$

(3) *Singularities.* At each node $x_{N,j}$, the first column of $\mathbf{P}$ is analytic and the second column of $\mathbf{P}$ has a simple pole, where the residue satisfies the condition

$$\operatorname*{Res}_{z = x_{N,j}} \mathbf{P}(z; N, k) = \lim_{z \to x_{N,j}} \mathbf{P}(z; N, k) \begin{pmatrix} 0 & w_{N,j} \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & w_{N,j} P_{11}(x_{N,j}; N, k) \\ 0 & w_{N,j} P_{21}(x_{N,j}, N, k) \end{pmatrix}, \tag{4.2}$$

for $j = 0, \ldots, N - 1$.

Proposition 4.1. Riemann-Hilbert Problem 1 has a unique solution when $0 \le k \le N - 1$. In this case,

$$\mathbf{P}(z; N, k) = \begin{pmatrix} \pi_{N,k}(z) & \displaystyle\sum_{j=0}^{N-1} \frac{w_{N,j} \pi_{N,k}(x_{N,j})}{z - x_{N,j}} \\ c_{N,k-1}^{(k-1)} p_{N,k-1}(z) & \displaystyle\sum_{j=0}^{N-1} \frac{w_{N,j} c_{N,k-1}^{(k-1)} p_{N,k-1}(x_{N,j})}{z - x_{N,j}} \end{pmatrix} \tag{4.3}$$

if $k > 0$ and

$$\mathbf{P}(z; N, 0) = \begin{pmatrix} 1 & \displaystyle\sum_{j=0}^{N-1} \frac{w_{N,j}}{z - x_{N,j}} \\ 0 & 1 \end{pmatrix}. \tag{4.4}$$

We analyze this Riemann-Hilbert problem asymptotically as $N, k \to \infty$ by adapting the Deift-Zhou procedure developed in [12] and subsequent work. Due to the conditions on the pole and the separation of the zeros of discrete orthogonal polynomials, we have two difficulties which we mentioned in Section 1.2. In the following two sections, we describe the main techniques we developed to overcome these difficulties. A complete asymptotic analysis of the above Riemann-Hilbert problem will appear in the full version of this paper.

In order to apply the usual Deift-Zhou method, we will transform the above Riemann-Hilbert problem into a Riemann-Hilbert problem with jump conditions on continuous contours: a transformation from $P \mapsto R$. For this new Riemann-Hilbert problem, the formal limit of accumulation of nodes can be rigorously justified. However, in addition to the continuum limit of $N \to \infty$ (N being the number of nodes), we simultaneously take the large degree limit $k \to \infty$. In the analysis of [10, 11], a method for this limit is to conjugate the Riemann-Hilbert problem with the so-called g-function that is defined as a log transform of the equilibrium measure. It was crucial in the analysis of [10, 11] for the continuous orthogonal polynomials that the equilibrium measure has only a lower constraint. Actually the lower constraint yields an exponentially decaying factor. Hence, the upper constraint condition (1.17) of the equilibrium measure for discrete orthogonal polynomials generates an exponentially growing factor. In order to replace an exponentially growing term with an exponentially decaying term, we introduce another transformation *before* we map P to R: we will introduce an intermediate Riemann-Hilbert problem for Q so that $P \mapsto Q \mapsto R$ as an exact sequence of transformations. In Section 4.2, we discuss the transformation $P \mapsto Q$ of reversing the triangularity of residue matrices, that will eventually work in our favor turning exponentially growing terms into exponentially decaying terms. In Section 4.3, we discuss the transformation $Q \mapsto R$ from a Riemann-Hilbert problem with residue conditions to a Riemann-Hilbert problem with jumps on continuous contours.

4.2   Selectively reversing triangularity of residue matrices

Riemann-Hilbert Problem 1 involves residue matrices that are upper-triangular. It will be advantageous in general to modify the matrix $P(z; N, k)$ in order to arrive at a new Riemann-Hilbert problem in which we have selectively reversed the triangularity of the residue matrices near certain individual nodes $x_{N,j}$. Let $\Delta \subset \mathbb{Z}_N$ where $\mathbb{Z}_N := \{0, 1, 2, \ldots, N - 1\}$ and denote the number of elements in Δ by $\#\Delta$ and the complementary set $\mathbb{Z}_N \setminus \Delta$ by ∇. We will reverse the triangularity for those nodes $x_{N,j}$ for which $j \in \Delta$. Consider the matrix $Q(z; N, k)$ related to the solution $P(z; N, k)$ of Riemann-Hilbert Problem 1 as

follows:

$$\mathbf{Q}(z; N, k) := \mathbf{P}(z; N, k) \left[\prod_{n \in \Delta} (z - x_{N,n}) \right]^{-\sigma_3}$$

$$= \mathbf{P}(z; N, k) \begin{pmatrix} \prod_{n \in \Delta} (z - x_{N,n})^{-1} & 0 \\ 0 & \prod_{n \in \Delta} (z - x_{N,n}) \end{pmatrix}. \tag{4.5}$$

It is direct to check that the matrix $\mathbf{Q}(z; N, k)$ is, for $k \in \mathbb{Z}_N$, the unique solution of the following Riemann-Hilbert problem.

Riemann-Hilbert Problem 2. Given a subset Δ of $\mathbb{Z}_N$ of cardinality $\#\Delta$, find a 2×2 matrix $\mathbf{Q}(z; N, k)$ with the following properties:

(1) *Analyticity.* $\mathbf{Q}(z; N, k)$ is an analytic function of z for $z \in \mathbb{C} \setminus X_N$.

(2) *Normalization.* As $z \to \infty$,

$$\mathbf{Q}(z; N, k) \begin{pmatrix} z^{\#\Delta - k} & 0 \\ 0 & z^{k - \#\Delta} \end{pmatrix} = \mathbb{I} + O\left(\frac{1}{z}\right). \tag{4.6}$$

(3) *Singularities.* At each node $x_{N,j}$, the matrix $\mathbf{Q}$ has a simple pole. If $j \in \nabla$ where $\nabla := \mathbb{Z}_N \setminus \Delta$, then the first column is analytic at $x_{N,j}$ and the pole is in the second column such that the residue satisfies the condition

$$\operatorname*{Res}_{z = x_{N,j}} \mathbf{Q}(z; N, k) = \lim_{z \to x_{N,j}} \mathbf{Q}(z; N, k), \begin{pmatrix} 0 & w_{N,j} \prod_{n \in \Delta} (x_{N,j} - x_{N,n})^2 \\ 0 & 0 \end{pmatrix} \tag{4.7}$$

for $j \in \nabla$. If $j \in \Delta$, then the second column is analytic at $x_{N,j}$ and the pole is in the first column such that the residue satisfies the condition

$$\operatorname*{Res}_{z = x_{N,j}} \mathbf{Q}(z; N, k) = \lim_{z \to x_{N,j}} \mathbf{Q}(z; N, k) \begin{pmatrix} 0 & 0 \\ \dfrac{1}{w_{N,j}} \prod_{\substack{n \in \Delta \\ n \neq j}} (x_{N,j} - x_{N,n})^{-2} & 0 \end{pmatrix}, \tag{4.8}$$

for $j \in \Delta$. $\qquad\qquad \square$

Note that the (21)-entry of the residue matrix in (4.8) is the reciprocal of the (12)-entry of the residue matrix in (4.7). When we make the choice that Δ contains the
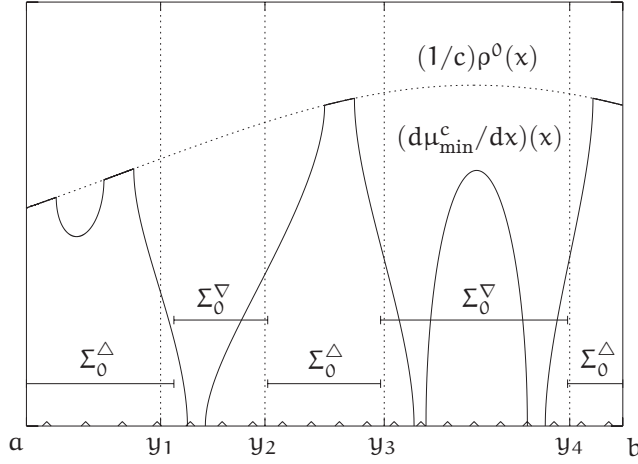
Figure 4.1   A schematic diagram showing the relation of
the minimizer $\mu_{min}^c(x)$ to the interval systems Σ_0^∇ and $\Sigma_0^\triangle$.
The nodes $x_{N,j} \in (a, b)$ are indicated on the x-axis with
triangles; their density is proportional to the upper con-
straint. The common endpoints of subintervals of Σ_0^∇ and
$\Sigma_0^\triangle$ converge as $N \to \infty$ to the points y_k indicated on the
x-axis.

saturated regions, the effect will be to turn exponentially growing factors into exponen-
tially decaying factors. Let us be more specific about how we choose Δ. In each band I_k
lying between a void and a saturated region, we choose a point y_k, and "quantize" these
to the lattice X_N by associating with each point a sequence $\{y_{k,N}\}_{N=0}^\infty$ converging to y_k as
$N \to \infty$ with elements given by

$$N \int_a^{y_{k,N}} \rho^0(x)dx = \left\lceil N \int_a^{y_k} \rho^0(x)dx \right\rceil, \tag{4.9}$$

where $\lceil u \rceil$ denotes the least integer greater than or equal to u. Thus $y_{k,N}$ lies asymptoti-
cally halfway between two consecutive nodes. For each N, these points are the common
endpoints of two complementary systems of subintervals of (a, b). We denote the union
of open subintervals delineated by these points and containing no saturated regions by
Σ_0^∇. The complementary system of subintervals contains no voids and is denoted by $\Sigma_0^\triangle$.
See Figure 4.1.

Based on this partitioning of (a, b), the specific choice we make is that Δ is the set
of indices $j \in \mathbb{Z}_N$ such that $x_{N,j} \in \Sigma_0^\triangle$.

Remark 4.2. After the completion of this work, we learned that an analogous transformation was used in [5] for a somewhat different asymptotic analysis of a special choice of discrete orthogonal polynomials (Askey-Lesky weights, or general Hahn weights). In [5], the choice of the region Δ to reverse the triangularity was made using intuition from representation theory. (We thank A. Borodin for bringing this to our attention.) For the limit of interest in this paper, we determine the region Δ in order to reverse the triangularity in saturated regions of the equilibrium measure while preserving triangularity in all voids.

4.2.1 Dual families of discrete orthogonal polynomials. The relation between Riemann-Hilbert Problem 1 and Riemann-Hilbert Problem 2 gives rise in a special case to a remarkable duality between pairs of weights $\{w_{N,j}\}$ defined on the same set of nodes and their corresponding families of discrete orthogonal polynomials that comes up in applications. Given nodes X_N and weights $\{w_{N,j}\}$, take $\Delta = \mathbb{Z}_N$ and let

$$\overline{\mathbf{P}}(z; N, \overline{k}) := \sigma_1 \mathbf{Q}(z; N, k)\sigma_1, \quad \overline{k} := N - k. \tag{4.10}$$

Thus, we are reversing the triangularity at all of the nodes, and swapping rows and columns of the resulting matrix. It is easy to check that $\overline{\mathbf{P}}(z; N, \overline{k})$ satisfies

$$\overline{\mathbf{P}}(z; N, \overline{k}) \begin{pmatrix} z^{-\overline{k}} & 0 \\ 0 & z^{\overline{k}} \end{pmatrix} = \mathbb{I} + O\left(\frac{1}{z}\right), \quad \text{as } z \longrightarrow \infty, \tag{4.11}$$

and is a matrix with simple poles in the second column at all nodes, such that

$$\operatorname*{Res}_{z = x_{N,j}} \overline{\mathbf{P}}(z; N, \overline{k}) = \lim_{z \to x_{N,j}} \overline{\mathbf{P}}(z; N, \overline{k}) \begin{pmatrix} 0 & \overline{w}_{N,j} \\ 0 & 0 \end{pmatrix} \tag{4.12}$$

holds for $j \in \mathbb{Z}_N$, where the "dual weights" $\{\overline{w}_{N,j}\}$ are defined by the identity

$$w_{N,j}\overline{w}_{N,j} \prod_{\substack{n=0 \\ n \neq j}}^{N-1} \left(x_{N,j} - x_{N,n}\right)^2 = 1. \tag{4.13}$$

Comparing with Riemann-Hilbert Problem 1, we see that $\overline{P}_{11}(z; N, \overline{k})$ is the monic orthogonal polynomial $\overline{\pi}_{N,\overline{k}}(z)$ of degree $\overline{k}$ associated with the dual weights $\{\overline{w}_{N,j}\}$. In this sense,

families of discrete orthogonal polynomials always come in dual pairs. An explicit relation between the dual polynomials comes from the representation of $\mathbf{P}(z; N, k)$ given by Proposition 4.1:

$$
\begin{aligned}
\overline{\pi}_{N,\overline{k}}(z) &= \overline{P}_{11}\left(z; N, \overline{k}\right) \\
&= P_{22}(z; N, k) \prod_{n=0}^{N-1} \left(z - x_{N,n}\right) \\
&= \sum_{j=0}^{N-1} w_{N,j} \left[c_{N,k-1}^{(k-1)}\right]^2 \pi_{N,k-1}\left(x_{N,j}\right) \prod_{\substack{n=0 \\ n \neq j}}^{N-1} \left(z - x_{N,n}\right).
\end{aligned}
\tag{4.14}
$$

Since the left-hand side is a monic polynomial of degree $\overline{k} = N - k$ and the right-hand side is apparently a polynomial of degree $N - 1$, equation (4.14) furnishes k relations among the weights and the normalization constants $c_{N,k}^{(k)}$.

In particular, if we evaluate (4.14) for $z = x_{N,l}$ for some $l \in \mathbb{Z}_N$, then only one term from the sum on the right-hand side survives and we find

$$
\overline{\pi}_{N,\overline{k}}\left(x_{N,l}\right) = \left[c_{N,k-1}^{(k-1)}\right]^2 w_{N,l} \prod_{\substack{n=0 \\ n \neq l}}^{N-1} \left(x_{N,l} - x_{N,n}\right) \cdot \pi_{N,k-1}\left(x_{N,l}\right),
\tag{4.15}
$$

an identity relating values of each discrete orthogonal polynomial and a corresponding dual polynomial at any given node. The identity (4.15) has also been derived by Borodin [2].

Remark 4.3. We want to point out that the notion of duality described here is different from that explained in [25]. The latter generally involves relationships between families of discrete orthogonal polynomials with two different sets of nodes of orthogonalization. For example, the Hahn polynomials are orthogonal on a lattice of equally spaced points, and the polynomials dual to the Hahn polynomials by the scheme of [25] are orthogonal on a quadratic lattice for which $x_{N,n} - x_{N,n-1}$ is proportional to n. However, the polynomials dual to the Hahn polynomials under the scheme described above are the associated Hahn polynomials, which are orthogonal on the same equally-spaced nodes as are the Hahn polynomials themselves. The notion of duality we use in this paper coincides with that described in [2] and is also equivalent to the "hole/particle transformation" considered by Johansson [18].
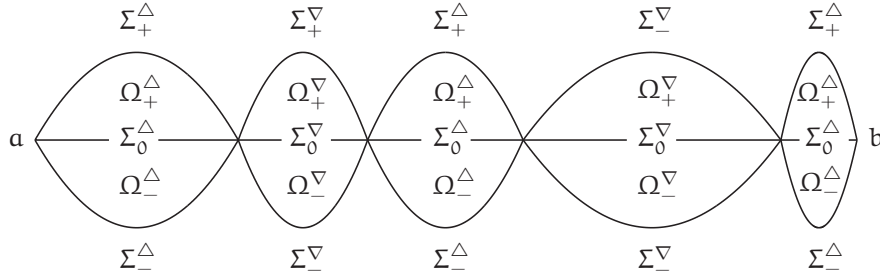
Figure 4.2   The contour Σ consists of the subintervals Σ_0^∇ and $\Sigma_0^\triangle$ as in Figure 4.1 and associated contour segments Σ_+^∇ and $\Sigma_+^\triangle$ in the upper half-plane, and Σ_-^∇ and $\Sigma_-^\triangle$ in the lower half-plane. The enclosed regions $\Omega_\pm^\nabla$ and $\Omega_\pm^\triangle$ are also indicated. The contour Σ lies entirely in the region of analyticity of $V(x)$ and $\rho^0(x)$. All components of Σ are taken to be oriented from left to right.

4.3   Removal of poles in favor of jumps on contours

The deformations in this section are based on similar ones first introduced by one of the authors in [24]. Let the analytic functions $\beta_\pm(z)$ be given by

$$\beta_\pm(z) := \pm i \exp\left(\mp i\pi N \int_z^b \rho^0(s)ds\right). \tag{4.16}$$

Note that by definition, $\beta_+(x_{N,j}) = \beta_-(x_{N,j}) = (-1)^{N-1-j}$ for all $N \in \mathbb{N}$ and $j \in \mathbb{Z}_N$. Consider the contour Σ illustrated in Figure 4.2. From the solution of Riemann-Hilbert Problem 2 we define a new matrix $\mathbf{R}(z)$ as follows. Set

$$\mathbf{R}(z) := \mathbf{Q}(z;N,k)\begin{pmatrix} 1 & -\beta_\pm(z)e^{-NV_N(z)}\dfrac{\prod\limits_{j\in\Delta}(z-x_{N,j})}{\prod\limits_{j\in\nabla}(z-x_{N,j})} \\ 0 & 1 \end{pmatrix}, \quad \text{for } z \in \Omega_\pm^\nabla, \tag{4.17}$$

$$\mathbf{R}(z) := \mathbf{Q}(z;N,k)\begin{pmatrix} 1 & 0 \\ -\beta_\pm(z)e^{NV_N(z)}\dfrac{\prod\limits_{j\in\nabla}(z-x_{N,j})}{\prod\limits_{j\in\Delta}(z-x_{N,j})} & 1 \end{pmatrix}, \quad \text{for } z \in \Omega_\pm^\triangle, \tag{4.18}$$

and for all other z set $\mathbf{R}(z) := \mathbf{Q}(z;N,k)$.

The matrix $\mathbf{R}(z)$ is, for arbitrary $N \in \mathbb{N}$ and $k \in \mathbb{Z}_N$, the unique solution of the following Riemann-Hilbert problem.

Riemann-Hilbert Problem 3. Find a 2×2 matrix $\mathbf{R}(z)$ with the following properties:

(1) *Analyticity*. $\mathbf{R}(z)$ is an analytic function of z for $z \in \mathbb{C} \setminus \Sigma$.

(2) *Normalization*. As $z \to \infty$,

$$\mathbf{R}(z) \begin{pmatrix} z^{\#\Delta - k} & 0 \\ 0 & z^{k - \#\Delta} \end{pmatrix} = \mathbb{I} + O\left(\frac{1}{z}\right). \tag{4.19}$$

(3) *Jump Conditions*. $\mathbf{R}(z)$ takes continuous boundary values on Σ from each connected component of $\mathbb{C}\setminus\Sigma$. Denoting the boundary values taken on the left (right) by $\mathbf{R}_+(z)$ ($\mathbf{R}_-(z)$), we have

$$\mathbf{R}_+(z) = \mathbf{R}_-(z) \begin{pmatrix} 1 & \pm\beta_\pm(z)e^{-NV_N(z)} \dfrac{\prod\limits_{j\in\Delta}(z - x_{N,j})}{\prod\limits_{j\in\nabla}(z - x_{N,j})} \\ 0 & 1 \end{pmatrix}, \quad \text{for } z \in \Sigma_\pm^\nabla, \tag{4.20}$$

$$\mathbf{R}_+(z) = \mathbf{R}_-(z) \begin{pmatrix} 1 & 0 \\ \pm\beta_\pm(z)e^{NV_N(z)} \dfrac{\prod\limits_{j\in\nabla}(z - x_{N,j})}{\prod\limits_{j\in\Delta}(z - x_{N,j})} & 1 \end{pmatrix}, \quad \text{for } z \in \Sigma_\pm^\Delta, \tag{4.21}$$

$$\mathbf{R}_+(z) = \mathbf{R}_-(z) \begin{pmatrix} 1 & (\beta_-(z) - \beta_+(z))e^{-NV_N(z)} \dfrac{\prod\limits_{j\in\Delta}(z - x_{N,j})}{\prod\limits_{j\in\nabla}(z - x_{N,j})} \\ 0 & 1 \end{pmatrix}, \quad \text{for } z \in \Sigma_0^\nabla, \tag{4.22}$$

$$\mathbf{R}_+(z) = \mathbf{R}_-(z) \begin{pmatrix} 1 & 0 \\ (\beta_-(z) - \beta_+(z))e^{NV_N(z)} \dfrac{\prod\limits_{j\in\nabla}(z - x_{N,j})}{\prod\limits_{j\in\Delta}(z - x_{N,j})} & 1 \end{pmatrix}, \quad \text{for } z \in \Sigma_0^\Delta. \tag{4.23}$$

$\square$

Note that all off-diagonal entries of the jump matrices are analytic nonvanishing functions on their respective contours.

The significance of passing from Riemann-Hilbert Problem 2 to Riemann-Hilbert Problem 3 is that all poles have completely disappeared from the problem. All boundary values of $\mathbf{R}(z)$ and the corresponding jump matrices relating them are analytic functions. This means that Riemann-Hilbert Problem 3 is sufficiently similar to that introduced in [14] for the continuous weight case that it may, in principle, be analyzed by methods like those used in [10, 11]. The main obstruction at this point is that the off-diagonal elements of the jump matrices for $\mathbf{R}(z)$ are not exactly of the form $e^{NW(z)}$ for some $W(z)$. This is a consequence of the fact that the sequence of transformations $\mathbf{P} \mapsto \mathbf{Q} \mapsto \mathbf{R}$ is *exact*, and one may observe at this point that the desired form $e^{NW(z)}$ can be achieved by carefully taking a natural continuum limit based on the assumptions on the nodes and weights set out at the beginning of this announcement. In other words, while the jump matrix for $\mathbf{R}$ does not have the desired form, one may introduce an approximate Riemann-Hilbert problem for a matrix $\dot{\mathbf{R}}(z)$ for which the jump matrix indeed has the desired form; part of the analysis then becomes the task of showing that $\mathbf{R}(z)$ and $\dot{\mathbf{R}}(z)$ are "close."

Remark 4.4. The difficulty that prevents us from obtaining asymptotic results near the band edges can be traced back precisely to the fact that the off-diagonal elements of the jump matrices for $\mathbf{R}(z)$ are not exactly of the form $e^{NW(z)}$, but only approximately so. In the local analysis, the (small) discrepancy between the approximate and the exact form prevents us from obtaining the necessary asymptotics up to the boundary of the contours. We hope to be able to overcome this problem in a future work.

The final aspect of our analysis that we would like to briefly describe is our choice of a g-function with which we stabilize the (approximate) Riemann-Hilbert problem for $\dot{\mathbf{R}}(z)$. We introduce a new matrix unknown by the transformation $\dot{\mathbf{S}}(z) := \dot{\mathbf{R}}(z)e^{(\#\Delta-k)g(z)\sigma_3}$ where the g-function is given by

$$g(z) = \int_a^b \log(z-x)\rho(x)dx \tag{4.24}$$

with density determined differently in the two types of intervals Σ_0^{∇} and Σ_0^{Δ}:

$$\rho(x) := \begin{cases} \dfrac{c}{c-d} \dfrac{d\mu_{\min}^c}{dx}(x), & x \in \Sigma_0^{\nabla}, \\[2mm] \dfrac{c}{c-d}\left(\dfrac{d\mu_{\min}^c}{dx}(x) - \dfrac{1}{c}\rho^0(x)\right), & x \in \Sigma_0^{\Delta}. \end{cases} \tag{4.25}$$

Here of course $\mu^c_{\min}$ is the equilibrium measure, $c = k/N$, and

$$d := \int_{\Sigma^\Delta_0} \rho^0(x)dx = \frac{\#\Delta}{N}. \tag{4.26}$$

With this choice of $g(z)$, in conjunction with the choice of the set Δ described above, the jump matrices for $\dot{S}(z)$ are precisely of the type for which the steepest-descent factorization technique can be applied. Also, $\dot{S}(z)$ is now normalized to the identity matrix for large z; the power asymptotics have been removed. The complete details of the subsequent analysis, including rigorous error estimates, will appear in the full version of the paper corresponding to this announcement.

Acknowledgments

References

[1]   M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions*, Dover Publications, New York, 1965.

[2]   A. Borodin, *Duality of orthogonal polynomials on a finite set*, preprint, 2001, http://www.arxiv.org/abs/math.CA/0101125.

[3]   A. Borodin and D. Boyarchenko, *Distribution of the first particle in discrete orthogonal polynomial ensembles*, to appear in Comm. Math. Phys.

[4]   A. Borodin, A. Okounkov, and G. Olshanski, *Asymptotics of Plancherel measures for symmetric groups*, J. Amer. Math. Soc. **13** (2000), no. 3, 481–515.

[5]   A. Borodin and G. Olshanski, *Harmonic analysis on the infinite-dimensional unitary group and determinantal point processes*, preprint, 2001, http://www.arxiv.org/abs/math.RT/0109194.

[6]   H. Cohn, M. Larsen, and J. Propp, *The shape of a typical boxed plane partition*, New York J. Math. **4** (1998), 137–165.

[7]   P. Deift, *Integrable operators*, Differential Operators and Spectral Theory, Amer. Math. Soc. Transl. Ser. 2, vol. 189, American Mathematical Society, Rhode Island, 1999, pp. 69–84.

[8] ———, *Orthogonal Polynomials and Random Matrices: A Riemann-Hilbert Approach*, Courant Lecture Notes in Mathematics, vol. 3, New York University Courant Institute of Mathematical Sciences, New York, 1999.

[9] P. Deift, T. Kriecherbauer, and K. T.-R. McLaughlin, *New results on the equilibrium measure for logarithmic potentials in the presence of an external field*, J. Approx. Theory **95** (1998), no. 3, 388–475.

[10] P. Deift, T. Kriecherbauer, K. T.-R. McLaughlin, S. Venakides, and X. Zhou, *Strong asymptotics of orthogonal polynomials with respect to exponential weights*, Comm. Pure Appl. Math. **52** (1999), no. 12, 1491–1552.

[11] ———, *Uniform asymptotics for polynomials orthogonal with respect to varying exponential weights and applications to universality questions in random matrix theory*, Comm. Pure Appl. Math. **52** (1999), no. 11, 1335–1425.

[12] P. Deift and X. Zhou, *A steepest descent method for oscillatory Riemann-Hilbert problems. Asymptotics for the MKdV equation*, Ann. of Math. (2) **137** (1993), no. 2, 295–368.

[13] P. D. Dragnev and E. B. Saff, *A problem in potential theory and zero asymptotics of Krawtchouk polynomials*, J. Approx. Theory **102** (2000), no. 1, 120–140.

[14] A. S. Fokas, A. R. Its, and A. V. Kitaev, *Discrete Painlevé equations and their appearance in quantum gravity*, Comm. Math. Phys. **142** (1991), no. 2, 313–344.

[15] A. R. Its, A. G. Izergin, V. E. Korepin, and N. A. Slavnov, *Differential equations for quantum correlation functions*, Internat. J. Modern Phys. B **4** (1990), no. 5, 1003–1037.

[16] K. Johansson, *Shape fluctuations and random matrices*, Comm. Math. Phys. **209** (2000), no. 2, 437–476.

[17] ———, *Discrete orthogonal polynomial ensembles and the Plancherel measure*, Ann. of Math. (2) **153** (2001), no. 1, 259–296.

[18] ———, *Non-intersecting paths, random tilings and random matrices*, Probab. Theory Related Fields **123** (2002), no. 2, 225–280.

[19] S. Kamvissis, K. T.-R. McLaughlin, and P. D. Miller, *Semiclassical soliton ensembles for the focusing nonlinear Schrödinger equation*, to appear in Ann. of Math. Stud.

[20] A. B. J. Kuijlaars, *On the finite-gap ansatz in the continuum limit of the Toda lattice*, Duke Math. J. **104** (2000), no. 3, 433–462.

[21] A. B. J. Kuijlaars and E. A. Rakhmanov, *Zero distributions for discrete orthogonal polynomials*, J. Comput. Appl. Math. **99** (1998), no. 1-2, 255–274.

[22] P. A. MacMahon, *Combinatory Analysis*, Two volumes (bound as one), Chelsea Publishing Co., New York, 1960.

[23] M. L. Mehta, *Random Matrices*, 2nd ed., Academic Press, Massachusetts, 1991.

[24] P. D. Miller, *Asymptotics of semiclassical soliton ensembles: rigorous justification of the WKB approximation*, Internat. Math. Res. Notices **2002** (2002), no. 8, 383–454.

[25] A. F. Nikiforov, S. K. Suslov, and V. B. Uvarov, *Classical Orthogonal Polynomials of a Discrete Variable*, Springer Series in Computational Physics, Springer-Verlag, Berlin, 1991.

[26] E. B. Saff and V. Totik, *Logarithmic Potentials with External Fields*, Grundlehren der Mathematischen Wissenschaften, vol. 316, Springer-Verlag, Berlin, 1997.

[27]   G. Szegő, *Orthogonal Polynomials*, 4th ed., American Mathematical Society, Colloquium Publi-
          cations, vol. 23, American Mathematical Society, Rhode Island, 1975.

[28]   C. A. Tracy and H. Widom, *On orthogonal and symplectic matrix ensembles*, Comm. Math.
          Phys. **177** (1996), no. 3, 727–754.

[29]   ———, *Correlation functions, cluster functions, and spacing distributions for random ma-
          trices*, J. Statist. Phys. **92** (1998), no. 5-6, 809–835.

J. Baik: Department of Mathematics, Princeton University, Princeton, NJ 08544, USA; Department of
Mathematics, University of Michigan, Ann Arbor, MI 48109, USA
E-mail address: jbaik@math.princeton.edu

T. Kriecherbauer: Fakultät für Mathematik, Ruhr-Universität Bochum, D-44780 Bochum, Germany
E-mail address: thomas.kriecherbauer@ruhr-uni-bochum.de

K. T.-R. McLaughlin: Department of Mathematics, University of North Carolina at Chapel Hill, Chapel
Hill, NC 27599-3250, USA; Department of Mathematics, The University of Arizona, Tucson, AZ 85721,
USA
E-mail address: mcl@amath.unc.edu

P. D. Miller: Department of Mathematics, University of Michigan, Ann Arbor, MI 48109, USA
E-mail address: millerpd@umich.edu