

Large deviations for the log-Gamma polymer

Tom Claeys and Julian Mauersberger

October 24, 2024

Abstract

We conjecture an explicit expression for the lower tail large deviation rate function of the partition function of the log-Gamma polymer. We rigorously prove our result, except for one step for which we only provide heuristic evidence. We show that the large deviation rate function matches with that of last passage percolation with exponential weights in the zero-temperature limit, and with the lower tail of the Tracy-Widom distribution for moderate deviations.

1 Introduction and main results

The log-Gamma polymer is a model of directed lattice paths in a random environment, introduced by Seppäläinen [41]. The random environment is special in the sense that it allows exact expressions for several statistics of interest. In particular, the Laplace transform of the partition function admits an exact characterization in terms of a Fredholm determinant for finite lattice size. This characterization enables one to study the asymptotic behavior of the partition function in the limit of a large lattice. In particular, it was shown in [2, 7] that the partition function has Tracy-Widom fluctuations. Other polymer models with similar features are the O’Connell-Yor polymer, the mixed polymer, and the continuum directed polymer or Kardar-Parisi-Zhang equation, as well as their zero-temperature counterparts which are last passage percolation or corner growth models. While a Fredholm determinant identity for the Laplace transform is effective to study the typical large n behavior, it is much more challenging to extract precise information about the lower tails of the probability distribution and to obtain accurate large deviation estimates. Such estimates have been recently established for the narrow wedge solution of the Kardar-Parisi-Zhang equation [10, 12, 16, 39, 44], but remain to be understood for other exactly solvable polymer models.

In this paper, we conjecture an explicit expression for the large deviation rate function of the log-Gamma polymer partition function. Our approach is based on a connection between the log-Gamma polymer partition function and a signed biorthogonal measure found in [11], which leads to a Fredholm determinant which we analyze asymptotically. Our analysis is rigorous, except for one step, namely an estimate for the difference between two Fredholm determinants. We provide heuristic evidence for this estimate, and explain the obstacles for a rigorous proof. We believe that a similar method can lead to the large deviation rate functions in other exactly solvable polymer models, like the inhomogeneous log-Gamma polymer on a rectangular lattice, the O’Connell-Yor polymer [40], and the mixed polymer [8].

Definition of the model. Consider the square lattice $\{1, 2, \dots, n\} \times \{1, 2, \dots, n\}$. To each point (i, j) in the grid we assign an independent positive random variable $d_{i,j}(\theta)$, which is *inverse-Gamma distributed*, that is, it has probability density

$$\frac{x^{-2\theta-1}}{\Gamma(2\theta)} e^{-1/x}, \quad x > 0,$$

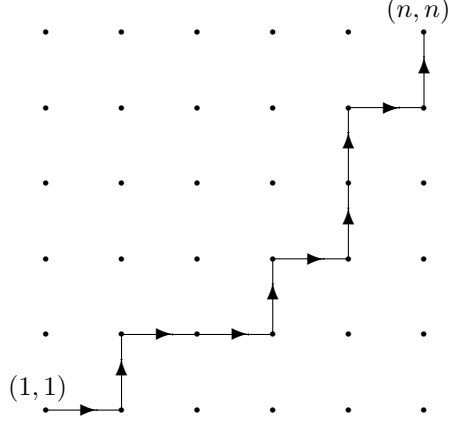


Figure 1: The log-Gamma polymer lattice for $n = 6$, with a possible up-right path.

depending on a parameter $\theta > 0$.

Given an up-right path π connecting $(1, 1)$ with (n, n) , we define its weight as

$$w_{n,\theta}(\pi) = \prod_{(i,j) \in \pi} d_{i,j}(\theta).$$

The *log-Gamma polymer partition function* is defined as the sum of the weights of all such paths:

$$Z_n(\theta) = \sum_{\pi: (1,1) \rightarrow (n,n)} w_{n,\theta}(\pi) = \sum_{\pi: (1,1) \rightarrow (n,n)} \prod_{(i,j) \in \pi} d_{i,j}(\theta), \quad (1.1)$$

where the sum is taken over all up-right directed paths π in the grid $\{1, \dots, n\}^2$ starting at $(1, 1)$ and ending at (n, n) , as illustrated in Figure 1. The model can be generalized by considering a rectangular instead of a square lattice, and by allowing different parameters $\theta = \theta_{i,j}$ for different vertices (i, j) , but we will restrict ourselves to the square lattice with homogeneous weights here.

For $\theta > 0$ small, the density of the weights $d_{i,j}(\theta)$ has a heavy tail, such that one expects large variations between the weights of different paths, implying that the partition function will be dominated by few paths with large weights for large n . More precisely, as $\theta \rightarrow 0$, the random variables $u_{i,j}(\theta) := 2\theta \log d_{i,j}(\theta)$ converge in distribution to independent exponential random variables $u_{i,j}$, i.e., $\mathbb{P}(u_{i,j} \leq s) = 1 - e^{-u_{i,j}}$. As $\theta \rightarrow 0$ (with n fixed), we then obtain the weak convergence [41]

$$2\theta \log Z_n(\theta) \longrightarrow F_n^{\text{LPP}} := \max_{\pi: (1,1) \rightarrow (n,n)} \sum_{(i,j) \in \pi} u_{i,j}.$$

The right hand side is the maximum additive weight of an up-right path in random exponential environment. This model is known as *last passage percolation* or *corner growth* with exponential weights, see [9, 27, 34]. From the physics point of view, $\theta > 0$ is a parameter proportional to the temperature of the model.

Known asymptotic results. Seppäläinen [41] proved that the average of $\log Z_n(\theta)$ behaves for large n like $-2n\psi(\theta)$, where $\psi = \frac{\Gamma'}{\Gamma}$ is the di-Gamma function, and that its variance is of order $n^{2/3}$. Later, in 2013, Borodin, Corwin, and Remenik [7, Corollary 1.8] showed that the Laplace transform of $Z_n(\theta)$ is equal to a Fredholm determinant, see also [6]: we have for $u \in \mathbb{C}$ with $\Re u > 0$ that

$$\mathbb{E}(e^{-uZ_n(\theta)}) = \det(I + K_n^{u,\theta})_{L^2(\Sigma)}, \quad (1.2)$$

where Σ is a positively oriented circle in the complex plane around $-\theta$ with radius $< \theta$, and $K_n^{u,\theta} : L^2(\Sigma) \rightarrow L^2(\Sigma)$ is the integral kernel operator with kernel

$$K_n^{u,\theta}(v, v') := \frac{1}{(2\pi i)^2} \int_{\ell} dw \frac{\pi u^{w-v}}{\sin \pi(v-w)} \frac{W(w)}{W(v)} \frac{1}{w-v'}, \quad v, v' \in \Sigma,$$

where ℓ is a vertical line at the right of Σ and at the left of θ , and

$$W(z) = W(z; n, \theta) = \frac{\Gamma(\theta - z)^n}{\Gamma(\theta + z)^n}. \quad (1.3)$$

We recall that the Fredholm determinant can be evaluated through the *Fredholm series*

$$\det(I + K)_{L^2(\gamma)} = 1 + \sum_{k=1}^{\infty} \frac{1}{k!} \int_{\gamma^k} \det(K(x_i, x_j))_{i,j=1}^k dx_1 \cdots dx_k. \quad (1.4)$$

The above Fredholm determinant identity combined with a saddle point analysis of the kernel $K_n^{u,\theta}$ allowed the authors to prove that the fluctuations of $\log Z_n(\theta)$ are described by the *Tracy-Widom distribution*. More precisely [7, Theorem 1],

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\frac{\log Z_n(\theta) + 2n\psi(\theta)}{(-\psi''(\theta)n)^{1/3}} \leq r \right) = F_{\text{TW}}(r), \quad r \in \mathbb{R} \quad (1.5)$$

for sufficiently small $\theta > 0$, where $F_{\text{TW}}(r)$ is the $\beta = 2$ Tracy-Widom distribution [43]. This distribution describes among others the largest eigenvalue distribution of large Hermitian random matrices with soft edges, and can be written as an Airy kernel Fredholm determinant,

$$F_{\text{TW}}(r) = \det(1 - K^{\text{Ai}})_{L^2(r, \infty)}, \quad K^{\text{Ai}}(x, x') = \frac{\text{Ai}(x)\text{Ai}'(x') - \text{Ai}(x')\text{Ai}'(x)}{x - x'}.$$

This result was subsequently generalized in [37, 2]. One can see (1.5) as the non-zero temperature analogue of a well-known limit theorem for last passage percolation with exponential weights [34], stating that

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\frac{F_n^{\text{LPP}} - 4n}{2^{4/3}n^{1/3}} \leq r \right) = F_{\text{TW}}(r), \quad r \in \mathbb{R}. \quad (1.6)$$

While (1.5) describes the large n behavior for probabilities of typical events, i.e., fluctuations of $\log Z_n(\theta)$ around $-2n\psi(\theta)$ of order $n^{1/3}$, it does not contain information about probabilities of rare events, like fluctuations of larger order. For instance, it does not contain any information about $\mathbb{P} \left(\frac{\log Z_n(\theta) + 2n\psi(\theta)}{(-\psi''(\theta)n)^{1/3}} \leq r \right)$ as $r \rightarrow \pm\infty$ together with $n \rightarrow \infty$. The upper tail asymptotics as $r \rightarrow +\infty$ are well understood [31], but the lower tail ($r \rightarrow -\infty$) of the probability distribution for $\log Z_n(\theta)$ is not understood and cannot be analyzed with current techniques. Such lower tail probabilities are up to now only understood at zero temperature. Then, remarkably, F_n^{LPP} is for finite n equal in distribution to the largest eigenvalue of an $n \times n$ Laguerre-Wishart random matrix [34, Proposition 1.4]. As a consequence, the distribution of F_n^{LPP} can be expressed as a Hankel determinant which can be analyzed in detail for large n , in order to obtain lower tail asymptotics for $\mathbb{P} \left(\frac{F_n^{\text{LPP}} - 4n}{2^{4/3}n^{1/3}} \leq r \right)$ as $r \rightarrow -\infty$ together with $n \rightarrow \infty$. In particular, we then have the large deviation result

$$\lim_{n \rightarrow \infty} \frac{-1}{n^2} \log \mathbb{P} [F_n^{\text{LPP}}(\theta) \leq 4ns] = -\frac{s^2}{2} - \frac{3}{2} + 2s - \log s, \quad 0 < s < 1, \quad (1.7)$$

see e.g. [20, Formula (162)] or [46]. Observe that this is formally consistent with (1.6): setting $s = 1 + r(2n)^{-2/3}$ with $r \rightarrow -\infty$, we find from (1.7) that $\log \mathbb{P} \left(\frac{F_n^{\text{LPP}} - 4n}{2^{4/3} n^{1/3}} \leq r \right) \sim -\frac{|r|^3}{12}$, which is indeed the leading order of $\log F_{\text{TW}}(r)$ as $r \rightarrow -\infty$, see [20, Theorem 1]. Lower tail deviations in the large temperature limit $\theta \rightarrow \infty$ can in principle be analyzed via the weak noise theory methods developed in [36, 45]. A large deviation principle and a variational problem for the stochastic six-vertex model, which degenerates to the log-Gamma polymer in an appropriate limit, was recently obtained in [18], but without explicit expression for the rate function.

Main result. Before formulating our main result, a conjecture for the large deviation rate function of the log-Gamma polymer partition function, we need to introduce several quantities. First, for $0 < s < -\theta\psi(\theta)$ with $\psi = \Gamma'/\Gamma$ as before, we define $b = b(s, \theta) > 0$ as the unique positive number solving the equation

$$\int_0^1 (\theta\psi(\theta + iu\theta b/2) + \theta\psi(\theta - iu\theta b/2) + 2s) \frac{du}{\pi\sqrt{1-u^2}} = 0. \quad (1.8)$$

Next, we define

$$f(s, \theta) = b^2 \int_0^1 (\theta\psi(\theta + iu\theta b/2) + \theta\psi(\theta - iu\theta b/2) + 2s) \sqrt{1-u^2} \frac{du}{2\pi}, \quad 0 < s < -\theta\psi(\theta) \quad (1.9)$$

and

$$F(s, \theta) = - \int_s^{-\theta\psi(\theta)} f(t; \theta) dt, \quad 0 < s < -\theta\psi(\theta). \quad (1.10)$$

Conjecture 1.1. *There exists $\theta_0 > 0$ such that uniformly for $0 < \theta < \theta_0$ and $\epsilon < s < -\theta\psi(\theta)$ for any $\epsilon > 0$, we have*

$$\lim_{n \rightarrow \infty} \frac{-1}{n^2} \log \mathbb{P} \left[\log Z_n(\theta) \leq \frac{2n}{\theta} s \right] = F(s, \theta). \quad (1.11)$$

Remark 1.2. We will prove in Proposition 3.4 that equation (1.8) has indeed a unique solution for $\theta > 0, 0 < s < -\theta\psi(\theta)$. Moreover, we will show that as $s \rightarrow -\theta\psi(\theta)$ with $\theta > 0$ fixed, we have

$$b(s, \theta) \sim 4 \sqrt{\frac{s + \theta\psi(\theta)}{\theta^3 \psi''(\theta)}}. \quad (1.12)$$

On the other hand, as $\theta \rightarrow 0$ with s fixed, we obtain after a straightforward computation, using the fact that $\psi(z) \sim -1/z$ as $z \rightarrow 0$, that

$$b(s, 0) := \lim_{\theta \rightarrow 0} b(s, \theta) = 2 \sqrt{\frac{1}{s^2} - 1}. \quad (1.13)$$

Remark 1.3. From the $s \rightarrow -\theta\psi(\theta)$ behavior for $b(s, \theta)$, a careful Taylor expansion of $f(s, \theta)$ yields

$$F(s, \theta) \sim \frac{2}{3\theta^3 \psi''(\theta)} (s + \theta\psi(\theta))^3 \quad \text{as } s \rightarrow -\theta\psi(\theta). \quad (1.14)$$

Setting $s = -\theta\psi(\theta) + \frac{\theta}{2}(-\psi''(\theta))^{1/3} r n^{-2/3}$, we see that this is consistent with the cubic lower tail of the Tracy-Widom distribution in (1.5), $\log F_{\text{TW}}(r) \sim -\frac{|r|^3}{12}$ as $r \rightarrow -\infty$.

Remark 1.4. Using the fact that $\psi(z) \sim -1/z$ as $z \rightarrow 0$ and the $\theta \rightarrow 0$ behavior for $b(s, \theta)$, we see that

$$f(s, 0) := \lim_{\theta \rightarrow 0} f(s, \theta) = 2 + \frac{s}{4} b(s, 0)^2 - \sqrt{4 + b(s, 0)^2} = 2 - s - \frac{1}{s}. \quad (1.15)$$

Consequently,

$$F(s, 0) := \lim_{\theta \rightarrow 0} F(s, \theta) = -\frac{s^2}{2} - \frac{3}{2} + 2s - \log s. \quad (1.16)$$

This provides an interesting consistency check: as it must, the above formula is precisely the large deviation rate function for last passage percolation with exponential weights given in (1.7).

Methodology. The starting point of our analysis is an alternative Fredholm determinant representation, equal to but of a different form than (1.2), which was established recently in [11, Corollary 5.2]. For $\theta \in (0, 1)$, this result, with $N = n$, $a_k = 0$, $\alpha_j = 2\theta$, states that

$$\mathbb{E}(e^{-e^{-s} Z_n(\theta)}) = \det(1 - \sigma_s L_n^\theta)_{L^2(\mathbb{R})}, \quad (1.17)$$

where

$$L_n^\theta(x, x') = \frac{1}{(2\pi i)^2} \int_{\mathcal{C}} du \int_{\delta + i\mathbb{R}} dv \frac{\Gamma(2\theta - v)^n \Gamma(u)^n}{\Gamma(2\theta - u)^n \Gamma(v)^n} \frac{e^{-vx + ux'}}{v - u}, \quad (1.18)$$

with $\theta < \delta < \min\{2\theta, 1\}$, $\mathcal{C}$ a positively oriented circle of radius $< \theta$ around 0, and with

$$\sigma_s(x) = \frac{1}{1 + e^{-x+s}}. \quad (1.19)$$

Equivalently,

$$L_n^\theta(x, x') = \frac{1}{(2\pi i)^2} \int_{\Sigma} du \int_{c + i\mathbb{R}} dv \frac{W(v)}{W(u)} \frac{e^{-vx + ux'}}{v - u}, \quad (1.20)$$

with W as in (1.3), $c = \delta - \theta \in (0, \min\{\theta, 1 - \theta\})$, $\Sigma = -\theta + \mathcal{C}$. In (1.17), the right hand side is simply a short-hand notation for the Fredholm series (1.4), here with kernel $K(x, x') = -\sigma_s(x) L_n^\theta(x, x')$, without requiring at this point that K is the kernel of a trace-class operator.

A useful feature of L_n^θ is that it is the correlation kernel of a signed biorthogonal measure $d\mu_n^\theta(x_1, \dots, x_n)$ on $\mathbb{R}^n$ [5], more specifically a signed polynomial ensemble of derivative type. In other words, L_n^θ has the reproducing property

$$\int_{\mathbb{R}} L_n^\theta(x, t) L_n^\theta(t, x') dt = L_n^\theta(x, x'), \quad (1.21)$$

and

$$d\mu_n^\theta(x_1, \dots, x_n) := \frac{1}{n!} \det(L_n^\theta(x_i, x_j))_{i,j=1}^n dx_1 \cdots dx_n \quad (1.22)$$

has the form

$$d\mu_n^\theta(x_1, \dots, x_n) = \frac{1}{Z_n} \left(\prod_{1 \leq j < i \leq n} (x_j - x_i) \right) \det(G^{(j-1)}(x_i; \theta))_{i,j=1}^n dx_1 \cdots dx_n, \quad (1.23)$$

where $G(x; \theta)$ is given by

$$G(x; \theta) = \frac{1}{2\pi i} \int_{c + i\mathbb{R}} \frac{\Gamma(\theta - v)^n}{\Gamma(\theta + v)^n} \frac{e^{-vx}}{(\theta + v)^n} dv, \quad (1.24)$$

see [11, Section 5]. Alternatively, G can be expressed in terms of the Meijer G -function, but we will not use this fact.

As we will point out, this biorthogonal measure lives on scales of order n/θ for large n and small θ , which leads us to defining the re-scaled kernel

$$\widehat{L}_n^\theta(y, y') := \frac{2n}{\theta} L_n^\theta\left(\frac{2n}{\theta}y, \frac{2n}{\theta}y'\right) = \frac{1}{(2\pi i)^2} \int_{\Sigma} du \int_{c+i\mathbb{R}} dv \frac{\Gamma(\theta - \frac{\theta v}{2n})^n \Gamma(\theta + \frac{\theta u}{2n})^n}{\Gamma(\theta + \frac{\theta v}{2n})^n \Gamma(\theta - \frac{\theta u}{2n})^n} \frac{e^{-vy+uy'}}{v-u}, \quad (1.25)$$

where Σ is now a circle of radius $< 2n$ around $-2n$, and $0 < c < \min\{2n, 2n\frac{1-\theta}{\theta}\}$. Then, the limit $\theta \rightarrow 0$ makes sense, and

$$\widehat{L}_n^0(y, y') = \frac{1}{(2\pi i)^2} \int_{\Sigma} du \int_{i\mathbb{R}} dv \frac{(1 + \frac{v}{2n})^n (1 - \frac{u}{2n})^n}{(1 - \frac{v}{2n})^n (1 + \frac{u}{2n})^n} \frac{e^{-vy+uy'}}{v-u}$$

is a well-known double integral expression for the kernel for the Laguerre-Wishart random matrix ensemble (see, e.g., [26, Formula (4.2)] or [30, Formula (5.158)]), which satisfies the Marchenko-Pastur law on $[0, 1]$:

$$\lim_{n \rightarrow \infty} \frac{1}{n} \widehat{L}_n^0(y, y) = \frac{2}{\pi} \sqrt{\frac{1-y}{y}} 1_{(0,1)}(y), \quad y \in \mathbb{R}. \quad (1.26)$$

From (1.17), we obtain the identity

$$\mathbb{E}(e^{-e^{-\frac{2n}{\theta}s} Z_n(\theta)}) = \det(1 - \sigma_{s,n,\theta} \widehat{L}_n^\theta)_{L^2(\mathbb{R})}, \quad \sigma_{s,n,\theta}(y) = \frac{1}{1 + e^{-\frac{2n}{\theta}(y-s)}}. \quad (1.27)$$

As $n \rightarrow \infty$, $\sigma_{s,n,\theta}(y)$ converges to the step function $1_{(s,+\infty)}(y)$, such that one may expect that the logarithm of the *step function Fredholm determinant*

$$Q_n^\theta(s) := \det(1 - 1_{(s,+\infty)} \widehat{L}_n^\theta)_{L^2(\mathbb{R})} = \det(1 - \widehat{L}_n^\theta)_{L^2(s,+\infty)} \quad (1.28)$$

is a good approximation of the logarithm of the *smoothed Fredholm determinant*

$$\widetilde{Q}_n^\theta(s) := \det(1 - \sigma_{s,n,\theta} \widehat{L}_n^\theta)_{L^2(\mathbb{R})} = \det(1 - \sigma_{\frac{2ns}{\theta}} L_n^\theta)_{L^2(\mathbb{R})}, \quad (1.29)$$

for large n . We can however not prove this fact without further assumptions, and this is the only reason why Conjecture 1.1 is not a Theorem. It leads us to the following ansatz.

Ansatz 1.5. *There exists $\theta_0 > 0$ such that*

$$\log Q_n^\theta(s) \sim \log \widetilde{Q}_n^\theta(s), \quad \text{as } n \rightarrow \infty, \quad (1.30)$$

uniformly for $\epsilon < s < -\theta\psi(\theta)$ and $0 < \theta < \theta_0$, for any $\epsilon > 0$.

We will explain in Section 8 why we are confident that our ansatz holds true. In Section 7, we will prove Conjecture 1.1 under Ansatz 1.5.

The advantage of the step function Fredholm determinant $Q_n^\theta(s)$ from (1.28), compared to (1.29), is that it can be transformed to an *integrable* Fredholm determinant, which one can characterize in terms of a Riemann-Hilbert (RH) problem thanks to a method developed by Its, Izergin, Korepin, and Slavnov [33]. We derive this RH characterization in Section 2.

A major part of this paper is devoted to the large n asymptotic analysis of this RH problem. In Section 3, we initiate the RH analysis. For $s > -\theta\psi(\theta) + \epsilon$ with $\epsilon > 0$, the asymptotic analysis of the RH problem is simple and only requires a suitable deformation of jump contours. For $\epsilon \leq$

$s \leq -\theta\psi(\theta) + \epsilon$ however, we need to introduce a g -function and construct local parametrices. For $s < -\theta\psi(\theta) - \epsilon$, we need to construct a global parametrix and two local Airy parametrices. We do this in Section 4. For $-\theta\psi(\theta) - \epsilon \leq s \leq -\theta\psi(\theta) + \epsilon$, the two local Airy parametrices come together and merge to a single local parametrix that we can build using a model RH problem associated to the Painlevé II equation. This is the content of Section 5. This analysis is in essence a *closing of a gap* transition similar to the one studied by Baik, Deift, and Johansson in their study of the longest increasing subsequence of a random permutation [1], and later in [4, 15] in a random matrix context. The asymptotic analysis of the RH problem yields the following result, which we will prove in Section 6.

Theorem 1.6. *There exists $\theta_0 > 0$ such that*

$$\log Q_n^\theta(s) = -n^2 F(s, \theta) + O(n^{2/3})$$

as $n \rightarrow +\infty$ uniformly for $\epsilon < s < -\theta\psi(\theta)$ and $0 < \theta < \theta_0$, for any $\epsilon > 0$. The function $F(s, \theta)$ is given by (1.9)–(1.10) with $b(s, \theta) > 0$ the unique positive number solving the equation (1.8). Furthermore, the function $F(s, \theta)$ is positive for $0 < s < -\theta\psi(\theta)$ and $0 < \theta < \theta_0$.

Remark 1.7. As we will explain later, we could with some more effort prove the Tracy-Widom convergence

$$\lim_{n \rightarrow \infty} Q_n^\theta \left(-\theta\psi(\theta) + \frac{\theta}{2}(-\psi''(\theta))^{1/3} r n^{-2/3} \right) = F_{\text{TW}}(r), \quad (1.31)$$

uniformly for $r \in (r_0, +\infty)$ for any $r_0 \in \mathbb{R}$. This is consistent with (1.5) and Ansatz 1.5. The error term $O(n^{2/3})$ in Theorem 1.6 is not sharp: we believe that it can be improved to $O(1)$. It is however not clear whether the approximation (1.30) is valid up to that order.

Combining Theorem 1.6 with Ansatz 1.5, we will prove Conjecture 1.1 in Section 7.

2 RH characterization of the step function Fredholm determinant

In this section, we work with a more general class of kernels L_n than the ones arising in the context of the log-Gamma Fredholm determinant. We let

$$L_n(x, x') = \frac{1}{(2\pi i)^2} \int_{\Sigma} du \int_{c+i\mathbb{R}} dv \frac{W_1(v)}{W_2(u)} \frac{e^{-vx+ux'}}{v-u}, \quad c \in \mathbb{R}, \quad (2.1)$$

where $c \in \mathbb{R}$, Σ is a simple closed positively oriented curve in the half plane $\Re z < c$, $W_1 : c+i\mathbb{R} \rightarrow \mathbb{C}$ is smooth, bounded, and not identically 0, and $W_2 : \Sigma \rightarrow \mathbb{C} \setminus \{0\}$ is smooth and bounded.

If $W_1 = W_2 = W$ given by (1.3) and $0 < c < \min\{\theta, 1-\theta\}$, we have that L_n is equal to L_n^θ . If we take $W_1(z) = W_2(z) = \widehat{W}(z)$, with

$$\widehat{W}(z) = \frac{\Gamma\left(\theta\left(1 - \frac{z}{2n}\right)\right)^n}{\Gamma\left(\theta\left(1 + \frac{z}{2n}\right)\right)^n}, \quad (2.2)$$

and $0 < c < \min\{2n, 2n\frac{1-\theta}{\theta}\}$, then L_n is equal to the re-scaled kernel $\widehat{L}_n^\theta$ from (1.25).

The goal of this section is to relate the step function Fredholm determinant $\det(1 - 1_{(-\infty, s)} L_n)_{L^2(\mathbb{R})}$ to a 2×2 matrix-valued RH problem. For our purposes, it would be enough to consider W_1, W_2 given by (1.3) or (2.2). However, the results and proofs in this section do not rely on the form of W_1 and W_2 , and other choices of W_1, W_2 are relevant in other models, like the non-homogeneous

log-Gamma polymer on a rectangular lattice, the O'Connell-Yor polymer, the mixed polymer, and more general biorthogonal ensembles of derivative type. We therefore believe that Proposition 2.1 below, for general W_1, W_2 , is of independent interest, and will turn out useful to study other models. This section follows similar lines as [13, Section 2] and is based on a general method developed in [3] and built on the theory of integrable operators from [33, 21].

The RH problem of interest is the following, and depends on parameters $s \in \mathbb{R}$ as well as on the function W , but we will omit the latter dependence in our notation. We will write $Y(z) = Y(z; s)$ for the solution of this RH problem.

RH problem for Y

(Y1) $Y : \mathbb{C} \setminus ((c + i\mathbb{R}) \cup \Sigma) \rightarrow \mathbb{C}^{2 \times 2}$ is analytic.

(Y2) For $z \in (c + i\mathbb{R}) \cup \Sigma$, Y satisfies the jump condition

$$Y_+(z) = Y_-(z)J(z),$$

where

$$J(z) = \begin{cases} \begin{pmatrix} 1 & 0 \\ e^{-sz}W_1(z) & 1 \end{pmatrix} & \text{for } z \in c + i\mathbb{R}, \\ \begin{pmatrix} 1 & -e^{sz}W_2(z)^{-1} \\ 0 & 1 \end{pmatrix} & \text{for } z \in \Sigma. \end{cases} \quad (2.3)$$

(Y3) There exists a 2×2 matrix $Y_1 = Y_1(s)$ depending on s and also on W_1, W_2 , such that $Y(z) = I + \frac{Y_1}{z} + \mathcal{O}(z^{-2})$ as $z \rightarrow \infty$.

In general, the boundary values $Y_{\pm}$ have to be understood in L^2 -sense and condition (Y3) is valid for $z \rightarrow \infty$ away from $c + i\mathbb{R}$, like in [21]. However, if $W_1(c + iy)$ decays sufficiently fast as $y \rightarrow \pm\infty$, say $W_1(c + iy) = O(|y|^{-1-\epsilon})$ as $y \rightarrow \pm\infty$ for some $\epsilon > 0$, then the boundary values are continuous, and condition (Y3) is uniform for $z \in \mathbb{C} \setminus (c + i\mathbb{R})$. To avoid technical complications, we will henceforth assume that $W_1(c + iy) = O(|y|^{-1-\epsilon})$ as $y \rightarrow \pm\infty$. Note that in our case of interest (2.2), it follows from Stirling's approximation that $W_1(c + iy) = O(|y|^{-c})$ as $y \rightarrow \pm\infty$, which means that we need to take $c > 1 + \epsilon$, in addition to the already standing condition $c < \min\{2n, 2n\frac{1-\theta}{\theta}\}$.

The rest of this section is dedicated to the proof of the following result, connecting the Fredholm determinant $\det(1 - 1_{(s, \infty)}L_n)_{L^2(\mathbb{R})}$ with the solution to the above RH problem.

Proposition 2.1 (Differential identity). *The RH problem for Y is solvable if and only if $\det(1 - L_n)_{L^2(s, +\infty)} \neq 0$, and we have the identity*

$$\frac{d}{ds} \log \det(1 - L_n)_{L^2(s, +\infty)} = (Y_1(s))_{11}.$$

In particular, $Q_n^\theta(s)$ from (1.28) satisfies

$$\frac{d}{ds} \log Q_n^\theta(s) = (Y_1(s))_{11}, \quad (2.4)$$

where $Y_1(s)$ corresponds to the RH solution Y with $W_1(z) = W_2(z) = \widehat{W}(z)$ given by (2.2).

To prove this result, we will use the two-sided Laplace transform $\mathcal{B} : L^2(\mathbb{R}) \rightarrow L^2(i\mathbb{R})$ defined by

$$\mathcal{B}[f](v) = \int_{-\infty}^{\infty} e^{-vx} f(x) dx,$$

with inverse $\mathcal{B}^{-1} : L^2(i\mathbb{R}) \rightarrow L^2(\mathbb{R})$ given by

$$\mathcal{B}^{-1}[\phi](x) = \frac{1}{2\pi i} \int_{i\mathbb{R}} e^{vx} \phi(v) dv,$$

as well as the integral operator $H_n^s : L^2(i\mathbb{R}) \rightarrow L^2(i\mathbb{R})$ with kernel

$$H_n^s(v, v') = \frac{1}{(2\pi i)^2} \int_{\mathcal{C}} du e^{s(u-v)} \frac{W_1(c+v')}{W_2(c+u)(v'-u)(v-u)},$$

where $\mathcal{C} = \Sigma - c$. Observe that H_n^s is a trace-class operator, as $H_n^s = J \circ K$ is the composition of the Hilbert-Schmidt operators $J : L^2(\mathcal{C}) \rightarrow L^2(i\mathbb{R})$, $K : L^2(i\mathbb{R}) \rightarrow L^2(\mathcal{C})$ with kernels

$$J(v, u) = \frac{e^{-sv}}{2\pi i W_2(c+u)(v-u)}, \quad K(u, v') = \frac{e^{su} W_1(c+v')}{2\pi i (v'-u)}.$$

Lemma 2.2. *For any $s \in \mathbb{R}$, we have the operator identity*

$$\mathcal{B}^{-1} \circ H_n^s \circ \mathcal{B} = \mathcal{L}_n^s,$$

where $\mathcal{L}_n^s : L^2(\mathbb{R}) \rightarrow L^2(\mathbb{R})$ is defined by

$$(\mathcal{L}_n^s)[f](x) = 1_{(s, +\infty)}(x) e^{-cx} \int_{\mathbb{R}} L_n(x', x) e^{cx'} f(x') dx',$$

and the Fredholm determinant identity

$$\det(1 - L_n)_{L^2(s, +\infty)} = \det(I - H_n^s)_{L^2(i\mathbb{R})}.$$

Proof. Let us first compute the kernel of the operator $\mathcal{B}^{-1} \circ J : L^2(\mathcal{C}) \rightarrow L^2(\mathbb{R})$: we have

$$\begin{aligned} (\mathcal{B}^{-1} \circ J)(x, u) &= \frac{1}{(2\pi i)^2 W_2(u+c)} \int_{i\mathbb{R}} \frac{e^{(x-s)v}}{v-u} dv \\ &= \frac{e^{(x-s)u}}{2\pi i W_2(u+c)} 1_{(s, +\infty)}(x), \end{aligned}$$

where we used a contour deformation and residue argument for the last identity. Next, we compute the kernel of the operator $K \circ \mathcal{B} : L^2(\mathbb{R}) \rightarrow L^2(\mathcal{C})$: we have

$$(K \circ \mathcal{B})(u, x) = \frac{e^{su}}{2\pi i} \int_{i\mathbb{R}} \frac{e^{-xv} W_1(v+c)}{v-u} dv.$$

Combining these expressions, we find that the kernel of $\mathcal{B}^{-1} \circ H_n^s \circ \mathcal{B} = \mathcal{B}^{-1} \circ J \circ K \circ \mathcal{B}$ is given by

$$\begin{aligned} (\mathcal{B}^{-1} \circ H_n^s \circ \mathcal{B})(x, x') &= 1_{(s, +\infty)}(x) \frac{1}{(2\pi i)^2} \int_{\mathcal{C}} \frac{e^{xu}}{W_2(u+c)} \int_{i\mathbb{R}} \frac{e^{-x'v} W_1(v+c)}{v-u} dv du \\ &= 1_{(s, +\infty)}(x) \frac{e^{-c(x-x')}}{(2\pi i)^2} \int_{\Sigma} du \int_{c+i\mathbb{R}} dv \frac{W_1(v)}{W_2(u)} \frac{e^{xu-x'v}}{v-u} dv du \\ &= 1_{(s, +\infty)}(x) e^{-c(x-x')} L_n(x', x). \end{aligned}$$

It follows that the kernel $1_{(s,+\infty)}(x)e^{-c(x-x')}L_n(x',x)$ defines a trace-class operator $\mathcal{L}_n^s$ (with $W_1 = W_2 = \widehat{W}$ as in (2.2), it is even of finite rank n , see [11], but we don't need this). We thus have the Fredholm determinant identities

$$\begin{aligned}\det(I - H_n^s)_{L^2(i\mathbb{R})} &= \det(I - \mathcal{B}^{-1} \circ H_n^s \circ \mathcal{B})_{L^2(\mathbb{R})} \\ &= \det(1 - \mathcal{L}_n^s)_{L^2(\mathbb{R})}.\end{aligned}$$

By (1.4), it is straightforward to check that the latter is precisely the Fredholm series $\det(1 - L_n)_{L^2(s,+\infty)}$, and the result is proven. $\square$

Next, we can write the Fredholm determinant in a different way, in terms of a Fredholm determinant of an integrable kernel operator.

Lemma 2.3. *Let $s \in \mathbb{R}$. Then*

$$\det(I - H_n^s)_{L^2(i\mathbb{R})} = \det(I - M_n^s)_{L^2((c+i\mathbb{R}) \cup \Sigma)},$$

where $M_n^s : L^2((c+i\mathbb{R}) \cup \Sigma) \rightarrow L^2((c+i\mathbb{R}) \cup \Sigma)$ is the integral operator induced by the kernel

$$M_n^s(z, z') = \frac{f(z)^t g(z')}{z - z'} \quad (2.5)$$

with

$$f(z) = \frac{1}{2\pi i} \begin{pmatrix} 1_\Sigma(z) W_2(z)^{-1} \\ 1_{c+i\mathbb{R}}(z) W_1(z) \end{pmatrix}, \quad g(z) = \begin{pmatrix} -e^{-sz} 1_{c+i\mathbb{R}}(z) \\ e^{sz} 1_\Sigma(z) \end{pmatrix}. \quad (2.6)$$

Proof. We have

$$\tilde{H}_n^s(v, v') := H_n^s(v - c, v' - c) = \frac{1}{(2\pi i)^2} \int_\Sigma du e^{s(u-v)} \frac{W_1(v')}{W_2(u)(v' - u)(v - u)}.$$

Hence, similarly to the factorization $H_n^s = J \circ K$ used before, we have

$$\tilde{H}_n^s = A \circ B,$$

where $A : L^2(\Sigma) \rightarrow L^2(c+i\mathbb{R})$ and $B : L^2(c+i\mathbb{R}) \rightarrow L^2(\Sigma)$ are the integral operators induced by the kernels

$$A(v, u) = \frac{e^{-sv}}{2\pi i W_2(u)(v - u)}, \quad B(u, v') = \frac{e^{su} W_1(v')}{2\pi i (v' - u)}.$$

Then A and B are Hilbert-Schmidt operators since $\int_\Sigma \int_{c+i\mathbb{R}} |A(v, u)|^2 |dv| |du| < \infty$ and $\int_{c+i\mathbb{R}} \int_\Sigma |B(u, v')|^2 |du| |dv'| < \infty$. We can decompose further $B = B_2 \circ B_1$ and $A = A_2 \circ A_1$, where the operators B_i, A_i are given by

$$\begin{aligned}B_1 : L^2(c+i\mathbb{R}) &\rightarrow L^2(\gamma), & B_1(w, v') &= \frac{W_1(v')}{2\pi i (v' - w)}, \\ B_2 : L^2(\gamma) &\rightarrow L^2(\Sigma), & B_2(u, w) &= \frac{e^{su}}{2\pi i (w - u)}, \\ A_1 : L^2(\Sigma) &\rightarrow L^2(\gamma), & A_1(w, u) &= \frac{1}{2\pi i W_2(u)(w - u)}, \\ A_2 : L^2(\gamma) &\rightarrow L^2(c+i\mathbb{R}), & A_2(v, w) &= \frac{e^{-sv}}{2\pi i (v - w)}.\end{aligned}$$

Here γ is a closed contour surrounding Σ in the half plane $\Re z < c$. All the operators above are Hilbert-Schmidt, hence A and B are trace-class. Considering M_n^s as an operator acting on the space $L^2(\Sigma) \oplus L^2(c + i\mathbb{R})$ and writing it as a 2×2 matrix of operators

$$M_n^s = \begin{pmatrix} 0 & A \\ B & 0 \end{pmatrix},$$

we can compute:

$$\begin{aligned} \det(1 - H_n^s)_{L^2(c+i\mathbb{R})} &= \det \left(1 - \begin{pmatrix} A \circ B & 0 \\ 0 & 0 \end{pmatrix} \right)_{L^2((c+i\mathbb{R}) \cup \Sigma)} \\ &= \det \left(1 - \begin{pmatrix} A \circ B & 0 \\ B & 0 \end{pmatrix} \right)_{L^2((c+i\mathbb{R}) \cup \Sigma)} \\ &= \det \left(1 + \begin{pmatrix} 0 & A \\ 0 & 0 \end{pmatrix} \right)_{L^2((c+i\mathbb{R}) \cup \Sigma)} \det \left(1 - \begin{pmatrix} 0 & A \\ B & 0 \end{pmatrix} \right)_{L^2((c+i\mathbb{R}) \cup \Sigma)} \\ &= \det \left(1 - \begin{pmatrix} 0 & A \\ B & 0 \end{pmatrix} \right)_{L^2((c+i\mathbb{R}) \cup \Sigma)} = \det(1 - M_n^s)_{L^2((c+i\mathbb{R}) \cup \Sigma)}, \end{aligned}$$

which proves the lemma. $\square$

Following the terminology of [33], an integral operator with kernel of the form (2.5) with $\mathbf{f}(u)^t \mathbf{g}(u) = 0$ is called integrable. The general theory behind such operators developed in [33, 21] and [3] enables us to relate the Fredholm determinant $\det(I - M_{n,s})_{L^2(i\mathbb{R} \cup \Sigma)}$ to the solution of a RH problem with jump matrix $I - 2\pi i \mathbf{f}(z) \mathbf{g}(z)^t$. In our case of interest, the relevant RH problem is given by conditions (Y1)–(Y3). Here and below, we write $\sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

Lemma 2.4. *Let $s \in \mathbb{R}$. It holds that*

$$\frac{d}{ds} \log \det(I - M_n^s)_{L^2(i\mathbb{R} \cup \Sigma)} = \int_{(c+i\mathbb{R}) \cup \Sigma} \frac{z}{2} \operatorname{Tr} (Y_+(z)^{-1} Y'_+(z) \sigma_3 - Y_-(z)^{-1} Y'_-(z) \sigma_3) \frac{dz}{2\pi i}.$$

Proof. By [3, Theorem 2.1], we have

$$\frac{d}{ds} \log \det(I - M_n^s)_{L^2((c+i\mathbb{R}) \cup \Sigma)} = \int_{(c+i\mathbb{R}) \cup \Sigma} z \operatorname{Tr} (Y_-(z)^{-1} Y'_-(z) \partial_s J(z) J(z)^{-1}) \frac{dz}{2\pi i}.$$

Using that $(I - J)J^{-1} = I - J$ and abbreviating $J = J(z)$, we get

$$\partial_s J(z) J(z)^{-1} = \begin{cases} -z \begin{pmatrix} 0 & 0 \\ e^{-sz} W_1(z) & 0 \end{pmatrix} J^{-1} = z(I - J)J^{-1} = z(I - J) = z(J - I)\sigma_3, & \text{for } z \in c + i\mathbb{R}, \\ z \begin{pmatrix} 0 & -e^{sz} W_2(z)^{-1} \\ 0 & 0 \end{pmatrix} J^{-1} = z(J - I)J^{-1} = z(J - I) = z(J - I)\sigma_3, & \text{for } z \in \Sigma, \end{cases}$$

which yields

$$\frac{d}{ds} \log \det(I - M_n^s)_{L^2((c+i\mathbb{R}) \cup \Sigma)} = \int_{(c+i\mathbb{R}) \cup \Sigma} z \operatorname{Tr} (Y_-(z)^{-1} Y'_-(z) (J - I)\sigma_3) \frac{dz}{2\pi i}.$$

Furthermore, we have

$$\begin{aligned} Y_+(z)^{-1} Y'_+(z) &= J(z)^{-1} Y_-(z)^{-1} (Y'_-(z) J(z) + Y_-(z) J'(z)) \\ &= J(z)^{-1} Y_-(z)^{-1} Y'_-(z) J(z) + J(z)^{-1} J'(z). \end{aligned}$$

Since $J'(z)$ is off-diagonal this implies

$$\mathrm{Tr}(Y_+(z)^{-1}Y'_+(z)\sigma_3) = \mathrm{Tr}(Y_-(z)^{-1}Y'_-(z)J(z)\sigma_3J(z)^{-1}).$$

Using $J^{-1}(z) = 2I - J(z)$ and $J(z)\sigma_3 = \sigma_3J(z)^{-1}$, we obtain

$$\begin{aligned} \mathrm{Tr}(Y_+(z)^{-1}Y'_+(z)\sigma_3) &= \mathrm{Tr}(Y_-(z)^{-1}Y'_-(z)J(z)\sigma_3(2I - J(z))) \\ &= 2\mathrm{Tr}(Y_-(z)^{-1}Y'_-(z)J(z)\sigma_3) - \mathrm{Tr}(Y_-(z)^{-1}Y'_-(z)J(z)\sigma_3J(z)) \\ &= 2\mathrm{Tr}(Y_-(z)^{-1}Y'_-(z)J(z)\sigma_3) - \mathrm{Tr}(Y_-(z)^{-1}Y'_-(z)\sigma_3). \end{aligned}$$

Hence

$$\begin{aligned} &\mathrm{Tr}(Y_+(z)^{-1}Y'_+(z)\sigma_3 - Y_-(z)^{-1}Y'_-(z)\sigma_3) \\ &= 2\mathrm{Tr}(Y_-(z)^{-1}Y'_-(z)J(z)\sigma_3) - 2\mathrm{Tr}(Y_-(z)^{-1}Y'_-(z)\sigma_3) \\ &= 2\mathrm{Tr}(Y_-(z)^{-1}Y'_-(z)(J - I)\sigma_3), \end{aligned}$$

which implies the result. $\square$

We are now ready to prove Proposition 2.1. By Lemma 2.4 we can write

$$\begin{aligned} \frac{d}{ds} \log \det(I - M_n^s)_{L^2((c+i\mathbb{R}) \cup \Sigma)} &= \int_{c+i\mathbb{R}} \frac{z}{2} \mathrm{Tr}(Y_+(z)^{-1}Y'_+(z)\sigma_3) \frac{dz}{2\pi i} - \int_{c+i\mathbb{R}} \frac{z}{2} \mathrm{Tr}(Y_-(z)^{-1}Y'_-(z)\sigma_3) \frac{dz}{2\pi i} \\ &\quad + \int_{\Sigma} \frac{z}{2} \mathrm{Tr}(Y_+(z)^{-1}Y'_+(z)\sigma_3) \frac{dz}{2\pi i} - \int_{\Sigma} \frac{z}{2} \mathrm{Tr}(Y_-(z)^{-1}Y'_-(z)\sigma_3) \frac{dz}{2\pi i}. \end{aligned}$$

The integral

$$\int_{\Sigma} \frac{z}{2} \mathrm{Tr}(Y_+(z)^{-1}Y'_+(z)\sigma_3) \frac{dz}{2\pi i}$$

vanishes since the integrand can be continued analytically inside Σ . The contour of the integral

$$- \int_{\Sigma} \frac{z}{2} \mathrm{Tr}(Y_-(z)^{-1}Y'_-(z)\sigma_3) \frac{dz}{2\pi i}$$

can be enlarged because the integrand can be continued analytically to the region outside Σ in the half plane $\Re z < c$. For instance, for $R > 0$ large and $\epsilon > 0$ small, we can write

$$- \int_{\Sigma} \frac{z}{2} \mathrm{Tr}(Y_-(z)^{-1}Y'_-(z)\sigma_3) \frac{dz}{2\pi i} = - \int_{\Sigma_{\epsilon,R}} \frac{z}{2} \mathrm{Tr}(Y(z)^{-1}Y'(z)\sigma_3) \frac{dz}{2\pi i},$$

where $\Sigma_{\epsilon,R}$ consists of the semi-circle of radius R around $c - \epsilon$ lying on the left of $c - \epsilon + i\mathbb{R}$ and of the interval $[c - \epsilon - iR, c - \epsilon + iR]$. As $\epsilon \rightarrow 0$ and $R \rightarrow \infty$, the integral on $[c - \epsilon - iR, c - \epsilon + iR]$ cancels out against

$$\int_{c+i\mathbb{R}} \frac{z}{2} \mathrm{Tr}(Y_+(z)^{-1}Y'_+(z)\sigma_3) \frac{dz}{2\pi i}.$$

Moreover, we can deform the contour $[c - iR, c + iR]$ in

$$\int_{c-iR}^{c+iR} \frac{z}{2} \mathrm{Tr}(Y_-(z)^{-1}Y'_-(z)\sigma_3) \frac{dz}{2\pi i}$$

by analytic continuation of the integrand to the semi-circle of radius R around c on the right of c . Hence

$$\frac{d}{ds} \log \det(I - M_n^s)_{L^2((c+i\mathbb{R}) \cup \Sigma)} = \lim_{R \rightarrow \infty} \int_{\{|z-c|=R\}} \frac{z}{2} \operatorname{Tr} (Y(z)^{-1} Y'(z) \sigma_3) \frac{dz}{2\pi i},$$

where the circle $\{|z-c|=R\}$ is oriented clockwise. Using that $Y(z) = I + \frac{Y_1(s)}{z} + \mathcal{O}(z^{-2})$ we arrive at

$$\frac{d}{ds} \log \det(I - M_n^s)_{L^2((c+i\mathbb{R}) \cup \Sigma)} = \int_{\{|z-c|=R\}} -\frac{Y_1(s) \sigma_3}{2z} \frac{dz}{2\pi i} = (Y_1(s))_{11},$$

where we used that $Y_1(s)$ has zero trace because $\det Y = 1$.

We now obtain Proposition 2.1 by combining this with Lemma 2.2 and Lemma 2.3.

3 Steepest descent analysis

In this section we carry out the first steps in the nonlinear steepest descent analysis for the RH problem (Y1)–(Y3) in the case of the log-Gamma polymer. This means that W_1 and W_2 in conditions (Y1)–(Y3) are given by

$$W_1(z) = W_2(z) = \widehat{W}(z) = \frac{\Gamma(\theta(1 - \frac{z}{2n}))^n}{\Gamma(\theta(1 + \frac{z}{2n}))^n}. \quad (3.1)$$

The associated kernel is given by

$$\widehat{L}_n^\theta(x, x') = \frac{1}{(2\pi i)^2} \int_\Sigma du \int_{c+i\mathbb{R}} dv \frac{\widehat{W}(v)}{\widehat{W}(u)} \frac{e^{-xv+x'u}}{v-u}. \quad (3.2)$$

As explained before, we need $1+\epsilon < c < \min\{2n, 2n\frac{1-\theta}{\theta}\}$, and Σ lies in the left half plane and encircles $-2n$ in the positive direction, but none of the other zeros of $\widehat{W}$, namely $-2n - \frac{2n}{\theta}, -2n - \frac{4n}{\theta} \dots$. A convenient choice for c will turn out to be $c = \delta n$ with $\delta > 0$ sufficiently small.

Following the Deift-Zhou nonlinear steepest descent method [24, 19], we will apply a series of transformations

$$Y \mapsto U \mapsto T \mapsto S \mapsto R$$

to the RH problem for Y , with the aim of arriving at a RH problem with small jumps as $n \rightarrow \infty$ and normalized to I at infinity. All of these RH solutions Y, U, T, S, R will depend on a complex variable denoted z or ζ , as well as on parameters s, θ, n . For the ease of notation, the dependence on the parameters s, θ, n will however not always be visible in our notations.

3.1 Transformation $Y \mapsto U$

Let us first introduce a rotation and dilation of the variable z of the RH problem, by setting $z = in\zeta$. This is convenient for two reasons: first, the scale of the new variable ζ will be convenient for the construction of a g -function, and secondly, we prefer to have the jump matrix on a horizontal line instead of a vertical one. We define

$$U(\zeta) = Y(in\zeta).$$

The RH conditions (Y1)–(Y3) for Y then translate to the following RH conditions for U .

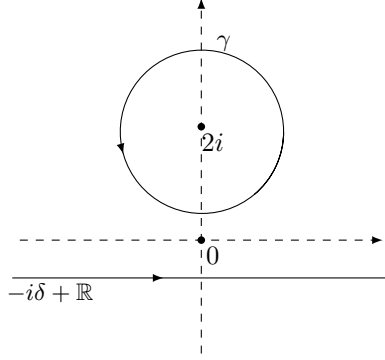


Figure 2: The shape of the jump contour $(-i\delta + \mathbb{R}) \cup \gamma$ for the RH problem for U .

RH problem for U

- (U1) $U : \mathbb{C} \setminus ((-i\delta + \mathbb{R}) \cup \gamma) \rightarrow \mathbb{C}^{2 \times 2}$ is analytic, where $\gamma = -\frac{i}{n}\Sigma$ is a closed positively oriented curve in the upper half plane surrounding $2i$ but not surrounding $2i + \frac{2i}{\theta}, 2i + \frac{4i}{\theta}, \dots$
- (U2) For $\zeta \in (-i\delta + \mathbb{R}) \cup \gamma$, U satisfies the jump condition

$$U_+(\zeta) = U_-(\zeta)J_U(\zeta),$$

where

$$J_U(\zeta) = \begin{cases} \begin{pmatrix} 1 & 0 \\ e^{nh(\zeta; s, \theta)} & 1 \end{pmatrix} & \text{for } \zeta \in -i\delta + \mathbb{R}, \\ \begin{pmatrix} 1 & -e^{-nh(\zeta; s, \theta)} \\ 0 & 1 \end{pmatrix} & \text{for } \zeta \in \gamma, \end{cases} \quad (3.3)$$

and

$$h(\zeta; s, \theta) = \log \Gamma \left(\theta \left(1 - \frac{i\zeta}{2} \right) \right) - \log \Gamma \left(\theta \left(1 + \frac{i\zeta}{2} \right) \right) - is\zeta. \quad (3.4)$$

- (U3) As $\zeta \rightarrow \infty$, we have $U(\zeta) = I - i \frac{Y_1(s, n, \theta)}{n\zeta} + \mathcal{O}(\zeta^{-2})$, where we write $Y_1(s, n, \theta)$ instead of $Y_1(s)$ to emphasize the dependence on n and θ .

Observe that $e^{\pm nh(\zeta; s, \theta)}$ is analytic for $\zeta \in \mathbb{C} \setminus \{\mp 2i, \mp 2i \mp \frac{2i}{\theta}, \mp 2i \mp \frac{4i}{\theta}, \dots\}$, and that

$$h'(\zeta; s, \theta) = -\frac{i\theta}{2}\psi \left(\theta \left(1 - \frac{i\zeta}{2} \right) \right) - \frac{i\theta}{2}\psi \left(\theta \left(1 + \frac{i\zeta}{2} \right) \right) - is, \quad (3.5)$$

and

$$h''(\zeta; s, \theta) = -\frac{\theta^2}{4}\psi' \left(\theta \left(1 - \frac{i\zeta}{2} \right) \right) + \frac{\theta^2}{4}\psi' \left(\theta \left(1 + \frac{i\zeta}{2} \right) \right). \quad (3.6)$$

In particular, $h''(\zeta; s, \theta)$ is antisymmetric in ζ , and for $\zeta > 0$, we have that

$$ih''(\zeta; s, \theta) = -\frac{\theta^2}{2}\Im \psi' \left(\theta \left(1 + \frac{i\zeta}{2} \right) \right) > 0. \quad (3.7)$$

The positivity follows for instance from the classical identity $\psi'(z) = \sum_{k=0}^{\infty} \frac{1}{(z+k)^2}$ for the trigamma function.

As $\theta \rightarrow 0$, we have

$$h(\zeta; s, 0) := \lim_{\theta \rightarrow 0} h(\zeta; s, \theta) = \log \frac{\zeta - 2i}{\zeta + 2i} - is\zeta. \quad (3.8)$$

The sign of $\Re h(\zeta; s, 0)$ as a function of $\zeta \in \mathbb{C}$ will be crucial in what follows. We denote $\mathcal{H}_+^s$ for the subset of the upper half plane containing all ζ such that $\Re h(\zeta; s, 0) > 0$, and $\mathcal{H}_-^s = \overline{\mathcal{H}_+^s} = -\mathcal{H}_+^s$ for the subset of the lower half plane on which $\Re h(\zeta; s, 0) < 0$. We illustrate the sign distribution of $\Re h(\zeta; s, 0)$ for $s < 1$, for $s = 1$, and for $s > 1$ in Figure 3.

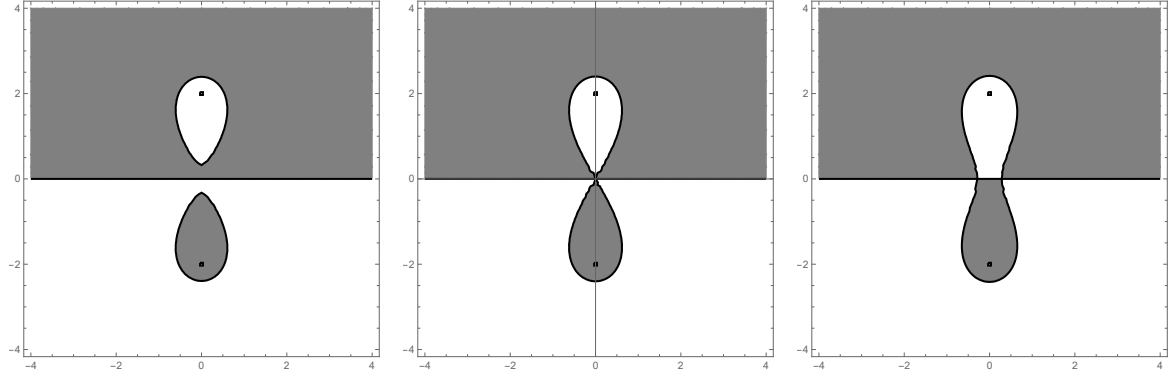


Figure 3: Plot of regions in the complex ζ -plane where $\Re h(\zeta; s, 0)$ is positive (shaded) and negative (white). The left figure corresponds to $s > 1$, the middle figure to $s = 1$, and the right figure to $s < 1$. $\mathcal{H}_+^s$ is the shaded region in the upper half plane, and $\mathcal{H}_-^s$ is the white region in the lower half plane.

We record the following result about the sets $\mathcal{H}_\pm^s$ for $s = 1$ and $s > 1$ for later use.

Proposition 3.1. *The sets $\mathcal{H}_\pm^s$ are increasing in s : for any $s_1 < s_2$, we have $\mathcal{H}_\pm^{s_1} \subset \mathcal{H}_\pm^{s_2}$. For $s = 1$, there exists a smooth 8-shaped closed curve with self-intersection at 0, such that the exterior of the curve is a subset of $\mathcal{H}_+^1 \cup \mathcal{H}_-^1$, and such that angle between the curve and the real line at 0 is $\pm\pi/3$. For any $s_0 > 1$, there exist two complex conjugate simply closed curves $\tilde{\gamma}_\pm$, with $\tilde{\gamma}_+$ strictly in the upper half plane and $\tilde{\gamma}_-$ strictly in the lower half plane, such that the exterior of $\tilde{\gamma}_+ \cup \tilde{\gamma}_-$ is a subset of $\mathcal{H}_+^s \cup \mathcal{H}_-^s$ for every $s > s_0$.*

Proof. The fact that $\mathcal{H}_\pm^s$ is increasing in s is a simple consequence of (3.8).

For $s = 1$, we compute

$$\Re h(\zeta; 1, 0) = \log \frac{|\zeta - 2i|}{|\zeta + 2i|} + \Im \zeta.$$

As $\zeta \rightarrow \infty$, this behaves like $\Im \zeta + O(\zeta^{-2})$, and it is straightforward to check that there exists $R > 0$ such that $\pm \Re h(\zeta; 1, 0) > 0$ for $|\zeta| \geq R$ and $\pm \Im \zeta > 0$. Next, given $R > 0$, we prove that there exists $t_0 = t_0(R) > 0$ such that $\pm \Re h(\zeta; 1, 0) > 0$ for $|\zeta| < R$, provided that $|\arg \zeta| < t_0$ or $|\arg(-\zeta)| < t_0$, and $\pm \Im \zeta > 0$. For that, we observe that

$$h'(\zeta; 1, 0) = \frac{4i}{\zeta^2 + 4} - i,$$

such that for $u \in \mathbb{R}$, $u \neq 0$,

$$\Re h(u; 1, 0) = 0, \quad \partial_v \Re h(u + iv; 1, 0)|_{v=0} = -\partial_u \Im h(u + iv; 1, 0)|_{v=0} > 0,$$

by the Cauchy-Riemann conditions. Hence, for $t_0 > 0$ sufficiently small, we have the required inequalities. With the values of $R, t_0 > 0$ at hand, we can now choose the eight-shaped in the region $\{|\zeta| \geq R\} \cup \{|\zeta| < R, \text{ and } |\arg \zeta| < t_0 \text{ or } |\arg(-\zeta)| < t_0\}$. From the cubic behavior of $h(\zeta; 1, 0)$ around 0, it is clear that we can take the curve such that it makes an angle $\pm\pi/3$ with the real line.

The case $s > 1$ can be treated similarly, the only difference being that a region containing 0 can be connected with the exterior of a large disk, such that we can choose $\tilde{\gamma}_+$ in a region of the form $\{|\zeta| \geq R\} \cup \{|\zeta| < R \text{ and } 0 < \Im \zeta < t_0\}$. $\square$

3.2 Transformation $U \mapsto T$

The phase diagrams in Figure 3 indicate how we need to proceed with the asymptotic RH analysis for $\theta > 0$ sufficiently small. In an ideal scenario, we would like to have a RH problem with jump matrices close to identity. To achieve this, we need to have the jump contour γ such that it lies in the shaded part $\mathcal{H}_+^s$ in the upper half plane, and we need the jump contour $-i\delta + \mathbb{R}$ to be contained in $\mathcal{H}_-^s$. We are able to deform the contours since the jump matrices $J_U(\zeta)$ are analytic for $\zeta \in \mathbb{C} \setminus \{\mp 2i, \mp 2i \mp \frac{2i}{\theta}, \mp 2i \mp \frac{4i}{\theta}, \dots\}$, but one should pay attention to the fact that we cannot cross singularities while deforming contours.

Let us first consider the case $\theta = 0$. Then, for $s > 1$, by Proposition 3.1, we can indeed take γ in the shaded region $\mathcal{H}_+^s$, and we can deform $\mathbb{R}$ to a curve in the white region $\mathcal{H}_-^s$ in the lower half plane. For $s = 1$, we can keep γ in the shaded region $\mathcal{H}_+^1$ only if we let it go through 0. Similarly, we then deform the horizontal line $-i\delta + \mathbb{R}$ to a contour in the white region $\mathcal{H}_-^1$ in the lower half plane but passing through 0. For $s < 1$, we are not in such an ideal scenario, as we cannot immediately make a convenient choice for the jump contours. We will then be forced to let γ coincide with the real line on an interval $[-b, b]$, before going into the upper half plane. Similarly, we will deform the horizontal line $-i\delta + \mathbb{R}$ to a suitable curve lying strictly in the lower half plane, except on the interval $[-b, b]$.

For $\theta > 0$ sufficiently small, the transition between the two phases will not happen at $s = 1$, but rather at $s = -\theta\psi(\theta)$, which is bigger than 1 for $\theta > 0$ small and which converges to 1 as $\theta \rightarrow 0$. In each of the three cases $s > -\theta\psi(\theta)$, $s = -\theta\psi(\theta)$, $s < -\theta\psi(\theta)$, we define a jump contour Σ as the closure of $\Sigma^+ \cup \Sigma^0 \cup \Sigma^-$, where Σ^+ is the part of Σ lying in the upper half plane, Σ^- the part of Σ lying in the lower half plane, and $\Sigma^0 = (-b, b)$ the (possibly empty) part of Σ lying on the real line. Depending on the value of s , we choose the jump contours as follows, and as illustrated in Figure 4.

- For $s > -\theta\psi(\theta)$: Σ^0 is empty, $\Sigma^+ = \gamma$ is a closed curve in $\mathcal{H}_+^s$, and Σ^- is a curve connecting $e^{i(\pi+\delta)\infty}$ with $e^{-i\delta}\infty$ through $\mathcal{H}_-^s$, with $\delta > 0$ sufficiently small. Note that for any $\epsilon > 0$, we can take $\Sigma^\pm$ independent of s for all $s > -\theta\psi(\theta) + \epsilon$, since the sets $\mathcal{H}_\pm^s$ are increasing in s .
- For $s = -\theta\psi(\theta)$: Σ^0 is empty, Σ^+ is a curve in $\mathcal{H}_+^1 \cup \{0\}$ closing at the point 0, and Σ^- consists of a curve connecting $e^{i(\pi+\delta)\infty}$ with 0 and one connecting 0 with $e^{-i\delta}\infty$ through $\mathcal{H}_-^1 \cup \{0\}$.
- For $0 < s < -\theta\psi(\theta)$: $\Sigma^0 = (-b, b)$ with $b = b(s, \theta) > 0$ to be determined later, Σ^+ is a curve connecting b with $-b$ in the upper half plane, and Σ^- consists of a curve connecting $e^{i(\pi+\delta)\infty}$ with $-b$ and one connecting b with $e^{-i\delta}\infty$ in the lower half plane.

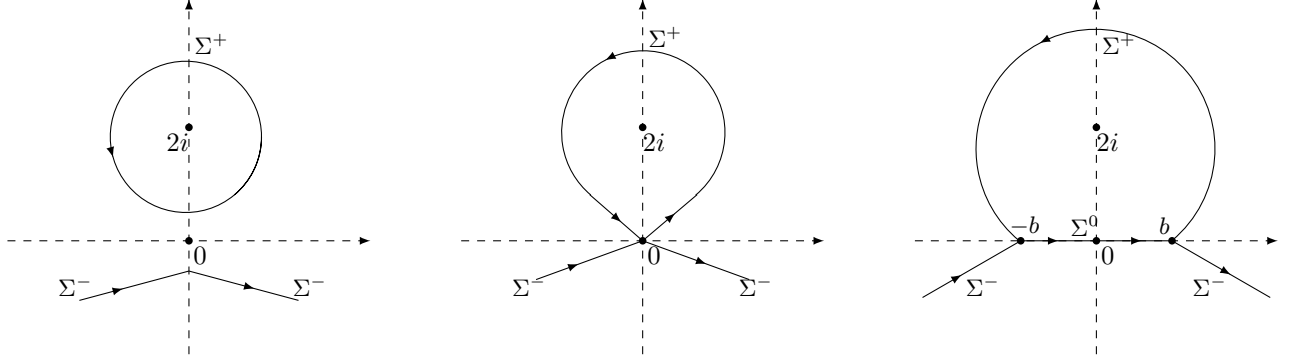


Figure 4: The shape of the jump contour $\Sigma = \Sigma^+ \cup \Sigma^0 \cup \Sigma^-$ for the RH problem for T , in the cases $s > -\theta\psi(\theta)$ (left), $-s = -\theta\psi(\theta)$ (middle), and $0 < s < -\theta\psi(\theta)$ (right).

Let us assume, without loss of generality and for simplicity, that γ is a small contour surrounding $2i$, such that it always lies in the interior of Σ^+ . In all three cases $s > -\theta\psi(\theta)$, $s = -\theta\psi(\theta)$, and $0 < s < -\theta\psi(\theta)$, we define T piecewise as follows:

- in the region below Σ^- , it is the analytic continuation of U from the region below $-i\delta + \mathbb{R}$;
- in the region above Σ^- and outside Σ^+ , it is the analytic continuation of U from the region above $-i\delta + \mathbb{R}$ and outside γ ;
- inside Σ^+ , it is the analytic continuation of U from the region inside γ .

In other words, we define

$$T(\zeta) = \begin{cases} U(\zeta) \begin{pmatrix} 1 & 0 \\ e^{nh(\zeta;s,\theta)} & 1 \end{pmatrix}, & \text{for } \Im \zeta < -\delta \text{ and } \zeta \text{ above } \Sigma^-, \\ U(\zeta) \begin{pmatrix} 1 & -e^{-nh(\zeta;s,\theta)} \\ 0 & 1 \end{pmatrix}, & \text{for } \zeta \text{ in between } \gamma \text{ and } \Sigma^+, \\ U(\zeta) \begin{pmatrix} 1 & 0 \\ -e^{nh(\zeta;s,\theta)} & 1 \end{pmatrix}, & \text{for } \Im \zeta > -\delta \text{ and } \zeta \text{ below } \Sigma^-, \\ U(\zeta), & \text{elsewhere.} \end{cases} \quad (3.9)$$

Then T solves the following RH problem.

RH problem for T

(T1) $T : \mathbb{C} \setminus \Sigma \rightarrow \mathbb{C}^{2 \times 2}$ is analytic.

(T2) Across the contour Σ , T satisfies the jump condition

$$T_+(\zeta) = T_-(\zeta) J_T(\zeta),$$

where

$$J_T(\zeta) = \begin{cases} \begin{pmatrix} 1 & 0 \\ e^{nh(\zeta;s,\theta)} & 1 \end{pmatrix} & \text{for } \zeta \in \Sigma^-, \\ \begin{pmatrix} 1 & -e^{-nh(\zeta;s,\theta)} \\ 0 & 1 \end{pmatrix} & \text{for } \zeta \in \Sigma^+, \\ \begin{pmatrix} 1 & -e^{-nh(\zeta;s,\theta)} \\ e^{nh(\zeta;s,\theta)} & 0 \end{pmatrix} & \text{for } \zeta \in \Sigma^0. \end{cases}$$

$$(T3) \quad T(\zeta) = I - i \frac{Y_1(s,n,\theta)}{n\zeta} + \mathcal{O}(\zeta^{-2}) \text{ as } \zeta \rightarrow \infty.$$

In the case $s > -\theta\psi(\theta) + \epsilon$, as a consequence of Proposition 3.1 and the construction of the jump contours, the jump matrices for T converge to I as $n \rightarrow \infty$ provided that $\theta > 0$ is sufficiently small.

Proposition 3.2. *Let $\epsilon > 0$. There exist $c, \theta_0 > 0$ such that*

$$J_T(\zeta; s, \theta) = I + O(e^{-cns|\zeta|}), \quad n \rightarrow \infty,$$

uniformly for $\zeta \in \Sigma$, $0 \leq \theta < \theta_0$, $s > -\theta\psi(\theta) + \epsilon$.

Proof. Since Σ^0 is empty, it is sufficient to prove that there exists $c > 0$ such that

$$\begin{aligned} \Re h(\zeta; s, \theta) &\geq cs|\zeta|, & \zeta \in \Sigma^+, \\ \Re h(\zeta; s, \theta) &\leq -cs|\zeta|, & \zeta \in \Sigma^-, \end{aligned}$$

for $0 \leq \theta < \theta_0$, $s > -\theta\psi(\theta) + \epsilon$. Using the large ζ behavior of h , we can verify in a straightforward way that the second inequality holds for $|\zeta| > R$, $\zeta \in \Sigma^-$, for some sufficiently large $R > 0$. Next, we consider $|\zeta| \leq R$, and we can assume that this region contains the whole curve Σ^+ . By construction, the above inequalities hold for $\theta = 0$. Moreover, by continuity in $\theta \geq 0$, they continue to hold for θ sufficiently small. $\square$

The above result implies that $\|J_T - I\|_1, \|J_T - I\|_2, \|J_T - I\|_\infty$ are $O(e^{-cns})$ as $n \rightarrow \infty$. It then follows from the general theory of RH problems, see [19, 22, 23], that the RH solution T is also close to identity as $n \rightarrow \infty$.

Corollary 3.3. *Let $\epsilon > 0$. There exist $c, \theta_0 > 0$ such that*

$$T(\zeta) = I + O\left(\frac{e^{-cns}}{|\zeta| + 1}\right), \quad n \rightarrow \infty,$$

uniformly for $\zeta \in \mathbb{C} \setminus \Sigma$, $0 \leq \theta < \theta_0$, $s > -\theta\psi(\theta) + \epsilon$. In particular, $Y_1(s, n, \theta) = O(ne^{-cns})$ as $n \rightarrow \infty$, uniformly for $0 \leq \theta < \theta_0$, $s > -\theta\psi(\theta) + \epsilon$.

This completes the asymptotic RH analysis in the case $s > -\theta\psi(\theta) + \epsilon$.

For s close to $-\theta\psi(\theta)$, the convergence of the jump matrices to I breaks down for ζ near 0, and we will be required to construct a local parametrix in that case.

For $s < -\theta\psi(\theta) - \epsilon$, the jumps are not small in a bigger neighborhood of 0, and we will need to construct a g -function to remedy this, before proceeding with the construction of a global parametrix and local parametrices.

3.3 Construction of the g -function for $0 < s < -\theta\psi(\theta)$

Our next transformation of the RH problem will rely on the construction of a g -function. As usual in asymptotic RH analysis, the g -function is the fundamental object that enables one to arrive, after the construction of parametrices, at a small-norm RH problem.

The g -function $g(\zeta) = g(\zeta; s, \theta)$ will have to satisfy the following conditions, for some $b = b(s, \theta) \geq 0$ that we will fix later on.

RH conditions for g

- (g1) $g : \mathbb{C} \setminus [-b, b] \rightarrow \mathbb{C}$ is analytic,
- (g2) $g_+(\zeta) + g_-(\zeta) = h(\zeta; s, \theta)$ for $\zeta \in (-b, b)$,
- (g3) $g(\zeta) = \frac{g_1}{\zeta} + O(\zeta^{-2})$ as $\zeta \rightarrow \infty$, for some $g_1 = g_1(s, \theta)$.

It is not hard to construct such a function g : we let

$$g(\zeta) = a(\zeta) \int_{-b}^b \frac{h(u; s, \theta)}{a_+(u)(u - \zeta)} \frac{du}{2\pi i}, \quad (3.10)$$

with

$$a(\zeta) = ((\zeta - b)(\zeta + b))^{1/2},$$

where the branch is chosen such that a is analytic in $\mathbb{C} \setminus [-b, b]$ and such that $a(\zeta) \sim \zeta$ as $\zeta \rightarrow \infty$, such that $a_+(u) = i\sqrt{b^2 - u^2}$ for $u \in (-b, b)$. Then, g is analytic except on $[-b, b]$ and, by Cauchy's theorem, g satisfies the jump relation (g2). We also see that

$$\lim_{\zeta \rightarrow \infty} g(\zeta) = \frac{1}{2\pi} \int_{-b}^b \frac{h(u; s, \theta)}{\sqrt{b^2 - u^2}} du = 0,$$

since h is antisymmetric. Moreover, we obtain

$$g(\zeta; s, \theta) = \frac{g_1(s, \theta)}{\zeta} + O(\zeta^{-2})$$

as $\zeta \rightarrow \infty$, with

$$\begin{aligned} g_1(s, \theta) &= \int_{-b}^b \frac{h(u; s, \theta)u}{\sqrt{b^2 - u^2}} \frac{du}{2\pi} \\ &= \int_{-b}^b h'(u; s, \theta) \sqrt{b^2 - u^2} \frac{du}{2\pi} \\ &= \frac{1}{\pi} \int_0^b h'(u; s, \theta) \sqrt{b^2 - u^2} du \\ &= \frac{b^2}{\pi} \int_0^1 h'(bv; s, \theta) \sqrt{1 - v^2} dv = -if(s, \theta), \end{aligned} \quad (3.11)$$

with f as in (1.9).

If we would proceed with arbitrary b , this g -function would not lead to jump matrices close to identity, and the local behavior of $g(\zeta)$ as $\zeta \rightarrow \pm b$ would not be suitable for the construction of local parametrices later on. We anticipate to this problem by choosing the value of $b = b(s, \theta) \geq 0$ appropriately. We define it, as announced in the introduction, to be the unique non-negative solution of the equation (1.8) if $0 < s \leq -\theta\psi(\theta)$. We now prove the existence and uniqueness of $b(s, \theta)$.

Proposition 3.4. *For any $\theta > 0$ and $s \in (0, -\theta\psi(\theta))$, there exists a unique number $b = b(s, \theta) > 0$ such that (1.8) holds. Moreover, $b(s, \theta)$ is differentiable and strictly decreasing in s . As $s \rightarrow -\theta\psi(\theta)$, (1.12) holds.*

Proof. Recall the expressions (3.4)–(3.6) for h and its first two derivatives. Observe that we can re-write (1.8) as

$$H(b) := \int_0^1 \frac{ih'(bu; s, \theta)}{\sqrt{1-u^2}} du = 0. \quad (3.12)$$

We have $H(0) = \frac{\pi}{2}(s + \theta\psi(\theta)) < 0$, and $H(\infty) = +\infty$, since the digamma function behaves like a logarithm at infinity. Moreover, H is strictly increasing in $b > 0$, since

$$H'(b) = \int_0^1 \frac{iuh''(bu)}{\sqrt{1-u^2}} du > 0.$$

We conclude that there is a unique solution $b(s, \theta)$ for any $\theta > 0$, $s \in (0, -\theta\psi(\theta))$. Since $ih'(bu; s, \theta)$ increases with s , $H(b)$ increases with s , such that $b(s, \theta)$ decreases with s . Moreover, by the implicit function theorem, $b(s, \theta)$ is differentiable in $s \in (0, \theta\psi(\theta))$. Taylor expanding H as $b \rightarrow 0$, we obtain

$$\begin{aligned} H(b) &= H(0) + H'(0)b + \frac{H''(0)}{2}b^2 + O(b^3) \\ &= \frac{\pi}{2}(s + \theta\psi(\theta)) + \frac{b^2}{2} \int_0^1 \frac{i u^2 h'''(0)}{\sqrt{1-u^2}} du + O(b^3) \\ &= \frac{\pi}{2}(s + \theta\psi(\theta)) - \frac{\pi\theta^3}{32}\psi''(\theta)b^2 + O(b^3). \end{aligned}$$

It then follows that (1.12) holds. $\square$

3.4 Transformation $T \mapsto S$

We now define S as follows:

$$S(\zeta) = T(\zeta) \begin{pmatrix} e^{-ng(\zeta)} & 0 \\ 0 & e^{ng(\zeta)} \end{pmatrix}, \quad (3.13)$$

where $g(\zeta) = g(\zeta; s, \theta)$ is the g -function which we just constructed if $\epsilon < s < -\theta\psi(\theta)$, and $g(\zeta) = 0$ if $-\theta\psi(\theta) \leq s \leq -\theta\psi(\theta) + \epsilon$. To state the jump relations for S in a simplified manner, it is convenient to introduce the function $q(\zeta) = q(\zeta; s, \theta)$ as

$$q(\zeta) = g(\zeta; s, \theta) - \frac{1}{2}h(\zeta; s, \theta). \quad (3.14)$$

This function is defined for $\zeta \in \mathbb{C} \setminus [-b, b]$ such that $h(\zeta)$ is defined. In particular, it is defined for $\zeta \in \Sigma^+ \cup \Sigma^-$. Using (3.10) and deforming the integration contour $[-b, b]$ to a positively oriented loop $\mathcal{C}$ around $[-b, b]$ and ζ , we obtain the following integral representation for q :

$$q(\zeta) = -a(\zeta) \int_{\mathcal{C}} \frac{h(u; s, \theta)}{a(u)(u - \zeta)} \frac{du}{4\pi i}. \quad (3.15)$$

For $\theta = 0$, we can compute $g'(\zeta; s, 0)$ and $q'(\zeta; s, 0)$ explicitly. We then have

$$g'(\zeta; s, 0) = \frac{h'(\zeta; s, 0)}{2} + \frac{is\zeta a(\zeta)}{2(\zeta^2 + 4)}, \quad (3.16)$$

$$q'(\zeta; s, 0) = \frac{is\zeta a(\zeta)}{2(\zeta^2 + 4)}. \quad (3.17)$$

Using the definitions of S and q , we obtain the following RH conditions for S .

RH problem for S

(S1) $S : \mathbb{C} \setminus \Sigma \rightarrow \mathbb{C}^{2 \times 2}$ is analytic.

(S2) Across the contour Σ , S satisfies the jump condition

$$S_+(\zeta) = S_-(\zeta)J_S(\zeta),$$

where

$$J_S(\zeta) = \begin{cases} \begin{pmatrix} 1 & 0 \\ e^{-2nq(\zeta)} & 1 \end{pmatrix} & \text{for } \zeta \in \Sigma^-, \\ \begin{pmatrix} 1 & -e^{2nq(\zeta)} \\ 0 & 1 \end{pmatrix} & \text{for } \zeta \in \Sigma^+, \\ \begin{pmatrix} e^{-n(q_+(\zeta)-q_-(\zeta))} & -1 \\ 1 & 0 \end{pmatrix} & \text{for } \zeta \in \Sigma^0. \end{cases} \quad (3.18)$$

(S3) $S(\zeta) = I + \frac{-iY_1(s, n, \theta) - n^2 g_1(s, \theta) \sigma_3}{n\zeta} + \mathcal{O}(\zeta^{-2})$ as $\zeta \rightarrow \infty$.

Recall that $\Sigma^0 = (-b, b)$ for $\epsilon \leq s < -\theta\psi(\theta)$, while Σ^0 is empty for $-\theta\psi(\theta) \leq s \leq -\theta\psi(\theta) + \epsilon$.

In analogy with the regions $\mathcal{H}_\pm^s$, we define $\mathcal{G}_+^s$ for the subset of the upper half plane containing all ζ such that $\Re q(\zeta; s, 0) < 0$, and $\mathcal{G}_-^s = \overline{\mathcal{G}_+^s}$ for the subset of the lower half plane on which $\Re q(\zeta; s, 0) > 0$. Let $\epsilon > 0$. We now claim that for any $\epsilon \leq s < -\theta\psi(\theta) - \epsilon$, there exist $\theta_0 > 0$ and curves $\Sigma^\pm$ in $\mathcal{G}_\pm^s$ such that Σ^+ is a curve connecting b with $-b$, and where Σ^- consists of a curve connecting $e^{i(\pi+\delta)}\infty$ with $-b$ and a curve connecting b with $e^{-i\delta}\infty$, for some small $\delta > 0$. Moreover, we can take $\Sigma^\pm$ such that they make an angle $\pm\pi/3$ with the real line at $\pm b$. In other words, we can take $\Sigma^\pm$ as illustrated in Figure 4 on the right.

To prove the claim, we proceed similarly as in the proof of Proposition 3.1. First, from the behavior of g and h as $\zeta \rightarrow \infty$, it follows that $\{\zeta \in \mathbb{C} : |\zeta| \geq R, \pm\Im \zeta > 0\}$ is a subset of $\mathcal{G}_\pm^s$ for sufficiently large $R > 0$. Moreover, from (3.17), it follows that $-iq'(\zeta; s, 0) > 0$ for $\zeta \in \mathbb{R} \setminus [-b, b]$. Using the Cauchy-Riemann equations, we can then again show that small sectors $0 < \arg(\zeta - b) < \delta$ and $\pi - \delta < \arg(\zeta + b) < \pi$ belong to $\mathcal{G}_+^s$. Similarly, small sectors $-\delta < \arg(\zeta - b) < 0$ and $\pi < \arg(\zeta + b) < \pi + \delta$ belong to $\mathcal{G}_-^s$.

By construction, we then obtain the following result.

Proposition 3.5. *Let $\epsilon > 0$. There exist $\eta, c, \theta_0 > 0$ such that*

$$J_S(\zeta) = \begin{cases} I + \mathcal{O}(e^{-cn|\zeta|}), & \zeta \in \Sigma^\pm, |\zeta \pm b| > \eta, \\ \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} + \mathcal{O}(e^{-cn}), & \zeta \in \Sigma^0, |\zeta \pm b| > \eta, \end{cases}$$

as $n \rightarrow \infty$, uniformly in ζ and in $0 \leq \theta < \theta_0$, $\epsilon \leq s < -\theta\psi(\theta) - \epsilon$.

Proof. We first observe that

$$\begin{aligned} (q_+ - q_-)(\zeta; s, 0) &\geq c, & \zeta \in \Sigma^0, |\zeta \pm b| > \eta, \\ \Re q(\zeta; s, 0) &\leq -c|\zeta|, & \zeta \in \Sigma^+, |\zeta \pm b| > \eta, \\ \Re q(\zeta; s, 0) &\geq c|\zeta|, & \zeta \in \Sigma^-, |\zeta \pm b| > \eta, \end{aligned}$$

for $\epsilon \leq s < -\theta\psi(\theta) - \epsilon$, by construction of the contours, by (3.17), and by the fact that $q(\pm b; s, 0) = 0$. By continuity in $\theta \geq 0$, these inequalities still hold for θ sufficiently small. $\square$

4 Construction of parametrices for $\epsilon \leq s < -\theta\psi(\theta) - \epsilon$

4.1 The global parametrix

We define P^∞ as the solution of the formal large n limit of the RH problem for S , obtained after ignoring exponentially small jumps and small neighborhoods of $\pm b$.

RH problem for P^∞

($P^\infty 1$) $P^\infty : \mathbb{C} \setminus [-b, b] \rightarrow \mathbb{C}^{2 \times 2}$ is analytic.

($P^\infty 2$) On $(-b, b)$, P^∞ satisfies the jump condition

$$P_+^\infty(\zeta) = P_-^\infty(\zeta) \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

($P^\infty 3$) $P^\infty(\zeta) = I + \frac{P_1^\infty}{\zeta} + \mathcal{O}(\zeta^{-2})$ as $\zeta \rightarrow \infty$.

The solution to this RH problem is explicit and it is given by (see [19, Chapter 7] for a similar construction)

$$P^\infty(\zeta) = P^\infty(\zeta; b) = \frac{1}{2} \begin{pmatrix} \gamma(\zeta) + \gamma(\zeta)^{-1} & i(\gamma(\zeta) - \gamma(\zeta)^{-1}) \\ -i(\gamma(\zeta) - \gamma(\zeta)^{-1}) & \gamma(\zeta) + \gamma(\zeta)^{-1} \end{pmatrix}, \quad (4.1)$$

where

$$\gamma(\zeta) = \gamma(\zeta; b) = \left(\frac{\zeta - b}{\zeta + b} \right)^{1/4} \quad (4.2)$$

with the fourth root defined and analytic on $\mathbb{C} \setminus [-b, b]$, with branch cut on $(-b, b)$, and such that $\gamma(\zeta) \rightarrow 1$ as $\zeta \rightarrow \infty$. Equivalently,

$$P^\infty(\zeta) = \begin{pmatrix} 1 & -1 \\ -i & -i \end{pmatrix} \left(\frac{\zeta - b}{\zeta + b} \right)^{\sigma_3/4} \begin{pmatrix} 1 & -1 \\ -i & -i \end{pmatrix}^{-1}, \quad \text{for } \zeta \in \mathbb{C} \setminus [-b, b]. \quad (4.3)$$

We easily compute

$$P_1^\infty = \begin{pmatrix} 0 & -\frac{i}{2}b \\ \frac{i}{2}b & 0 \end{pmatrix}. \quad (4.4)$$

4.2 Local Airy parametrices

In small disks $U_{\pm b}$ of radius $\eta > 0$ around the endpoints $-b$ and b , we need to construct local parametrices. These local parametrices need to satisfy the following conditions.

RH problem for $P^{(\pm b)}$

(P1) $P^{(\pm b)} : U_{\pm b} \setminus \Sigma \rightarrow \mathbb{C}^{2 \times 2}$ is analytic.

(P2) $P_+^{(\pm b)}(\zeta) = P_-^{(\pm b)}(\zeta) J_S(\zeta)$ for $\zeta \in U_{\pm b} \cap \Sigma$.

(P3) On $\partial U_{\pm b}$, we have the (uniform in ζ and in $\epsilon \leq s \leq -\theta\psi(\theta) - \epsilon$) matching condition $P^{(\pm b)}(\zeta) = (I + \mathcal{O}(\frac{1}{n})) P^{(\infty)}(\zeta)$ as $n \rightarrow \infty$.

For the construction of parametrices, it is important to understand the $\zeta \rightarrow \pm b$ behavior of the function $q(\zeta; s, \theta)$ appearing in the jump matrices (3.18). Using the fact that q/a is an even function, we compute

$$\begin{aligned} \frac{q(\zeta)}{a(\zeta)} &= \int_{\mathcal{C}} \frac{h(u; s, \theta)}{a(u)(\zeta - u)} \frac{du}{4\pi i} \\ &= \int_{\mathcal{C}} \frac{h(u; s, \theta)u}{a(u)(\zeta^2 - u^2)} \frac{du}{4\pi i} \\ &= \int_{\mathcal{C}} \frac{h(u; s, \theta)u}{a(u)((\zeta^2 - b^2) - (u^2 - b^2))} \frac{du}{4\pi i} \\ &= - \int_{\mathcal{C}} \frac{h(u; s, \theta)u}{a(u)(u^2 - b^2)} \left(1 + \frac{\zeta^2 - b^2}{u^2 - b^2}\right) \frac{du}{4\pi i} + R_b(\zeta), \end{aligned}$$

where $R_b(\zeta)$ is an error term which is $O(a(\zeta)^4)$ as $\zeta \rightarrow \pm b$. The term

$$\int_{\mathcal{C}} \frac{h(u; s, \theta)u}{a(u)(u^2 - b^2)} \frac{du}{4\pi i}$$

vanishes: one easily sees this by integrating by parts (3.12). Hence we have

$$q(\zeta) \sim ca(\zeta)^3, \quad \zeta \rightarrow \pm b. \quad (4.5)$$

Using this local behavior of q like a $3/2$ -power, we can construct local parametrices in terms of a model RH problem whose solution can be built out of the Airy function and its derivative. Such local Airy parametrices were constructed originally in [1, 19, 22, 23], and their construction is by now standard. We omit the details of this construction, and refer the interested reader to [13, Section 3.5], where Airy parametrices were constructed in a situation almost identical to ours. For our purposes, we will only need the existence of such local parametrices $P^{(\pm b)}$ satisfying conditions (P1)–(P3), the precise form of the parametrices is unimportant.

4.3 The transformation $S \mapsto R$

We define

$$R(\zeta) = \begin{cases} S(\zeta)P^{(\pm b)}(\zeta)^{-1}, & \text{if } \zeta \in U_{\pm b}, \\ S(\zeta)P^\infty(\zeta)^{-1}, & \text{otherwise.} \end{cases}$$

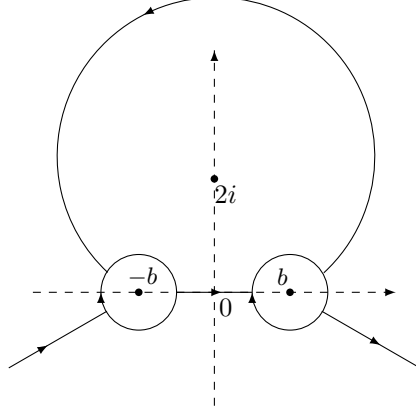


Figure 5: Jump contour Σ_R for R .

Then R solves the following RH problem.

RH problem for R

(R1) $R : \mathbb{C} \setminus \Sigma_R \rightarrow \mathbb{C}^{2 \times 2}$ is analytic, where $\Sigma_R = (\Sigma^+ \cup \Sigma^- \cup \Sigma^0 \cup \partial U_b \cup \partial U_{-b}) \setminus (U_b \cup U_{-b})$.

(R2) Across the contour Σ_R , R satisfies the jump condition

$$R_+(\zeta) = R_-(\zeta)J_R(\zeta),$$

where

$$J_R(\zeta) = \begin{cases} P^\infty(\zeta)J_S(\zeta)P^\infty(\zeta)^{-1} & \text{for } \zeta \in (\Sigma^+ \cup \Sigma^-) \setminus (U_b \cup U_{-b}), \\ P^\infty(\zeta)J_S(\zeta) \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} P^\infty(\zeta)^{-1} & \text{for } \zeta \in \Sigma^0 \setminus (U_b \cup U_{-b}), \\ P^{(\pm b)}(\zeta)P^\infty(\zeta)^{-1} & \text{for } \zeta \in \partial U_{\pm b}. \end{cases}$$

(R3) $R(\zeta) = I + \frac{R_1(n, s, \theta)}{\zeta} + \mathcal{O}(\zeta^{-2})$ as $\zeta \rightarrow \infty$ with

$$R_1(s, n, \theta) = S_1(s, n, \theta) - P_1^\infty = \frac{-iY_1(s, n, \theta) - n^2 g_1(s, \theta)\sigma_3}{n} - P_1^\infty. \quad (4.6)$$

By construction, the jump matrix for R is small as $n \rightarrow \infty$

Proposition 4.1. *Let $\epsilon > 0$. There exist $c, \theta_0 > 0$ such that*

$$J_R(\zeta) = I + O\left(\frac{1}{n(|\zeta|^2 + 1)}\right), \quad n \rightarrow \infty,$$

uniformly for $\zeta \in \Sigma_R$, $0 \leq \theta < \theta_0$, $\epsilon \leq s < -\theta\psi(\theta) - \epsilon$.

Proof. On $\partial U_{\pm b}$, the jump is $I + O(1/n)$ because of the matching condition (Pb3). Elsewhere, the jump matrix for J_S is I plus exponentially small terms which moreover decay as $\zeta \rightarrow \infty$, by Proposition 3.5 and the boundedness of $P^\infty(\zeta)$. $\square$

As in the case $s > -\theta\psi(\theta) + \epsilon$, it then follows from the general theory of RH problems that the RH solution R is also close to identity as $n \rightarrow \infty$.

Corollary 4.2. *Let $\epsilon > 0$. There exist $\theta_0 > 0$ such that*

$$R(\zeta; s, n, \theta) = I + O\left(\frac{1}{n(|\zeta| + 1)}\right), \quad n \rightarrow \infty,$$

uniformly for $\zeta \in \mathbb{C} \setminus \Sigma_R$, $0 \leq \theta < \theta_0$, $\epsilon \leq s < -\theta\psi(\theta) - \epsilon$. In particular, $R_1(s, n, \theta) = O(1/n)$ as $n \rightarrow \infty$, uniformly for $0 \leq \theta < \theta_0$, $\epsilon \leq s < -\theta\psi(\theta) - \epsilon$.

As a consequence, we have

$$Y_1(s, n, \theta) = in^2 g_1(s, \theta) \sigma_3 + in P_1^\infty + O(1), \quad n \rightarrow \infty \quad (4.7)$$

uniformly for $0 \leq \theta < \theta_0$, $\epsilon \leq s < -\theta\psi(\theta) - \epsilon$. By (3.11) and (4.4), we finally find

$$(Y_1(s, n, \theta))_{11} = n^2 f(s, \theta) + O(1), \quad n \rightarrow \infty \quad (4.8)$$

uniformly for $0 \leq \theta < \theta_0$, $\epsilon \leq s < -\theta\psi(\theta) - \epsilon$.

5 Construction of parametrices for $-\theta\psi(\theta) - \epsilon \leq s \leq -\theta\psi(\theta) + \epsilon$

Throughout this section, we assume that $\eta > 0$ is sufficiently small, and we take $\epsilon > 0$ sufficiently small such that $b(s, \theta) < \eta$ for $-\theta\psi(\theta) - \epsilon \leq s \leq -\theta\psi(\theta) + \epsilon$. We can do this by (1.12). We denote U_0 for the disk with radius η centered at 0.

5.1 The global parametrix

By Proposition 5.2, we know that the jump matrix J_S for S is close to I on $\Sigma^\pm \setminus U_0$ for n large. For this reason, we need a global parametrix which is analytic outside U_0 , and which tends to I at infinity. While we could simply set $P^\infty(\zeta) = I$ at this point, this is not a convenient choice for $-\theta\psi(\theta) - \epsilon \leq s < -\theta\psi(\theta)$: indeed, this would not allow for a good matching with the local parametrix near 0 later on.

For $-\theta\psi(\theta) - \epsilon \leq s < -\theta\psi(\theta)$, we rather define $P^\infty(\zeta) = P^\infty(\zeta; b)$ as before, given by (4.1), with $b = b(s, \theta)$ the unique solution of (1.8). We recall that P^∞ solves the RH conditions $(P^\infty 1) - (P^\infty 3)$, and that $b(s, \theta) > 0$ is small for s close to $-\theta\psi(\theta)$. For any $\eta > 0$, we can choose $\epsilon > 0$ sufficiently small such that $b(s, \theta) < \eta$, which implies that P^∞ is analytic for $|\zeta| > \eta$.

For $-\theta\psi(\theta) \leq s \leq -\theta\psi(\theta) + \epsilon$, we set $P^\infty(\zeta) = I$. Note that this is the same as (1.8) with $b = 0$.

5.2 Model RH problem associated to the Painlevé II equation

We will construct a local parametrix in U_0 , which satisfies the same jump condition as S and which matches with P^∞ on the boundary of U_0 .

RH problem for P

(P1) $P : \mathbb{C} \setminus U_0 \rightarrow \mathbb{C}^{2 \times 2}$ is analytic.

(P2) $P_+(\zeta) = P_-(\zeta) J_S(\zeta)$ for $\zeta \in U_0 \cup \Sigma$.

(P3) On ∂U_0 , we have the (uniform in ζ , $0 < \theta < \theta_0$, and in $-\theta\psi(\theta) - \epsilon \leq s \leq -\theta\psi(\theta) + \epsilon$) matching condition $P(\zeta) = (I + O(n^{-1/3})) P^\infty(z)$ as $n \rightarrow \infty$.

To construct the local parametrix P , we will use the solution $\Psi(z) = \Psi(z; r)$ of a well-known model RH problem, connected to the Hastings-McLeod solution of the Painlevé II equation, and depending on a parameter $r \in \mathbb{C}$. This RH problem is due to Flaschka and Newell [28] and was used and studied further in [1, 4, 15, 25, 29].

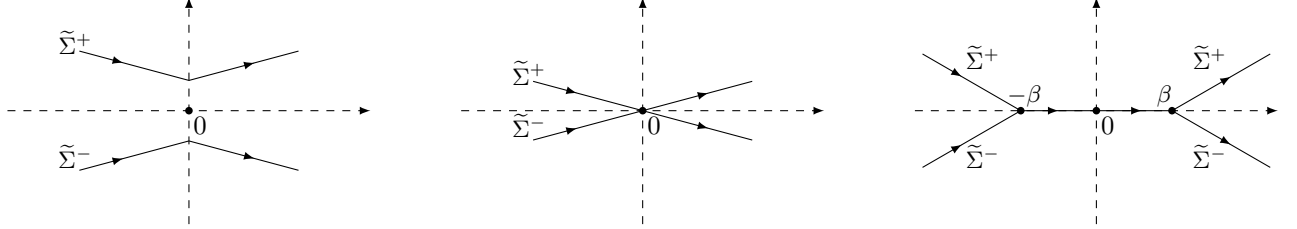


Figure 6: The shape of the jump contour $\tilde{\Sigma}$ for the RH problem for Ψ (on the left), for the RH problem for $\hat{\Psi}$ with $\beta = 0$ (in the middle), and for $\hat{\Psi}$ with $\beta > 0$ (on the right). Compare this with Figure 4.

RH problem for Ψ

(Ψ 1) $\Psi : \mathbb{C} \setminus \tilde{\Sigma}$ is analytic, with $\tilde{\Sigma} = \tilde{\Sigma}^+ \cup \tilde{\Sigma}^-$, and $\tilde{\Sigma}^+$ lies in the upper half plane, $\tilde{\Sigma}^-$ lies in the lower half plane, and $\tilde{\Sigma}$ lies for large z within the region $\{\delta < |\arg z| < \frac{\pi}{3} - \delta\} \cup \{\delta < |\arg(-z)| < \frac{\pi}{3} - \delta\}$ for some $\delta > 0$, like in the left picture of Figure 6.

(Ψ 2) Ψ satisfies the jump conditions

$$\Psi_+(z) = \Psi_-(z) = \begin{cases} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} & \text{if } z \in \tilde{\Sigma}^+, \\ \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} & \text{if } z \in \tilde{\Sigma}^-. \end{cases}$$

(Ψ 3) As $z \rightarrow \infty$, we have

$$\Psi(z)e^{i(4/3z^3 + rz)\sigma_3} = I + \Psi_1(r)z^{-1} + \Psi_2(r)z^{-2} + O(z^{-3}), \quad (5.1)$$

for some matrices $\Psi_1(r), \Psi_2(r)$.

There is an equivalent RH problem, obtained by letting $\tilde{\Sigma}^+$ and $\tilde{\Sigma}^-$ coincide on an interval $[-\beta, \beta]$ on the real line, with $\beta = \sqrt{-r/2}$, which will also be useful for us.

RH problem for $\hat{\Psi}$

($\hat{\Psi}$ 1) $\hat{\Psi} : \mathbb{C} \setminus \tilde{\Sigma}$ is analytic, with $\tilde{\Sigma} = \tilde{\Sigma}^+ \cup \tilde{\Sigma}^- \cup [-\beta, \beta]$ as in the right picture of Figure 6.

($\widehat{\Psi}2$) $\widehat{\Psi}$ satisfies the jump conditions

$$\widehat{\Psi}_+(z) = \widehat{\Psi}_-(z) = \begin{cases} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} & \text{if } z \in \widetilde{\Sigma}^+, \\ \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix} & \text{if } z \in (-\beta, \beta), \\ \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} & \text{if } z \in \widetilde{\Sigma}^-. \end{cases}$$

($\widehat{\Psi}3$) As $z \rightarrow \infty$, we have

$$\widehat{\Psi}(z)e^{i(4/3z^3+rz)\sigma_3} = I + \Psi_1(r)z^{-1} + \Psi_2(r)z^{-2} + O(z^{-3}). \quad (5.2)$$

($\widehat{\Psi}4$) $\widehat{\Psi}$ is bounded near $\pm\beta$.

These two equivalent RH problems have been studied intensively in the past decades and are by now well understood. Below, we summarize the relevant known results about its solution, for which we refer the reader to [29, 28, 15].

First, we have that the solutions $\Psi, \widehat{\Psi}$ exist for any s in a complex neighborhood of the real line, and that the matrices $\Psi_1(r), \Psi_2(r)$ are independent of the choice of jump contour $\widetilde{\Sigma}$. Secondly, the RH solution is connected to the Painlevé II equation in the following sense. We have

$$\Psi_1(r) = \frac{1}{2i} \begin{pmatrix} u(r) & q(r) \\ -q(r) & -u(r) \end{pmatrix}, \quad (5.3)$$

where q is the Hastings-McLeod solution of the Painlevé II equation, i.e.

$$q''(r) = rq(r) + 2q(r)^3 \quad \text{and } q(r) \sim \text{Ai}(r) \text{ as } r \rightarrow +\infty,$$

and $u(r) = q'(r)^2 - sq(r)^2 - q(r)^4$, such that $u'(r) = -q(r)^2$.

The expansion (5.1) in ($\Psi3$) is uniform for r in compact subsets of the complex r -plane in which the RH problem is solvable, and it is also uniform as $r \rightarrow \infty$ with $|\arg r| < \delta$ with $\delta > 0$ small. As $r \rightarrow -\infty$ or $r \rightarrow \infty$ with $|\arg r - \pi| < \delta$, it is not uniform, but one can rewrite it in order to obtain a uniform expansion. As $r \rightarrow \infty$ with $|\arg r - \pi| < \delta$, we have the following asymptotics, uniform for $|z| > \sqrt{-r}$ (see for instance [29, Chapter 8] or [25]):

$$\widehat{\Psi}(z, r)e^{\frac{4i}{3}(z^2+rz/2)^{3/2}\sigma_3} = \left(I + O\left(\frac{1}{(|r|+1)z}\right) \right) P^\infty(z; \sqrt{-r/2}), \quad z \rightarrow \infty, \quad (5.4)$$

with $P^\infty(z; \sqrt{-r/2})$ given by (4.1) with $b = \sqrt{-r/2}$.

5.3 Construction of the local parametrix for $-\theta\psi(\theta) \leq s \leq -\theta\psi(\theta) + \epsilon$

We first record the following result stating that the jump matrix $J_T(\zeta)$ is close to I as $n \rightarrow \infty$, except for ζ close to 0. We omit the proof, since it is the same as that of Proposition 3.2, for $|\zeta| > \rho$.

Proposition 5.1. *Let $\rho > 0$. There exist $\epsilon, c, \theta_0 > 0$ such that*

$$J_T(\zeta) = I + O(e^{-cn|\zeta|}), \quad \zeta \in \Sigma^\pm, \quad |\zeta| > \rho,$$

as $n \rightarrow \infty$, uniformly in ζ and in $0 \leq \theta < \theta_0$, $-\theta\psi(\theta) \leq s < -\theta\psi(\theta) + \epsilon$.

This is not sufficient to conclude that T is close to I , because the convergence of the jump matrices breaks down near 0. Therefore, we need to construct a local parametrix near 0, for which we will use $\Psi(z; r)$.

We will construct P of the following form:

$$P(\zeta) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \Psi(n^{1/3} f(\zeta), r_n(\zeta)) \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} e^{\frac{n}{2} h(\zeta) \sigma_3}, \quad (5.5)$$

where q is defined by (3.14). Here f will be a conformal map in U_0 with $f(0) = 0$ and $f'(0) > 0$, and r_n will be analytic in U_0 and uniformly bounded in n, s, ζ . First, we need P to have the same jump contour as S , and therefore we need to take $\Sigma, \tilde{\Sigma}$ and f such that $n^{1/3} f(\Sigma \cap U_0) \subset \tilde{\Sigma}$. If this holds, then it is straightforward to check that P and Ψ satisfy the same jump relations on $\Sigma \cap U_0$. The remaining freedom to define f and r_n will be needed to achieve the matching condition (P3). By condition (P3) and (5.5), it follows that the matching condition

$$P(\zeta) P^\infty(\zeta)^{-1} = P(\zeta) = I + O(n^{-1/3}), \quad n \rightarrow \infty, \quad (5.6)$$

holds uniformly for $\zeta \in \partial U_0$, $-\theta\psi(\theta) \leq s \leq -\theta\psi(\theta) + \epsilon$, provided that we have the identity

$$nh(\zeta; s, \theta) = -\frac{8i}{3} n f(\zeta)^3 - 2in^{1/3} r_n(\zeta) f(\zeta). \quad (5.7)$$

To achieve this, we define

$$f(\zeta) = \left(\frac{3i}{8} h(\zeta; -\theta\psi(\theta), \theta) \right)^{1/3}, \quad r_n(\zeta) = n^{2/3} (s + \theta\psi(\theta)) \frac{\zeta}{2f(\zeta)}. \quad (5.8)$$

It is then straightforward to check that (5.7) holds, by (3.4). Since

$$h(0; -\theta\psi(\theta), \theta) = h'(0; -\theta\psi(\theta), \theta) = h''(0; -\theta\psi(\theta), \theta) = 0, \quad h'''(0; -\theta\psi(\theta), \theta) = \frac{i\theta^3 \psi''(\theta)}{4},$$

we also have, as required, that f is a conformal map, with

$$f(0) = 0, \quad f'(0) = \frac{\theta(-\psi''(\theta))^{1/3}}{4} > 0. \quad (5.9)$$

Furthermore, r_n is analytic near 0, and we have

$$r_n(0) = \frac{2}{\theta(-\psi''(\theta))^{1/3}} (s + \theta\psi(\theta)) n^{2/3} \geq 0. \quad (5.10)$$

5.4 Construction of the local parametrix for $-\theta\psi(\theta) - \epsilon \leq s < -\theta\psi(\theta)$

For $s + \theta\psi(\theta) = O(n^{-2/3})$ as $n \rightarrow \infty$, we could have proceeded directly with the construction of a local Painlevé II parametrix like before, without the need to use a g -function transformation, following the method of [15]. However, since we need uniform asymptotics for $-\theta\psi(\theta) - \epsilon \leq s \leq -\theta\psi(\theta) + \epsilon$, we are forced to use a g -function transformation here as well, similarly to [1]. As already mentioned, we let $\eta > 0$ be a sufficiently small number, and we let $\epsilon > 0$ be such that $|b(s, \theta)| < \eta$ for $-\theta\psi(\theta) - \epsilon \leq s < -\theta\psi(\theta)$. We first prove that the jump matrix $J_S(\zeta)$ is close to I as $n \rightarrow \infty$, except for ζ close to 0.

Proposition 5.2. *Let $\eta > 0$. There exist $\epsilon, c, \theta_0 > 0$ such that*

$$J_S(\zeta) = I + O(e^{-cn|\zeta|}), \quad \zeta \in \Sigma^\pm, \quad |\zeta| > \eta,$$

as $n \rightarrow \infty$, uniformly in ζ and in $0 \leq \theta < \theta_0$, $-\theta\psi(\theta) - \epsilon \leq s < -\theta\psi(\theta) + \epsilon$.

Proof. For $\theta = 0$ and $s = 1$, we have

$$q(\zeta; 1, 0) = -\frac{1}{2}h(\zeta; 1, 0) = -\frac{1}{2}\log \frac{\zeta - 2i}{\zeta + 2i} + \frac{i}{2}\zeta.$$

Recall that we have chosen the contours $\Sigma^\pm$ such that

$$\Re q(\zeta; 1, 0) \geq C, \quad \zeta \in \Sigma^+, |\zeta| > \eta,$$

and

$$\Re q(\zeta; 1, 0) \leq -C, \quad \zeta \in \Sigma^-, |\zeta| > \eta,$$

for some $C > 0$. From the behavior of q at infinity, it follows moreover that we can replace C by $c|\zeta|$ with $c > 0$ sufficiently small in the above estimates. By continuity of $q(\zeta; s, \theta)$ in s, θ for $|\zeta| > \eta$, it follows that these estimates continue to hold for $0 < \theta < \theta_0$ with $\theta_0 > 0$ sufficiently small, and s sufficiently close to 1. Note moreover that any interval containing 1 contains also $-\theta\psi(\theta)$ for $\theta > 0$ small, since $\lim_{\theta \rightarrow 0}(-\theta\psi(\theta)) = 1$. The result follows. $\square$

We already know that q behaves like a $3/2$ -power near $\pm b$, but we are now in a situation where $b(s, \theta)$ can converge to 0 in the limit where $n \rightarrow \infty$. This is why we now need uniform control of q in a fixed neighborhood of 0, containing both $\pm b$.

Proposition 5.3. *There exist $\eta, \epsilon, \theta_0 > 0$ such that*

$$\frac{-iq(\zeta)}{a(\zeta)^3} = c(s) + O(a(\zeta)^2),$$

uniformly for $0 < \theta < \theta_0$, $-\theta\psi(\theta) - \epsilon \leq s < -\theta\psi(\theta)$, and $|\zeta| \leq \eta$. Furthermore,

$$c(s) = -\theta^3\psi''(\theta)/48 + O(s + \theta\psi(\theta)), \quad s \rightarrow -\theta\psi(\theta).$$

Proof. We already computed (see the computation before (4.5))

$$\frac{-iq(\zeta)}{a(\zeta)} = a(\zeta)^2 \int_{\mathcal{C}} \frac{h(u)u}{a(u)(u^2 - b^2)^2} \frac{du}{4\pi} + O(a(\zeta)^4),$$

uniformly for $|\zeta| < \eta$ with η sufficiently small. We define

$$c(s) = \int_{\mathcal{C}} \frac{h(u)u}{a(u)(u^2 - b^2)^2} \frac{du}{4\pi}.$$

Then, as $b = b(s, \theta) \rightarrow 0$, we find by a residue computation that

$$c(s) = \int_{\mathcal{C}} \frac{h(u)}{u^4} \frac{du}{4\pi} + O(b^2) = -\frac{\theta^3\psi''(\theta)}{48} + O(b^2).$$

By (1.12), the error term is $O(s + \theta\psi(\theta))$ as $n \rightarrow \infty$, and the result follows. $\square$

We will now construct P of the following form, similar to (5.5) but with $\widehat{\Psi}$ instead of Ψ and with q instead of h :

$$P(\zeta) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \widehat{\Psi}(n^{1/3}f(\zeta), r_n(\zeta)) \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} e^{-nq(\zeta)\sigma_3}. \quad (5.11)$$

Here f will be the same as in (5.8), and $r_n(\zeta)$ will be analytic in U_0 and uniformly bounded. As before, we will take $\Sigma, \tilde{\Sigma}$ and f such that $n^{1/3}f(\Sigma \cap U_0) \subset \tilde{\Sigma} = \Sigma^+ \cup \Sigma^- \cup [-\beta, \beta]$. This fixes in particular the value of $\beta = n^{1/3}f(b(s, \theta))$.

It is then again straightforward to verify that P and Ψ satisfy the same jump relations on $\Sigma \cap U_0$. The remaining freedom to define r_n will be needed to achieve the matching condition (P3). By (5.4) and (5.11), it follows that the matching condition

$$P(\zeta)P^\infty(\zeta)^{-1} = I + O\left(\frac{1}{n^{1/3}(|r_n(\zeta)| + 1)}\right), \quad n \rightarrow \infty, \quad (5.12)$$

holds uniformly for $\zeta \in \partial U_0$, $-\theta\psi(\theta) - \epsilon < s < -\theta\psi(\theta)$, is valid provided that we have the identity

$$nq(\zeta; s, \theta) = \frac{4i}{3} \left(n^{2/3}f(\zeta)^2 + \frac{1}{2}r_n(\zeta) \right)^{3/2}. \quad (5.13)$$

To achieve this, we define

$$r_n(\zeta) = 2 \left(\frac{-3i}{4} nq(\zeta; s, \theta) \right)^{2/3} - 2n^{2/3}f(\zeta)^2, \quad (5.14)$$

which we can rewrite as

$$r_n(\zeta) = 2n^{2/3} \left(\frac{-3i}{4} q(\zeta; s, \theta) \right)^{2/3} - 2n^{2/3} \left(\frac{-3i}{4} q(\zeta; -\theta\psi(\theta), \theta) \right)^{2/3}$$

By Proposition 5.3, r_n is analytic in U_0 , and using also (1.12), we obtain

$$\begin{aligned} r_n(\zeta) &= \frac{b^2}{8} \theta^2 (-\psi''(\theta))^{2/3} n^{2/3} + O(n^{2/3}b^4) + O(n^{2/3}b^2\zeta^2) \\ &= \frac{2}{\theta(-\psi''(\theta))^{1/3}} n^{2/3} (s + \theta\psi(\theta)) (1 + O(s + \theta\psi(\theta)) + O(\zeta^2)), \end{aligned}$$

uniformly in $\zeta \in U_0$, $0 < \theta < \theta_0$, $-\theta\psi(\theta) - \epsilon \leq s < -\theta\psi(\theta)$. For $\epsilon > 0$ small, it follows that $r_n(\zeta)$ lies in a small sector around the negative real line, hence the RH problem for $\hat{\Psi}(\cdot; r = r_n(\zeta))$ is solvable, and we can indeed use the asymptotic relation (5.4).

5.5 The transformation $S \mapsto R$

We define

$$R(\zeta) = T(\zeta) \begin{cases} P(\zeta)^{-1}, & \text{if } \zeta \in U_0, \\ I, & \text{otherwise,} \end{cases} \quad \text{if } -\theta\psi(\theta) \leq s \leq -\theta\psi(\theta) + \epsilon,$$

and

$$R(\zeta) = S(\zeta) \begin{cases} P(\zeta)^{-1}, & \text{if } \zeta \in U_0, \\ P^\infty(\zeta; b(s, \theta))^{-1}, & \text{otherwise,} \end{cases} \quad \text{if } -\theta\psi(\theta) - \epsilon \leq s < -\theta\psi(\theta).$$

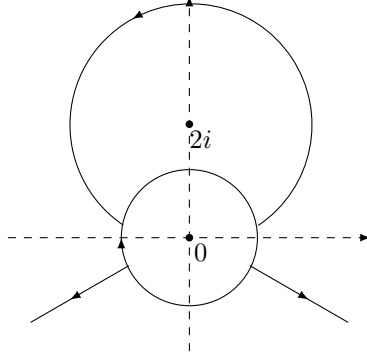


Figure 7: Jump contour Σ_R for R .

By construction, R solves the following RH problem.

RH problem for R

(R1) $R : \mathbb{C} \setminus \Sigma_R \rightarrow \mathbb{C}^{2 \times 2}$ is analytic.

(R2) Across the contour Σ_R , R satisfies the jump condition

$$R_+(\zeta) = R_-(\zeta)J_R(\zeta),$$

where

$$J_R(\zeta) = I + O(n^{-1/3}) \quad (5.15)$$

for ζ on Σ_R .

(R3) $R(\zeta) = I + \frac{R_1(s, n, \theta)}{\zeta} + \mathcal{O}(\zeta^{-2})$ as $\zeta \rightarrow \infty$ with

$$R_1(s, n, \theta) = S_1(s, n, \theta) - P_1^\infty = \frac{-iY_1(s, n, \theta) - n^2 g_1(s, \theta) 1_{(0, -\theta\psi(\theta))}(s) \sigma_3}{n} - P_1^\infty 1_{(0, -\theta\psi(\theta))}(s).$$

As before, it then follows that

$$R(\zeta) = I + O\left(\frac{1}{n^{1/3}(|z| + 1)}\right)$$

as $n \rightarrow +\infty$, uniformly in $z \in \mathbb{C} \setminus \Sigma_R$.

It follows that

$$R_1(s, n, \theta) = O(n^{-1/3}), \quad n \rightarrow \infty. \quad (5.16)$$

Hence,

$$Y_1(s, n, \theta) = in^2 g_1(s, \theta) 1_{(0, -\theta\psi(\theta))}(s) \sigma_3 + in P_1^\infty 1_{(0, -\theta\psi(\theta))}(s) + O(n^{2/3}), \quad (5.17)$$

and

$$(Y_1(s, n, \theta))_{11} = in^2 g_1(s, \theta) 1_{(0, -\theta\psi(\theta))}(s) + O(n^{2/3}) = n^2 f(s, \theta) 1_{(0, -\theta\psi(\theta))}(s) + O(n^{2/3}), \quad (5.18)$$

as $n \rightarrow \infty$, uniformly for $-\theta\psi(\theta) - \epsilon \leq s \leq -\theta\psi(\theta) + \epsilon$.

6 Proof of Theorem 1.6

We recall the result from Proposition 2.1:

$$\frac{d}{ds} \log Q_n^\theta(s) = (Y_1(s, n, \theta))_{11},$$

and we substitute the results from the large n asymptotic RH analysis in the right hand side. Let $\epsilon > 0$ be sufficiently small. For $s > -\theta\psi(\theta) + \epsilon$, we have by Corollary 3.3 that $Y_1(s, n, \theta) = O(ne^{-cns})$ as $n \rightarrow \infty$; for $-\theta\psi(\theta) - \epsilon \leq s \leq -\theta\psi(\theta) + \epsilon$, we have (5.18); finally, for $\epsilon < s < -\theta\psi(\theta) - \epsilon$, we have (4.8). We thus obtain the uniform in s large n asymptotics

$$\frac{d}{ds} \log Q_n^\theta(s) = \begin{cases} O(ne^{-cns}), & s > -\theta\psi(\theta) + \epsilon, \\ n^2 f(s, \theta) 1_{(0, -\theta\psi(\theta))}(s) + O(n^{2/3}), & -\theta\psi(\theta) - \epsilon \leq s \leq -\theta\psi(\theta) + \epsilon, \text{ as } n \rightarrow \infty. \\ n^2 f(s, \theta) + O(1), & \epsilon < s \leq -\theta\psi(\theta) - \epsilon, \end{cases}$$

Integrating in s from s to $+\infty$ and observing that $\log Q_n^\theta(+\infty) = 0$, we obtain

$$\log Q_n^\theta(s) = \begin{cases} O(e^{-c'ns}), & s > -\theta\psi(\theta) + \epsilon, \\ -n^2 F(s, \theta) 1_{(0, -\theta\psi(\theta))}(s) + O(n^{2/3}), & -\theta\psi(\theta) - \epsilon \leq s \leq -\theta\psi(\theta) + \epsilon, \\ -n^2 F(s, \theta) + O(n^{2/3}), & \epsilon < s \leq -\theta\psi(\theta) - \epsilon, \end{cases}$$

as $n \rightarrow \infty$. This proves Theorem 1.6. The RH analysis allows in fact to obtain much stronger large n asymptotics for $\log Q_n^\theta(s)$. For instance, we can expand the jump matrix for R as far as we want in powers of n as $n \rightarrow \infty$, and use this expansion to compute subleading asymptotic terms for R , and thus for Y . In this way, after long but straightforward computations, we could prove the stronger asymptotics

$$\log Q_n^\theta(s) = -\log F_{\text{TW}} \left(\frac{2n^{2/3}(s + \theta\psi(\theta))}{\theta(-\psi''(\theta))^{1/3}} \right) + o(1), \quad (6.1)$$

as $n \rightarrow \infty$ and $s \rightarrow -\theta\psi(\theta)$ with $n^{2/3}(s + \theta\psi(\theta)) = O(1)$. These stronger asymptotics can be used to reprove (1.5), but would not lead to a stronger result about the probability distribution of the log-Gamma polymer in Conjecture 1.1.

7 Proof of Conjecture 1.1 under Ansatz 1.5

Assuming that Ansatz 1.5 holds, we now know from (1.27), (1.29), and Theorem 1.6 that

$$\lim_{n \rightarrow \infty} \frac{1}{n^2} \log \mathbb{E} \left(e^{-e^{-\frac{2n}{\theta}s} Z_n(\theta)} \right) = \lim_{n \rightarrow \infty} \frac{1}{n^2} \log \tilde{Q}_n^\theta(s) = \lim_{n \rightarrow \infty} \frac{1}{n^2} \log Q_n^\theta(s) = -F(s, \theta), \quad (7.1)$$

where $Z_n(\theta)$ is the log-Gamma polymer partition function associated with θ and n . From this, we can draw conclusions for the log-Gamma polymer partition function. We are interested in the large n behavior of the probability

$$\mathbb{P} \left[\log Z_n(\theta) \leq \frac{2n}{\theta} s \right].$$

For any $s > s_0 \in \mathbb{R}$, we have the inequalities

$$\begin{aligned} & \mathbb{E} e^{-e^{\log Z_n(\theta) - \frac{2n}{\theta} s_0}} \\ &= \mathbb{E} \left[1_{\{\log Z_n(\theta) \leq \frac{2n}{\theta} s\}} e^{-e^{\log Z_n(\theta) - \frac{2n}{\theta} s_0}} + 1_{\{\log Z_n(\theta) > \frac{2n}{\theta} s\}} e^{-e^{\log Z_n(\theta) - \frac{2n}{\theta} s_0}} \right] \\ &\leq \mathbb{P} \left[\log Z_n(\theta) \leq \frac{2n}{\theta} s \right] + \mathbb{E} \left[1_{\{\log Z_n(\theta) > \frac{2n}{\theta} s\}} e^{-e^{\log Z_n(\theta) - \frac{2n}{\theta} s_0}} \right]. \end{aligned}$$

Note that, whenever $\log Z_n(\theta) > \frac{2n}{\theta} s$, it holds that

$$e^{\log Z_n(\theta) - \frac{2n}{\theta} s_0} = e^{\log Z_n(\theta) - \frac{2n}{\theta} s} e^{\frac{2n}{\theta}(s - s_0)} \geq e^{\frac{2n}{\theta}(s - s_0)}$$

and thus

$$e^{-e^{\log Z_n(\theta) - \frac{2n}{\theta} s_0}} \leq e^{-e^{\frac{2n}{\theta}(s - s_0)}},$$

which implies

$$\mathbb{P} \left[\log Z_n(\theta) \leq \frac{2n}{\theta} s \right] \geq \mathbb{E} e^{-e^{\log Z_n(\theta) - \frac{2n}{\theta} s_0}} - e^{-e^{\frac{2n}{\theta}(s - s_0)}}.$$

We can now set, for instance, $s = s_0 + \frac{3\theta \log n}{2n}$, such that

$$e^{-e^{\frac{2n}{\theta}(s - s_0)}} = e^{-e^{3 \log n}} = e^{-n^3},$$

and for sufficiently large n ,

$$\mathbb{P} \left[\log Z_n(\theta) \leq \frac{2n}{\theta} s \right] \geq \mathbb{E} e^{-e^{\log Z_n(\theta) - \frac{2n}{\theta} s_0}} - e^{-n^3} \geq \frac{1}{2} \mathbb{E} e^{-e^{\log Z_n(\theta) - \frac{2n}{\theta} s_0}}, \quad (7.2)$$

where we used (7.1) in the last step to conclude that the first term dominates the second for large n .

On the other hand, Chernoff's bound (or Markov's inequality) yields

$$\begin{aligned} \mathbb{P} \left[\log Z_n(\theta) \leq \frac{2n}{\theta} s \right] &= \mathbb{P} \left[e^{-e^{\log Z_n(\theta) - \frac{2n}{\theta} s}} \geq e^{-1} \right] \\ &\leq e \mathbb{E} \left[e^{-e^{\log Z_n(\theta) - \frac{2n}{\theta} s}} \right] \end{aligned}$$

implying

$$\log \mathbb{P} \left[\log Z_n(\theta) \leq \frac{2n}{\theta} s \right] \leq \log \mathbb{E} \left[e^{-e^{\log Z_n(\theta) - \frac{2n}{\theta} s}} \right] + 1. \quad (7.3)$$

Hence, by combining (7.2) and (7.3) with (7.1), we obtain

$$\lim_{n \rightarrow \infty} \frac{1}{n^2} \log \mathbb{P} \left[\log Z_n(\theta) \leq \frac{2n}{\theta} s \right] = -F(s, \theta),$$

which implies Conjecture 1.1.

8 Evidence for Ansatz 1.5

In order to understand why Ansatz 1.5 is plausible, we need to take a closer look at the signed biorthogonal measures $d\mu_n^\theta(x_1, \dots, x_n)$ defined in (1.22)–(1.23). As already mentioned in the introduction, it is convenient to consider the re-scaled biorthogonal measure

$$d\hat{\mu}_n^\theta(y_1, \dots, y_n) := d\mu_n\left(\frac{2n}{\theta}y_1, \dots, \frac{2n}{\theta}y_n\right) = \frac{1}{n!} \det\left(\hat{L}_n^\theta(y_i, y_j)\right)_{i,j=1}^n dy_1 \cdots dy_n, \quad (8.1)$$

with $\hat{L}_n^\theta$ given by (1.25).

Macroscopic limit of the one-point function. For $\theta = 0$, $d\hat{\mu}_n^0$ is the eigenvalue distribution of the Wishart-Laguerre random matrix ensemble, and it is well-known that $\frac{1}{n}\hat{L}_n^0(y, y)$ converges as $n \rightarrow \infty$ to the Marchenko-Pastur law, recall (1.26). In other words, the particles $y_1, \dots, y_n$ admit a limiting one-point function. There is no indication of a phase transition taking place at $\theta = 0$, so we expect that the limit

$$\lim_{n \rightarrow \infty} \frac{1}{n} \hat{L}_n^\theta(y, y) =: h^\theta(y)$$

exists also for $\theta > 0$ sufficiently small, at least when $y > 0$. This can presumably be proved using a classical saddle point analysis of the double integral expression (1.25) for the kernel $\hat{L}_n^\theta$. The Tracy-Widom limit (1.31) suggests in addition that the right-most endpoint of the support of h^θ is given by $-\theta\psi(\theta)$, and that $h^\theta(y)$ behaves like a square root as $y \rightarrow -\theta\psi(\theta)$. We know that

$$Q_n^\theta(s) = \int_{\mathbb{R}^n} \prod_{j=1}^n (1 - 1_{(s, +\infty)}(y_j)) d\hat{\mu}_n^\theta(y_1, \dots, y_n) = \int_{(-\infty, s)^n} d\hat{\mu}_n^\theta(y_1, \dots, y_n),$$

while

$$\tilde{Q}_n^\theta(s) = \int_{\mathbb{R}^n} \prod_{j=1}^n (1 - \sigma_{s,n,\theta}(y_j)) d\hat{\mu}_n^\theta(y_1, \dots, y_n), \quad \sigma_{s,n,\theta}(y) = \frac{1}{1 + e^{-\frac{2n}{\theta}(y-s)}}.$$

As $n \rightarrow \infty$, the function $1 - \sigma_{s,n,\theta}(y)$ converges point-wise to $1 - 1_{(s, +\infty)}(y)$. From a broad perspective, one may already expect at this point that $Q_n^\theta(s)$ and $\tilde{Q}_n^\theta(s)$ have the same rate of decay, because the large deviation rate function is a global quantity associated to a determinantal point process that typically does not depend on microscopic deformations of the involved test function, such that one expects

$$\lim_{n \rightarrow \infty} \frac{1}{n^2} \log \tilde{Q}_n^\theta(s) = \lim_{n \rightarrow \infty} \frac{1}{n^2} \log Q_n^\theta(s),$$

which implies Ansatz 1.5.

Jacobi's identity. We will now develop some more concrete analytical estimates in support of Ansatz 1.5. We define a deformation of Q_n^θ as follows: we define

$$Q_n^\theta(s; t) := \det\left(1 - \sigma_{s,n,\theta}^t \hat{L}_n^\theta\right)_{L^2(\mathbb{R})}, \quad \sigma_{s,n,\theta}^t(y) = \frac{1}{1 + e^{-\frac{2n}{\theta t}(y-s)}}, \quad (8.2)$$

such that $Q_n^\theta(s, 0) = Q_n^\theta(s)$ and $Q_n^\theta(s, 1) = \tilde{Q}_n^\theta(s)$. By Jacobi's variational identity, we have

$$\partial_t \log Q_n(s; t) = \text{Tr} \left(-(\partial_t \sigma_{s,n,\theta}^t) \hat{L}_n^\theta \left(1 - \sigma_{s,n,\theta}^t \hat{L}_n^\theta\right)^{-1} \right).$$

Abbreviating

$$K_{s,n,\theta}^t := \widehat{L}_n^\theta \left(1 - \sigma_{s,n,\theta}^t \widehat{L}_n^\theta\right)^{-1}, \quad (8.3)$$

and writing $K_{s,n,\theta}^t(y, y')$ for a kernel of this operator, we have

$$\partial_t \log Q_n(s; t) = \frac{2n}{\theta t^2} \int_{\mathbb{R}} \frac{e^{-\frac{2n}{\theta t}(y-s)}}{\left(1 + e^{-\frac{2n}{\theta t}(y-s)}\right)^2} (y-s) K_{s,n,\theta}^t(y, y) dy.$$

We can bound the integral on the right as follows:

$$|\partial_t \log Q_n(s; t)| \leq \frac{\theta}{2n} \sup_{y \in \mathbb{R}} |K_{s,n,\theta}^t(y, y)| \int_{\mathbb{R}} \frac{e^{-u}}{(1 + e^{-u})^2} |u| du. \quad (8.4)$$

We believe that the following holds, for reasons that we will explain below.

Ansatz 8.1. *We have the estimate*

$$\sup_{y \in \mathbb{R}} |K_{s,n,\theta}^t(y, y)| = O(n^2), \quad n \rightarrow \infty, \quad (8.5)$$

uniformly for $s > \epsilon$, $0 \leq t \leq 1$, and $0 < \theta < \theta_0$.

This would imply

$$|\partial_t \log Q_n(s; t)| \leq \frac{C\theta n}{2} \int_{\mathbb{R}} \frac{e^{-u}}{(1 + e^{-u})^2} |u| du = O(\theta n), \quad n \rightarrow \infty, \quad (8.6)$$

and hence by integrating between $t = 0$ and $t = 1$,

$$\log Q_n(s, 1) - \log Q_n(s, 0) = \log \widetilde{Q}_n(s) - \log Q_n(s) = O(\theta n), \quad n \rightarrow \infty. \quad (8.7)$$

Combining this with Theorem 1.6, we obtain Ansatz 1.5 and thus Conjecture 1.1, assuming Ansatz 8.1.

Evidence for Ansatz 8.1. The kernel $K_{s,n,\theta}^t(y, y')$ appearing in Ansatz 8.1 has itself an interpretation in terms of a biorthogonal measure. Whereas $\widehat{L}_n^\theta(y, y')$ is the kernel of the signed biorthogonal measure $d\widehat{\mu}_n^\theta$, $K_{s,n,\theta}^t$ is the kernel of a deformation of this biorthogonal measure. Consider the signed biorthogonal measure

$$d\widehat{\nu}_{n,s,\theta}^t(y_1, \dots, y_n) := \frac{1}{Z_n} \prod_{j=1}^n (1 - \sigma_{n,s,\theta}^t(y_j)) d\widehat{\mu}_n^\theta(y_1, \dots, y_n), \quad (8.8)$$

where

$$Z_n = \int_{\mathbb{R}^n} \prod_{j=1}^n (1 - \sigma_{n,s,\theta}^t(y_j)) d\widehat{\mu}_n^\theta(y_1, \dots, y_n).$$

It was proved in [11, Section 4] (as an application of a more general result from [14]) that $K_{s,n,\theta}^t$ is the kernel of this determinantal signed measure with respect to the measure $(1 - \sigma_{n,s,\theta}^t(y)) dy$. In other words,

$$d\widehat{\nu}_{n,s,\theta}^t(y_1, \dots, y_n) = \frac{1}{n!} \det \left(K_{s,n,\theta}^t(y_j, y_k) \right)_{j,k=1}^n \prod_{j=1}^n (1 - \sigma_{n,s,\theta}^t(y_j)) dy_j, \quad (8.9)$$

and $K_{s,n,\theta}^t$ has the reproducing property

$$\int_{\mathbb{R}} K_{s,n,\theta}^t(y, u) K_{s,n,\theta}^t(u, y') (1 - \sigma_{n,s,\theta}^t(u)) du = K_{s,n,\theta}^t(y, y').$$

Let us try to get some more insight in these deformed biorthogonal measures. For $\theta = 0$, we have

$$d\hat{\nu}_{n,s,\theta}^t(y_1, \dots, y_n) := \frac{1}{Z_n} \prod_{1 \leq j < i \leq n} (y_i - y_j)^2 \prod_{j=1}^n (1 - \sigma_{n,s,\theta}(y_j)) 1_{(0,+\infty)}(y_j) e^{-4ny_j} dy_j. \quad (8.10)$$

This is a deformation of the eigenvalue distribution of a Laguerre-Wishart random matrix, and it is clearly a positive measure. For $\theta = 0$, $t = 0$, it is the eigenvalue distribution of a Laguerre-Wishart random matrix under conditioning on the event that all eigenvalues $y_1, \dots, y_n$ are smaller than s . For $\theta = 0$, $t > 0$, the deformed biorthogonal ensemble can be constructed via a procedure of marking and conditioning as done in [14]. The deformed ensemble then assigns a smaller likelihood to eigenvalue configurations with eigenvalues bigger than s , and one can think of it as a *pushed* Coulomb gas, following the terminology of [17, 35]. Both for $t = 0$ and $t > 0$, the deformed ensembles are special cases of orthogonal polynomial ensembles, whose large n asymptotic behavior is well understood. See e.g. [38] for an overview of such results. In particular, it is understood that for $t = 0$, the bulk of the eigenvalues lies for large n between 0 and s , and that the ensemble has a hard edge at 0 and another hard edge at s , if $s < 1$. Scaling limits of the eigenvalue correlation kernel $K_{s,n,\theta}^t(y, y')$ lead to the sine kernel in the bulk, and to the Bessel kernel near the hard edges. In particular, we have

$$K_{s,n,0}^0(y, y) = O(n), \quad n \rightarrow \infty, \quad (8.11)$$

for $\epsilon < y < s - \epsilon$, and

$$K_{s,n,0}^0(y, y) = O(n^2), \quad n \rightarrow \infty, \quad (8.12)$$

for $y \leq \epsilon$ and for $s - \epsilon \leq y$. This gives Ansatz 8.1 for $\theta = 0$, $t = 0$. As t increases, the hard edge at s turns into a soft edge, such that one expects that the above bounds still hold, and even that the bound near $y = s$ becomes non-optimal for bigger t .

Let us now consider the case $\theta > 0$. Then the deformed biorthogonal measure $d\hat{\nu}_{n,s,\theta}^t$ is not necessarily positive; however we still expect that the one-point function $K_{s,n,\theta}^t(y, y)$ is positive for $y > \epsilon$, θ sufficiently small, and n sufficiently large. For $t = 0$, the deformed measure $d\hat{\nu}_{n,s,\theta}^0$ is the restriction of $d\hat{\mu}_n^\theta(y_1, \dots, y_n)$ to the region $(-\infty, s)^n$, renormalized to have total mass 1. For $t > 0$, the deformed measure $d\hat{\nu}_{n,s,\theta}^t$ gives less (positive or negative) weight to configurations $(y_1, \dots, y_n)$ for which $y_j s$ lie in $(s, +\infty)$, so we can still think of it as a *pushed* measure, where weight is pushed towards $(-\infty, s)^n$. We do not expect a qualitative difference between the correlation kernel $K_{s,n,0}^t(y, y)$ and $K_{s,n,\theta}^t(y, y)$ for $\theta > 0$ sufficiently small; in particular, we expect (but cannot prove) that (8.11)–(8.12) continue to hold for $\theta > 0$ sufficiently small and for $0 < t < 1$, such that Ansatz 8.1 holds.

Unfortunately, we have no explicit formulas or tools to analyze $K_{s,n,\theta}^t(y, y)$ asymptotically as $n \rightarrow \infty$: there is no double contour integral representation like for L_n^θ , and there is no convenient RH characterization for it. We do however have access to a special case of this quantity, corresponding to $t = 0$ and $y = s$. We indeed have by Jacobi's identity that

$$\partial_s \log Q_n(s; t) = K_{s,n,\theta}^0(s, s).$$

Differentiating the result from Theorem 1.6 on the other hand, we obtain that

$$K_{s,n,\theta}^0(s, s) = \partial_s \log Q_n(s; t) = O(n^2), \quad n \rightarrow \infty,$$

which is consistent with and confirms the order predicted in Ansatz 8.1.

Comparison of Fredholm determinants. Alternatively, one may attempt to prove Ansatz 1.5 by using a comparison bound for Fredholm determinants. For trace-class operators A, B , we have [42, Chapter 5]

$$|\det(1 + A) - \det(1 + B)| \leq \|A - B\|_1 e^{\max\{\|A\|_1, \|B\|_1\} + 1}, \quad (8.13)$$

where $\|\cdot\|_1$ denotes the trace norm. Setting $B = \mathcal{L}_n^s$ with $\mathcal{L}_n^s$ given by

$$(\mathcal{L}_n^s)[f](x) = 1_{(s, +\infty)}(y) e^{-cy} \int_{\mathbb{R}} \widehat{L}_n^\theta(y', y) e^{cy'} f(y') dy',$$

like in the proof of Lemma 2.2, and A given by

$$(A)[f](y) = \sigma_{n,s,\theta}(y) e^{-cy} \int_{\mathbb{R}} \widehat{L}_n^\theta(y', y) e^{cy'} f(y') dy',$$

the left hand side of (8.13) becomes $|\widetilde{Q}_n(s) - Q_n(s)|$. It is not immediate to compute the order of the right hand side: since the operators A and B are not Hermitian, the trace norm is not necessarily equal to the trace and thus hard to compute. One may nevertheless expect that $\|A - B\|_1 = O(1/n)$ as $n \rightarrow \infty$. Unfortunately, for $0 < s < -\theta\psi(\theta)$, one also expects that $\|A\|_1, \|B\|_1 = O(n)$ as $n \rightarrow \infty$, such that the right hand side of (8.13) is not small. While the inequality (8.13) is sharp for small operators A, B , its quality becomes poor when A, B are close to being singular, which is our situation. This approach consequently does not lead to a proof nor to further support for Ansatz 1.5.

Conclusion. Let us summarize the line of arguments in support of Conjecture 1.6. Above, we explained why we strongly believe that the technical Ansatz 8.1 holds: firstly because it holds for $\theta = 0$ and there is no sign of a dramatic change of behavior for $\theta > 0$, and secondly because of the interpretation of the kernel $K_{s,n,\theta}^t(y, y')$ in terms of a deformed signed biorthogonal measure along with its interpretation as a pushed measure. Then, we proved that the technical Ansatz 8.1 implies the less technical Ansatz 1.5. Furthermore, we proved Theorem 1.6, and we proved that it implies Conjecture 1.1 under Ansatz 1.5.

Acknowledgements

The authors are grateful to Mattia Cafasso, Reda Chhaibi, Ivan Corwin, Yuchen Liao, and Jiyuan Zhang for useful discussions. They acknowledge support by *FNRS Research Project T.0028.23* and by the *Fonds Spécial de Recherche* of UCLouvain.

References

- [1] J. Baik, P. Deift, and K. Johansson, On the distribution of the length of the longest increasing subsequence of random permutations, *J. Amer. Math. Soc.* 12 (1999), 1119–1178.
- [2] G. Barraquand, I. Corwin, E. Dimitrov, Fluctuations of the log-gamma polymer free energy with general parameters and slopes, *Prob. Th. Rel. Fields* 181 (1) (2021), 113–195.
- [3] M. Bertola and M. Cafasso, The transition between the gap probabilities from the Pearcey to the Airy process - a Riemann–Hilbert approach, *Int. Math. Res. Not.*, 2012 (2012), 1519–1568.
- [4] P. Bleher and A. Its, Double scaling limit in the random matrix model: the Riemann-Hilbert approach, *Comm. Pure Appl. Math.* 56 (2003), 433–516.
- [5] A. Borodin. Biorthogonal ensembles, *Nucl. Phys. B* 536 (1998), no. 3, 704–732.

- [6] A. Borodin and I. Corwin, Macdonald processes. *Probab. Theory Related Fields* 158 (2014), no. 1, 225–400.
- [7] A. Borodin, I. Corwin and D. Remenik, Log-gamma polymer free energy fluctuations via a Fredholm determinant identity, *Comm. Math. Phys.* **324** (2013), 215–323.
- [8] A. Borodin, I. Corwin, P. Ferrari, and B. Vető, Height fluctuations for the stationary KPZ equation, *Math. Phys. Anal. Geom.* 18 (2015), 1–95.
- [9] A. Borodin and S. Péché, Airy kernel with two sets of parameters in directed percolation and random matrix theory, *J. Stat. Phys.* 132 (2008), 275–290.
- [10] M. Cafasso and T. Claeys, A Riemann–Hilbert approach to the lower tail of the KPZ equation, *Comm. Pure Appl. Math.* 75(3), 493–540 (2022).
- [11] M. Cafasso and T. Claeys, Biorthogonal measures, polymer partition functions, and random matrices, to appear in *Ann. Inst. H. Poincaré. Prob. Stat.*, arXiv:2401.10130.
- [12] C. Charlier, T. Claeys, and G. Ruzza, Uniform tail asymptotics for Airy kernel determinant solutions to KdV and for the narrow wedge solution to KPZ, *J. Funct. Anal.* 283, no. 8, article 109608 (2022).
- [13] T. Claeys, M. Girotti and D. Stivigny, Large Gap Asymptotics at the Hard Edge for Product Random Matrices and Muttalib–Borodin Ensembles, *Int. Math. Res. Not.* 2017 (2017), doi: 10.1093/imrn/rnx202, 48 pages.
- [14] T. Claeys and G. Glesner, Determinantal point processes conditioned on randomly incomplete configurations, *Ann. Inst. H. Poincaré. Prob. Stat.* 59 (2023), no. 4, 2189–2219.
- [15] T. Claeys and A. Kuijlaars, Universality of the double scaling limit in random matrix models, *Comm. Pure Appl. Math.* 59, no. 11 (2006), 1573–1603.
- [16] I. Corwin and P. Ghosal, Lower tail of the KPZ equation, *Duke Math. J.* 169, no. 7 (2020).
- [17] I. Corwin, P. Ghosal, A. Krajenbrink, P. Le Doussal, and L.C. Tsai, Coulomb-gas electrostatics controls large fluctuations of the Kardar-Parisi-Zhang equation, *Phys. Rev. Letters* 121(6), 060201 (2018).
- [18] S. Das, Y. Liao, and M. Muccicone, Lower tail large deviations of the stochastic six vertex model, arXiv:2407.08530.
- [19] P. Deift, *Orthogonal Polynomials and Random Matrices: A RiemannHilbert Approach*, Courant Lecture Notes in Mathematics 3, Amer. Math. Soc., Providence, RI, 1999.
- [20] P. Deift, A. Its, and I. Krasovsky, Asymptotics of the Airy-kernel determinant, *Comm. Math. Phys.* 278 (2008), no. 3, 643–678.
- [21] P. Deift, A. Its, and X. Zhou, A Riemann-Hilbert approach to asymptotic problems in the theory of random matrix models, and also in the theory of integrable statistical mechanics, *Ann. Math.* 146 (1997), 149–235.
- [22] 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), 1491–1552.

- [23] P. Deift, T. Kriecherbauer, K.T-R McLaughlin, S. Venakides, and X. Zhou, 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), 1335–1425.
- [24] P. Deift and X. Zhou, A steepest descent method for oscillatory Riemann-Hilbert problems. Asymptotics for the MKdV equation, *Ann. Math.* 137 (1993), 295–368.
- [25] P. Deift and X. Zhou, Asymptotics for the Painlevé II equation, *Comm. Pure Appl. Math.* 48 (1995), no. 3, 277–337.
- [26] P. Desrosiers and P.J. Forrester, Asymptotic correlations for Gaussian and Wishart matrices with external source, *Int. Math. Res. Notices* 2006 (2006), 27395.
- [27] A.B. Dieker and J. Warren, On the largest-eigenvalue process for generalized Wishart random matrices, *ALEA Lat. Am. J. Probab. Math. Stat.* 6 (2009), 369–376.
- [28] H. Flaschka and A.C. Newell, Monodromy and spectrum-preserving deformations I, *Comm. Math. Phys.* 76 (1980), 65–116.
- [29] A.S. Fokas, A.R. Its, A.A. Kapaev, and V.Yu. Novokshenov, Painlevé transcendents: the Riemann-Hilbert approach, *AMS Mathematical Surveys and Monographs* 128, 2006.
- [30] P.J. Forrester, *Log-Gases and Random Matrices*, London Mathematical Society Monographs Series (34), Princeton University Press, 2010.
- [31] N. Georgiou and T. Seppäläinen, Large deviation rate functions for the partition function in a log-gamma distributed random potential, *Ann. Probab.* 41 (2013), no. 6, 4248–4286.
- [32] T. Imamura and T. Sasamoto, Determinantal structures in the O’Connell-Yor directed random polymer model, *J. Stat. Phys.* 163 (2016), 675–713.
- [33] A. Its, A.G. Izergin, V.E. Korepin, and N.A. Slavnov, Differential equations for quantum correlation functions, In *proceedings of the Conference on Yang-Baxter Equations, Conformal Invariance and Integrability in Statistical Mechanics and Field Theory, Volume 4* (1990), 1003–1037.
- [34] K. Johansson, Shape fluctuations and random matrices, *Comm. Math. Phys.* 209 (2000), 437–476.
- [35] A. Krajenbrink and P. Le Doussal, Linear statistics and pushed Coulomb gas at the edge of β -random matrices: Four paths to large deviations, *Europhysics Letters* 125 (2019), no. 2.
- [36] A. Krajenbrink and P. Le Doussal, The weak noise theory of the O’Connell-Yor polymer as an integrable discretisation of the nonlinear Schrödinger equation, *arxiv:2307.01172*.
- [37] A. Krishnan and J. Quastel, Tracy–Widom fluctuations for perturbations of the log-gamma polymer in intermediate disorder, *Ann. Appl. Probab.* **28** (2018), 3736–3764.
- [38] A. Kuijlaars, Universality, Chapter 6 in *The Oxford Handbook on Random Matrix Theory* (G. Akemann, J. Baik, and P. Di Francesco, eds.), Oxford University Press, 2011, pp. 103–134.
- [39] P. Le Doussal, Large deviations for the Kardar–Parisi–Zhang equation from the Kadomtsev–Petviashvili equation, *J. Stat. Mech. Theory Exp.* 2020.4 (2020): 043201.
- [40] N. O’Connell and M. Yor, Brownian analogues of Burke’s theorem, *Stoch. Process. Appl.* 96 (2001), 285–304.

- [41] T. Seppäläinen, Scaling for a one-dimensional directed polymer with boundary conditions, *Ann. Probab.* 40 (2012), 19–73.
- [42] B. Simon, Trace ideals and their applications, *Mathematical Surveys and Monographs* 120, AMS (2005).
- [43] C.A. Tracy and H. Widom, Level-spacing distributions and the Airy kernel, *Comm. Math. Phys.* 159 (1994), 151–174.
- [44] L.C. Tsai, Exact lower tail large deviations of the KPZ equation, *Duke Math. J.* 171 (9) (2022), 1879–1922.
- [45] L.C. Tsai, Integrability in the weak noise theory, [arxiv:2204.00614](https://arxiv.org/abs/2204.00614).
- [46] P. Vivo, S. Majumdar, and O. Bohigas, Large deviations of the maximum eigenvalue in Wishart random matrices, *J. Phys. A: Math. Theor.* 40(16) (2007), 4317–4337.