MAXIMAL FREE ENERGY OF THE LOG-GAMMA POLYMER

GUILLAUME BARRAQUAND, IVAN CORWIN, AND EVGENI DIMITROV

ABSTRACT. We prove a phase transition for the law of large numbers and fluctuations of $F_N$, the maximum of the free energy of the log-gamma directed polymer with parameter $\theta$, maximized over all possible starting and ending points in an $N \times N$ square. In particular, we find an explicit critical value $\theta_c = 2\Psi^{-1}(0) > 0$ ($\Psi$ is the digamma function) such that:

- For $\theta < \theta_c$, $F_N + 2\Psi(\theta/2)N$ has order $N^{1/3}$ GUE Tracy-Widom fluctuations;
- For $\theta = \theta_c$, $F_N = \Theta(N^{1/3}(\log N)^{2/3})$;
- For $\theta > \theta_c$, $F_N = \Theta(\log N)$.

Using the same techniques for analyzing $F_N$, we also show that an analogous phase transition occurs in a certain free start/end-point polymer measure with inverse gamma weights. By exploiting a connection between the log-gamma polymer and a certain random operator on the honeycomb lattice, recently found by Kotowski and Virág (Comm. Math. Phys. 370, 2019), we also deduce a similar phase transition for the asymptotic behavior of the smallest positive eigenvalue of the aforementioned random operator.

CONTENTS

1. Introduction and main results 1
2. Applications 7
3. Preliminary results 13
4. Proof of Theorem 1.2 21
5. Proof of Theorem 1.3 28
6. Proof of Theorem 1.4 34
7. Proofs of Propositions 2.2, 2.3, and 2.4 35
Appendix A. Upper tail estimates 46
References 52

1. INTRODUCTION AND MAIN RESULTS

1.1. Maximal free energy of the log-gamma polymer. The log-gamma polymer is an exactly solvable model for a directed polymer on the $\mathbb{Z}^2$ lattice introduced in [Sep12]. A variety of probabilistic tools and exact formulas have been developed to study this model, allowing one to obtain very precise information about the asymptotic behavior of its free energy and the geometry of paths under the polymer measure. In this paper we consider $F_N$, the maximum of the point-to-point free energies of the log-gamma polymer as the starting and ending point vary inside an $N \times N$ square in the $\mathbb{Z}^2$ lattice. The log-gamma polymer depends on $\theta > 0$, the shape parameter of the inverse-gamma distributed weights used to define the model. We prove that the asymptotic behavior of the maximal free energy obeys a sharp transition depending on the value of $\theta$.

G. Barraquand, Laboratoire de physique de l’école normale supérieure, ENS, Université PSL, CNRS, Sorbonne Université, Université de Paris, Paris, France
I. Corwin, Department of Mathematics, Columbia University, New York, NY 10027, USA
E. Dimitrov, Department of Mathematics, University of Southern California, Los Angeles, CA 90089, USA
E-mail addresses: guillaume.barraquand@ens.fr, ivan.corwin@gmail.com, edimitro52gmail.edu.
In the remainder of this section we define the polymer model and state our main results about $F_N$. In Section 2.2 below we provide two applications of our work, which we believe to be of separate interest. In Section 2.1 we define a certain free start/end-point polymer measure in the $N \times N$ square in the $\mathbb{Z}^2$ lattice with inverse gamma weights. We call this the log-gamma polymer measure and show that it mirrors the phase transition of $F_N$; for example, we prove a sharp transition in the polymer length as one varies $\theta$. Our second application comes from a connection established in [KV19] between $F_N$ and the smallest positive eigenvalue of a certain random operator defined on a finite honeycomb lattice. Based on this connection, one can directly apply our results about $F_N$ to obtain an analogous description of this smallest positive eigenvalue. We will discuss this further in Section 2.2.

**Definition 1.1.** A continuous random variable $X$ is said to have the inverse-gamma distribution with parameter $\theta > 0$ if its density is given by

$$f_\theta(x) = \frac{1\{x > 0\}}{\Gamma(\theta)} \cdot x^{-(\theta-1)} \cdot \exp(-x^{-1}),$$

where $\Gamma$ is the Euler gamma function. We let $w = (w_{i,j} : i, j \in \mathbb{Z})$ denote the random matrix such that $w_{i,j}$ are i.i.d. with density $f_\theta$ as in (1.1). We denote by $(\Omega, \mathcal{F}, \mathbb{P})$ the probability space on which $w$ is defined. A directed lattice path is a sequence of vertices $(x_1, y_1), \ldots, (x_k, y_k) \in \mathbb{Z}^2$ such that $x_1 \leq x_2 \leq \cdots \leq x_k$, $y_1 \leq y_2 \leq \cdots \leq y_k$ and $(x_i - x_{i-1}) + (y_i - y_{i-1}) = 1$ for $i = 2, \ldots, k$. In words, a directed lattice path is an up-right path on $\mathbb{Z}^2$, which makes unit steps in the coordinate directions (see Figure 1).

![Figure 1](image)

**Figure 1.** A directed lattice path $\pi \in \Pi(m_1, n_1; m_2, n_2)$ for the log-gamma polymer.

For any $m_1, n_1, m_2, n_2 \in \mathbb{Z}$ we let $\Pi(m_1, n_1; m_2, n_2)$ denote the set of directed paths $\pi$ in $\mathbb{Z}^2$ from $(m_1, n_1)$ to $(m_2, n_2)$. Observe that $\Pi(m_1, n_1; m_2, n_2)$ is non-empty if and only if $m_1 \leq m_2$ and $n_1 \leq n_2$ (which we write as $(m_1, n_1) \leq (m_2, n_2)$). Given a directed path $\pi$ we define its weight

$$w(\pi) = \prod_{(i,j) \in \pi} w_{i,j}.$$  

For $m_1, n_1, m_2, n_2 \in \mathbb{Z}$ with $(m_1, n_1) \leq (m_2, n_2)$ we define the point-to-point partition function as

$$Z(m_1, n_1; m_2, n_2) = \sum_{\pi \in \Pi(m_1, n_1; m_2, n_2)} w(\pi).$$

For each $p, q \in \mathbb{Z}$ with $p \leq q$ we write $[p, q] = \{p, p+1, \ldots, q\}$. For each $N \in \mathbb{N}$ we define

$$F_N = \max_{(m_1, n_1) \leq (m_2, n_2)} \max_{m_1, n_1, m_2, n_2 \in [1, N]} \log Z(m_1, n_1; m_2, n_2),$$

the maximal point-to-point free energy of the log-gamma polymer in a square of side length $N$. 


In this paper we are interested in understanding the asymptotic behavior of \( F_N \) as \( N \) tends to infinity. Before we state our main result we introduce a bit of notation.

Let \( \Psi(x) \) denote the digamma function, i.e.

\[
\Psi(z) = \frac{\Gamma'(z)}{\Gamma(z)} = -\gamma_E + \sum_{n=0}^{\infty} \left( \frac{1}{n+1} - \frac{1}{n+z} \right),
\]

where \( \gamma_E \) is the Euler constant. From the formula

\[
\Psi'(z) = \sum_{n=0}^{\infty} \frac{1}{(n+z)^2},
\]

we know that \( \Psi \) is strictly increasing on \((0, \infty)\). Moreover, \( \lim_{z \to 0} \Psi(z) = -\infty \) and \( \lim_{z \to \infty} \Psi(z) = \infty \) (see [Luk69], (8), pp. 33). This means that there is a unique point \( \gamma_* \in (0, \infty) \) such that \( \Psi(\gamma_*) = 0 \), and we let \( \theta_c := 2\gamma_* \) so that

\[
(1.6) \quad \Psi(\theta/2) < 0 \text{ for } \theta \in (0, \theta_c), \quad \Psi(\theta/2) = 0 \text{ for } \theta = \theta_c, \text{ and } \Psi(\theta/2) > 0 \text{ for } \theta > \theta_c.
\]

Finally, for any \( \theta > 0 \) we define

\[
(1.7) \quad \sigma_\theta := \left[ -\Psi''(\theta/2) \right]^{1/3} = \left[ 2 \sum_{n=0}^{\infty} \frac{1}{(n+\theta/2)^3} \right]^{1/3}.
\]

With the above notation we are ready to state our main results (see also Figure 2). It should be noted that the limit theorem in the subcritical case (Theorem 1.2) is the same one that is satisfied when \( F_N \) is replaced by \( \log Z(1,1;N,N) \) (i.e., the maximal separation of endpoints). That result has been previously shown in the works [BCR13, KQ18, BCD21a] and is recalled and used here as Proposition 3.1. We also mention that while the results below are formulated for the square domain \( N \times N \), one could adapt the techniques of this paper to more general rectangular domains \( N \times M \) and obtain analogous results for those. We will not pursue this generalization in the present paper and work exclusively with square domains.

**Theorem 1.2. [Subcritical case]** Suppose that \( \theta \in (0, \theta_c) \), where \( \theta_c \) is as in (1.6). For all \( y \in \mathbb{R} \)

\[
(1.8) \quad \lim_{N \to \infty} \mathbb{P} \left( \frac{F_N + 2N\Psi(\theta/2)}{\sigma_\theta N^{1/3}} \leq y \right) = \mathbb{F}_{\text{GUE}}(y),
\]

where \( \mathbb{F}_{\text{GUE}} \) is the GUE Tracy-Widom distribution [TW94].

**Theorem 1.3. [Critical case]** Suppose that \( \theta = \theta_c \), where \( \theta_c \) is as in (1.6). There are positive constants \( c_2 > c_1 > 0 \) such that

\[
(1.9) \quad \lim_{N \to \infty} \mathbb{P} \left( c_1 N^{1/3}(\log N)^{2/3} \leq F_N \leq c_2 N^{1/3}(\log N)^{2/3} \right) = 1.
\]

**Theorem 1.4. [Supercritical case]** Suppose that \( \theta > \theta_c \), where \( \theta_c \) is as in (1.6). There exist constants \( c_2 > c_1 > 0 \) (depending on \( \theta \)) such that

\[
(1.10) \quad \lim_{N \to \infty} \mathbb{P}(c_1 \log N \leq F_N \leq c_2 \log N) = 1.
\]

[KV19] Theorem 1.1] proved (1.8) in the case that \( \theta \in (0, \theta_c/2) \). The limitation of their result to half of the subcritical range can be understood in terms of the energy / entropy balance. We call the energy of a single path the logarithm of its weight, or equivalently the sum of \( \log w_{i,j} \) along the path. The free energy \( \log Z(m_1,n_1; m_2, n_2) \) will exceed the single path energy from \((m_1,n_1)\) to \((m_2, n_2)\). This is due to entropy – the fact that there are exponentially many different paths with the same starting and ending points. The balance of energy and entropy is quite non-trivial and accounts for the complicated nature of the free energy. The range \( \theta \in (0, \theta_c/2) \) is exactly when \( \mathbb{E}[\log w_{i,j}] > 0 \), and hence the single path energy is strictly positive. To maximize the single path energy in that case, one maximizes the length of the path. Entropy only further encourages longer
path lengths and hence, for $\theta \in (0, \theta_c/2)$, both energy and entropy push towards having longer paths. The specifics of the proof of [KV19] Theorem 1.1 rely on the strict positivity $\mathbb{E}[\log w_{i,j}] > 0$ and this logic. Hence the argument does not extend to $\theta \geq \theta_c/2$.

For $\theta > \theta_c/2$, the single path energy becomes negative. However, up to $\theta_c$, the free energy remains positive since the number of paths overwhelms the average negativity of each one. Beyond $\theta_c$, the balance shifts and the entropy is not strong enough of a force to keep the free energy positive.

This represents the transition to the supercritical phase. In order to address the full range of $\theta$ (in particular, $\theta \geq \theta_c/2$) we develop a different approach than used in [KV19], one which relies on the free energy directly and not the single path energy.

The following heuristic shows how to use the free energy to argue that the critical value should be $\theta_c$ (as we show) and not $\theta_c/2$. For points $(s_1, s_2) \leq (t_1, t_2)$ in the unit square $[0,1]^2$, the free energy $N^{-1} \log Z([s_1 N], [s_2 N]; [t_1 N], [t_2 N])$ converges to an explicit function $g_\theta(s_1, s_2; t_1, t_2)$ (we do not define it here), which is determined by the law of large numbers proved in [Sep12] Theorem 2.4. One expect that $N^{-1}F_N$ behaves like the maximum of $g_\theta(s_1, s_2; t_1, t_2)$ over endpoints varying in the unit square. Due to convexity of $g_\theta$ (a general property of such limit shape functions), the maximizing endpoints will lie on a diagonal. For $(s_1, s_2) = (a, a)$ and $(t_1, t_2) = (b, b)$, $g_\theta(a, a; b, b) = -2\Psi(\theta/2)(b - a)$. If $-2\Psi(\theta/2) > 0$, this will be maximized for $a = 0$ and $b = 1$ and take value $-2\Psi(\theta/2)$; otherwise it will be maximized for any choice of $a = b$ and value $0$. One easily checks that $\theta_c$ is the value for $\theta$ which divides between these two cases. For $\theta \in (0, \theta_c)$, the maximizer will be $a = 0$ and $b = 1$ whereas for $\theta > \theta_c$, it will be $a = b$. This explains the limit for $N^{-1}F_N$. We mention here that near the end of [KV19] Section 1 there is a related (abstract) discussion for the existence of a critical value and a phase transition. This phase transition is not formulated in terms of the polymer model, but through a certain related random operator (we discuss this relation in detail in Section 2.2 below). Our results above can be viewed as a confirmation of the existence of such a phase transition with an explicit formula for the critical value where it occurs.

Let us turn now to the higher resolution information contained in our results, beyond the limit of $N^{-1}F_N$. In the subcritical regime, $\theta \in (0, \theta_c)$, we show that $F_N$ and $\log Z(1, 1; N, N)$ behave the same down to order $N^{1/3}$, and thus conclude our theorem by known asymptotic results for the later random variable [BCR13, KQ18, BCD21a]. In the supercritical regime, $\theta > \theta_c$ one might hope that the maximizing points will be order one apart, and hence the $\log(N)$ behavior for $F_N$ makes sense from standard extreme value theory. In the critical case, $\theta = \theta_c$, the $N^{1/3} \log(N)^{2/3}$ behavior is a bit more subtle and comes from a competition between many roughly independent possible maximizers on different scales. The $\log(N)^{2/3}$ factor is related to the upper tail $3/2$ exponent for the point-to-point free energy.
1.2. Ideas in the proof. In this section we explain what are the main challenges and ideas needed to overcome them in the proofs of our theorems. We will focus on the subcritical case since its proof is the most involved and makes use of all of the tools employed in the other proofs. We will also explain the key roles of our recent works [BCD21a] and [BCD21b] in the proof. These works are both based upon special integrable structure (see, e.g. [COSZ14][BC14]) enjoyed by the log-gamma polymer – the first relies on a Fredholm determinant formula for the Laplace transform of the polymer partition function while the second relies on the Gibbsian line ensemble (or equivalently, the Whittaker process) into which the polymer free energy embeds. Let us assume below that $\theta \in (0, \theta_c)$, which is the subcritical case.

The heuristic law of large numbers argument presented above suggests that the maximizing points defining $F_N$ should lie within a $o(N)$ window of $(1, 1)$ and $(N, N)$. By taking into account the KPZ scaling which describes the fluctuations around the law of large numbers, we can see that the maximizing points should really lie in windows of side length $k = N^{1/3+\epsilon}$ (for any positive $\epsilon$) touching $(1, 1)$ and $(N, N)$. This can be seen from the strict convexity of the law of large numbers which suggests that even on a more local scale, there is a linear drift which pushes the maximizer towards the extreme points $(1, 1)$ and $(N, N)$. This drift is balanced once the fluctuations of the free energy are commensurable with it, which happens in a scale of order $N^{1/3}$ from the extreme points. [BCD21a, Theorem 1.2] (quoted here as Proposition 3.1) implies the $N^{1/3}$ fluctuations (see also earlier work of [BCR13][KQ18]). In order to discount all other starting and ending points besides those in the $(k \times k)$-windows around the extreme points, we use [BCD21a, Theorem 1.7] (quoted here as Proposition 3.2) which provides a moderate deviation upper bound on the upper-tail of the free energy fluctuations. This upper bound enables us to take a union bound over all starting and ending points outside the windows.

Under suitable centering and KPZ scaling, the free energy profile for varying starting and ending points is believed to converge to the directed landscape (constructed in [DOV18]). If we had such a result at our disposal, then estimates on the local modulus of continuity for the directed landscape [DOV18, Proposition 1.6] would imply that the maximum over the $(k \times k)$-windows is, in the $N^{1/3}$ scale, the same as $\log Z(1, 1; N, N)$. Such convergence to the directed landscape has, however, only been established for zero-temperature models [DOV18, DV21].

For the log-gamma polymer we can control a much weaker type of local modulus of continuity. Namely, if we fix the starting point and then vary the ending point along a single down-right path, then we can show that the free energy profile (properly centered and scaled) looks similar (up to a Radon-Nikodym derivative) to a centered random walk with log-gamma distributed jump distribution. This result is shown in [BCD21b] as a consequence of the fact that this free energy profile can be embedded as the lowest labeled curve of a Gibbsian line ensemble (see also [COSZ14][JO20][Wu20]). Specifically, [BCD21b, Theorem 1.10], quoted here as Proposition 3.6, shows tightness of the free energy profile.

The problem for us is that both the starting and ending points vary in their respective $(k \times k)$-windows. With a bit of work, we can restrict to only consider varying the starting and ending points along the one-dimensional down-right path on the boundary of the $(k \times k)$-windows (we call these frames). But still, even if we know that for each given starting point, the free energy profile is tight as the endpoint varies, this does not imply that the tightness holds simultaneously as both vary. This is because there are order $k \gg 1$ different starting points. Without more quantitative control on the modulus of continuity, such as moment or tail bounds, this approach is not feasible. The proof of [CGH21, Theorem 1.3] provides a route (in the case of the KPZ equation) for using the Brownian Gibbs property to establish such moment / tail bounds on the modulus of continuity. Almost all of the pieces used there are present for the log-gamma polymer, except for a lower tail bound on the free energy fluctuations. For the KPZ equation, that piece is supplied by [CG20]. Unfortunately, such a lower tail bound for the log-gamma polymer is currently unavailable (and the approach of [CG20] does not readily apply). As such, we do not currently know how to establish...
modulus of continuity bounds when we simultaneously vary the starting and ending points. (The key identity [DOV18, Proposition 4.1] in the construction of the directed landscape in [DOV18] is generalized in [Cor21] to the log-gamma polymer – a glimmer of hope that the methods of [DOV18] may eventually be lifted to polymers.)

So, the remaining challenge in our work is how to only use the one-free endpoint tightness result of Proposition 3.6 to deduce that varying both starting and ending points along the frames of the $(k \times k)$-windows cannot result in $F_N$ exceeding log $Z(1,1;N,N)$ on the $N^{1/3}$ scale. This is achieved by expanding to a larger window and considering log $Z(-K,-K; N+K, N+K)$ for $K$ such that $K \ll N$ and $K^{2/3} \gg k$. We can lower bound $Z(-K,-K; N+K, N+K)$ by a summation indexed by points on the two frames, with summands given by the product of three partition functions – one from $(-K,-K)$ to the point on the frame near $(1,1)$, one from that point to the point on the frame near $(N,N)$, and one from that frame point to $(N+K, N+K)$. Since $(-K,-K)$ and $(N+K, N+K)$ are fixed, the two outer partition functions in this summand involve a fixed starting point and varying ending point. Thus, we can use Proposition 3.6 to control how much these can vary as the points on the frame vary. Here we use that $k \ll K^{2/3}$ which means that on the frame we are zooming into the local modulus of continuity is as discussed above. Using this, we can reverse the inequality to upper bound the maximal frame-to-frame partition function by $Z(-K,-K; N+K, N+K)$ minus well-controlled point-to-frame partition functions. Since $K \ll N$, the contribution from the point-to-frame partition functions ends up being negligible (the free energy fluctuations are order $K^{1/3} \ll N^{1/3}$). Similarly, since $K \ll N$, the free energy fluctuations of log $Z(-K,-K; N+K, N+K)$ on the scale $N^{1/3}$ have the same law as those of $Z(1,1;N,N)$. This completes the sketch of the argument used to prove Theorem 1.2.

1.3. Further questions. We have not attempted to optimize the values of the constants $c_1, c_2$ in Theorems 1.3 and 1.4 or determine the scale and distribution of fluctuations of $F_N$. For instance, our results do not rule out the possibility that $F_N$ actually fluctuates on a smaller scale than the window specified by $c_1$ and $c_2$. We can easily deduce some numerical bounds on $c_1$ and $c_2$ from our proofs. In the critical case, we show (see Section 5.1) that one may choose $c_1 = 20^{-1} \sigma_\theta$, where $\sigma_\theta$ is as in (1.7). The value of $c_2$ that we may choose depends on the value of the constant $D_2$ in Lemma 5.11 for which we do not have a numerical estimate. In the supercritical case, the proof of Lemma 6.1 shows that one may take any $c_1 < 2/\theta$. In the case $\theta > 3$ (note that $\theta_c < 3$), we may choose any constant $c_2 > 2$ (see Remark 5.3 for more details). When $\theta_c < \theta \leq 3$ the constant $c_2$ depends again on the value of the constant $D_2$ in Lemma 3.11 as in the critical case.

Another natural problem concerns the location of the maximizing starting and ending points. While the proof of Theorem 1.2 tells us that the maximum free energy is attained for very long polymers with starting and ending points located near the corners of the $N \times N$ square (see for instance Lemma 4.2 below), the location of maximizing starting and ending points in the critical and supercritical is less clear. In the supercritical case $\theta > \theta_c$ we expect that the maximizing starting and ending points are close to each other. Note that, even for a large (but fixed) $\theta$, the probability that the starting point and ending point of the maximal partition function are not the same does not go to zero as $N$ goes to infinity. Indeed, let us call $v \in [1,N]^2$ the vertex where the maximal weight is attained. Consider the event that on one of the vertices adjacent to $v$, one of the weights is larger than $1$, and call $w$ this weight. The probability of this event does not vanish as $N$ goes to infinity. Moreover, on this event, we have $F_N \geq \log w^\max + \log w$ and the maximal path hence cannot be reduced to a single point.

Outline of the paper. In Section 2 we introduce the log-gamma polymer measure and the random operator from [KV19], and state our results about these models. In Section 3 we recall some preliminary results from our recent papers [BCD21a] and [BCD21b], and establish some estimates on the distribution of log $Z(1,1;M,N)$, in particular the important tail estimate from Lemma 3.11 that we will use in the following sections. Section 4 is devoted to the proof of Theorem 1.2.
(subcritical case). Section 5 is devoted to the proof of Theorem 1.3 (the critical case) and Section 6 is devoted to the proof of Theorem 1.3 (supercritical case). Section 7 contains the proofs of our results about the log-gamma polymer measure from Section 2. In Appendix A, we prove Proposition 6.3, an estimate for the speed of convergence of \( \log Z(1,1; M, N) \) to the Tracy-Widom distribution that we use in Section 5.

Acknowledgments. We thank Marcin Kotowski and Bálint Virág for useful conversations about their work [KV19]. G.B. also thanks Herbert Spohn for illuminating explanations about the relations between random operators and directed polymers. I.C. is partially supported by the NSF grants DMS:1811143 and DMS:1664650 as well as a Packard Foundation Fellowship for Science and Engineering. G.B. was partially supported by NSF grant DMS:1664650 as well. E.D. is partially supported by the Minerva Foundation Fellowship.

2. Applications

In Section 2.1, we consider a polymer measure that naturally arises in the log-gamma polymer model and we investigate the asymptotic behavior of paths sampled according to this measure. Specifically, we show that as one varies the \( \theta \) parameter of the log-gamma polymer the polymer measure behaves quite differently depending on whether \( \theta \in (0, \theta_c) \), \( \theta = \theta_c \), or \( \theta \in (\theta_c, \infty) \), where \( \theta_c \) is as in (1.6). In Section 2.2, we recall the random operator from [KV19] and show how our results from Section 1.1 can be directly applied to study the asymptotic behavior of the smallest positive eigenvalue of this operator. Throughout this section, we use freely notation from Section 1.1.

2.1. The log-gamma polymer measure. We proceed to define our main object of study in this section.

**Definition 2.1.** Fix \( \theta > 0 \), \( N \in \mathbb{N} \) and continue with the same notation as in Definition 1.1. Denote \( \Pi_N = \bigcup_{m_1,n_1,m_2,n_2 \in [1,N]} \Pi(m_1,n_1;m_2,n_2) \), which is the collection of all directed lattice paths, which start and end in the square \([1,N] \times [1,N]\). We define the log-gamma polymer measure to be the (depending on \( w \) and hence random) probability measure on \( \Pi_N \) defined by

\[
\mathbb{P}_N^w(\pi) = (Z_N^w)^{-1} \cdot w(\pi), \quad \text{where} \quad Z_N^w = \sum_{\pi \in \Pi_N} w(\pi) = \sum_{(m_1,n_1) \leq (m_2,n_2)} Z(m_1,n_1;m_2,n_2).
\]

By possibly enlarging the probability space \((\Omega, \mathcal{F}, \mathbb{P})\), we may assume that in addition to \( w \) it supports, for each \( N \in \mathbb{N} \), a random \( \pi_N \) in \( \Pi_N \), whose conditional distribution given \( w \) is \( \mathbb{P}_N^w(\cdot) \). We continue to call this new space \((\Omega, \mathcal{F}, \mathbb{P})\).

If \( \pi \) is given by the sequence \((x_1, y_1), \ldots, (x_k, y_k)\) as in Definition 1.1, we denote its starting point by \((X^s(\pi), Y^s(\pi)) = (x_1, y_1)\) and its ending point by \((X^e(\pi), Y^e(\pi)) = (x_k, y_k)\). We also denote the length of the path by \( L(\pi) = k - 1 = (x_k - x_1) + (y_k - y_1) \). Finally, we define the slope of the path by \( S(\pi) = \frac{y_k - y_1 + 1}{x_k - x_1 + 1} \in (0, \infty) \). Using the random up-right path \( \pi_N \), we can now define the following random variables on \((\Omega, \mathcal{F}, \mathbb{P})\):

\[
X_N^s = X^s(\pi_N), \quad Y_N^s = Y^s(\pi_N), \quad X_N^e = X^e(\pi_N), \quad Y_N^e = Y^e(\pi_N), \quad L_N = L(\pi_N), \quad S_N = S(\pi_N).
\]

In view of Definition 1.1, we have that \( X_N^s, Y_N^s, X_N^e, Y_N^e \) are random variables in \([1,N]\), \( L_N \) is a random variable in \([0, 2N - 2]\) and \( S_N \) is a random variable in \([1/N, N]\).

We are interested in describing the asymptotic behavior of the variables in (2.2) as \( N \to \infty \). We summarize our results in the next sequence of propositions, whose proofs are given in Section 7.

**Proposition 2.2.** [Subcritical case] Suppose that \( \theta \in (0, \theta_c) \), where \( \theta_c \) is as in (1.6). For any \( \epsilon > 0 \) we have that the following sequences

\[
\frac{X_N^s}{N^{1/3+\epsilon}}, \quad \frac{Y_N^s}{N^{1/3+\epsilon}}, \quad \frac{X_N^e - N}{N^{1/3+\epsilon}}, \quad \frac{Y_N^e - N}{N^{1/3+\epsilon}}, \quad \frac{L_N - 2N}{N^{1/3+\epsilon}}, \quad \frac{N(S_N - 1)}{N^{1/3+\epsilon}}.
\]
all converge to 0 in probability as $N \to \infty$.

**Proposition 2.3. [Critical case]** Suppose that $\theta = \theta_c$, where $\theta_c$ is as in (1.0). There exists a constant $c_1 \in (0, 1)$ such that
\begin{equation}
\lim_{N \to \infty} P(c_1 N \leq L_N \leq 2N) = 1.
\end{equation}
In addition, for each $\epsilon > 0$ we have that $\frac{N(S_N - 1)}{N^{2/3 + \epsilon}}$ converges to 0 in probability as $N \to \infty$.

**Proposition 2.4. [Supercritical case]** Suppose that $\theta \in (\theta_c, \infty)$, where $\theta_c$ is as in (1.0). There exists a constant $c_2 > 0$ (depending on $\theta$) such that
\begin{equation}
\lim_{N \to \infty} P(L_N \leq c_2 \log N) = 1.
\end{equation}
If $\theta \in (3, \infty)$, then the sequence $L_N$ is tight. In addition, for any $\theta \in (\theta_c, \infty)$ the sequence of random vectors $(N^{-1}X_N^\theta, N^{-1}Y_N^\theta)$ converges weakly to the uniform measure on $(0, 1)^2$.

**Remark 2.5.** From [AS65] page 259 we have that $\theta_c \approx 2.92326$. Consequently, Proposition 2.4 states that for $\theta \in (\theta_c, 3]$ the path length $L_N$ is with high probability $O(\log N)$, while for $\theta > 3$ it is with high probability $O(1)$. It would be very interesting to see if there is a range of parameters in $(\theta_c, 3]$ where the length $L_N$ is indeed of logarithmic order, or if for any $\theta > \theta_c$ the length is of unit order. At this point our methods appear to be insufficient to address this question as they can only provide an upper bound for the range $\theta \in (\theta_c, 3]$.

**Remark 2.6.** A great variety of polymer models, directed or not, have been considered in polymer physics [?]. Directed polymer models with variable length may arise for instance in the context of polymer melts [?]. The log-gamma polymer measure that we consider is hardly a model of polymer melt, since we are considering a single polymer chain. However, the weights in our model can be thought of as the exponential of some monomer-monomer bond energy, and we prove the existence of a sharp transition in polymer length, as the temperature varies, in presence of quenched disorder.

### 2.2. A random operator on the honeycomb lattice

Part of our motivation for studying the asymptotic behavior of $F_N$ comes from a random operator, introduced by Kotowski and Virág in [KV19]. The authors showed that the smallest positive eigenvalue of this operator is related to the random variable $F_N$ for the log-gamma polymer model, and this allowed them to deduce Tracy-Widom fluctuations for the logarithm of the smallest positive eigenvalue in a certain range of parameters. Before explaining the exact relation between the random operator and the log-gamma polymer (see Section 2.2.1), let us first define the model and state our results.

**Definition 2.7.** For all $N \geq 1$, we define a planar graph $G_N$, depicted in Figure 3, as follows. Consider a tiling of the Euclidean plane by regular hexagons of sidelength 1. To identify this tiling among its translations and rotations, assume that three hexagons meet at the origin, and their edges form angles $0, 2\pi/3, 4\pi/3$. A graph $G = (V, E)$ is associated to this tiling, where the vertex set $V$ consists of all points in the plane where three hexagons meet, and the edge set $E$ consists of pairs of vertices connected by the common edge of two adjacent hexagons. We define $G_N = (V_N, E_N)$ as the subgraph of $G$ consisting of vertices (we identify the plane with $\mathbb{C}$)

$$x_{xy} := x(1 + e^{-1\pi/3}) + y(1 + e^{\pi/3}) \quad \text{and} \quad \Omega_{x,y} := 1 + x_{xy},$$

for all $x, y \in [1, N]$. As can be seen in Figure 3, $G_N$ is a portion of a honeycomb lattice. The edge set $E_N$ consists of edges in $E$ connecting two vertices in $V_N$.

The graph $G_N$ is bipartite, as depicted in Figure 3 where black vertices (of the form $x_{xy}$) are connected to white vertices (of the form $\Omega_{x,y}$). We associate weights to the edges of $G_N$, in such a way that for each (blue) horizontal edge $e = e_{x,y}$ in $E_N$ connecting $x_{xy}$ to $\Omega_{x,y}$, its weight $w(e) = 1/w_{x,y}$ is a gamma random variable with shape parameter $\theta$, that is with density
Figure 3. The planar graph $G_N$ for $N = 4$.

\[
\frac{1}{\Gamma(\theta)} x^{\theta - 1} e^{-x} 1_{x > 0}, \text{ and all gamma random variables are independent. All other edges (the red ones) have weight 1.}
\]

Because $G_N$ is bipartite, its weights adjacency matrix $A_N$ can be written in the form

\[
A_N = \begin{pmatrix} 0 & \tilde{A} \\ \tilde{A}^t & 0 \end{pmatrix}
\]

where $\tilde{A}$ is an $N^2$ by $N^2$ matrix whose rows are indexed by black vertices and columns are indexed by white vertices. We may consider that rows and columns are both indexed by points $(x, y) \in [1, N]^2$. The matrix $A_N$ may then be thought of as an operator acting on functions $[1, N]^2 \to \mathbb{C}^2$. The matrix $\tilde{A}$ acts on functions $f : [1, N]^2 \to \mathbb{C}$ as

\[
\tilde{A} = \nabla_x - x + \nabla_y + w
\]

where $w$ acts by pointwise multiplication $w f(x, y) = \frac{1}{w_{x,y}} f(x, y)$, and the $\nabla^\pm$ are shift operators

\[
\nabla^\pm_x f(x, y) = \begin{cases} f(x - 1, y) & \text{for } x > 0, \\ 0 & \text{for } x = 0, \end{cases}
\]

and

\[
\nabla^\pm_y f(x, y) = \begin{cases} f(x, y - 1) & \text{for } y > 0, \\ 0 & \text{for } y = 0. \end{cases}
\]

$A_N$ is almost surely invertible [KV19, Lemma 2.1] with spectrum consisting of pairs of eigenvalues $\pm \lambda_1, \ldots, \pm \lambda_{N^2}$, $0 < \lambda_1 < \cdots < \lambda_{N^2}$. It was shown in [KV19] that when $\theta \in (0, \theta_c/2)$, the smallest positive eigenvalue $\lambda_1$ is exponentially small, with GUE Tracy-Widom distributed fluctuations. Our results above about the asymptotics of $F_N$ allow us to extend this result and prove the following immediate corollaries (of our results from Section 1.1) about the asymptotic behavior of $A_N$. For $\theta \in (0, \theta_c/2)$, the result of Corollary 2.8 below is already present in [KV19 Theorem 1.1]. Our results provide a complete phase diagram and show that $\theta_c$ is the critical point dividing exponential decay from polynomial decay of $\lambda_1$ with $N$.

Corollary 2.8. [Subcritical case] Suppose that $\theta \in (0, \theta_c)$, where $\theta_c$ is as in (1.6). For all $r \in \mathbb{R}$,

\[
\lim_{N \to \infty} \mathbb{P} \left( \frac{-\log \lambda_1 + 2N \Psi(\theta/2)}{\sigma_N N^{1/3}} \leq r \right) = F_{\text{GUE}}(r).
\]

Corollary 2.9. [Critical case] Suppose that $\theta = \theta_c$, where $\theta_c$ is as in (1.6). With the same constants $c_2 > c_1 > 0$ as in Theorem 1.3, we have

\[
\lim_{N \to \infty} \mathbb{P} \left( c_1 N^{1/3} (\log N)^{2/3} \leq -\log \lambda_1 \leq c_2 N^{1/3} (\log N)^{2/3} + 4 \log N \right) = 1.
\]
Corollary 2.10. [Supercritical case] Suppose that \( \theta > \theta_c \), where \( \theta_c \) is as in (1.6). With the same constants \( c_2 > c_1 > 0 \) as in Theorem 1.4 we have

\[
\lim_{N \to \infty} \mathbb{P} \left( c_1 \log N \leq -\log \lambda_1 \leq (4 + c_2) \log N \right) = 1.
\]

The \( 4 \log(N) \) which arises in the critical and supercritical case follows from Proposition 2.12 which itself comes from the \( k = 1 \) case of (2.10). For the subcritical case, this factor of \( 4 \log(N) \) is still present but can be absorbed into the \( N^{1/3} \) fluctuation scale. In Section 2.2.1 we explain how these corollaries follow from Theorems 1.2, 1.3 and 1.4.

2.2.1. Relation between the log-gamma polymer and the random operator. Corollaries 2.8, 2.9 and 2.10 are deduced respectively from Theorems 1.2, 1.3 and 1.4 using an exact algebraic relation between the log-gamma polymer from Definition 1.1 and the random operator \( A_N \) from Definition 2.7. This relation was discovered in [KV19] and is stated in the following proposition.

Proposition 2.11 (Equivalent to [KV19 Proposition 2.3]). The inverse of the matrix \( \tilde{A} \), introduced in (2.6), satisfies for all \( S, T \in \llbracket 1, N \rrbracket^2 \),

\[
(\tilde{A}^{-1})_{T,S} = Z(S; T)(-1)^{|T-S|},
\]

where for \( S = (s_1, s_2) \), \( T = (t_1, t_2) \) we define \( |T - S|_1 = |t_1 - s_1| + |t_2 - s_2| \) and we recall from (1.3) that \( Z(S; T) = Z(s_1, s_2; t_1, t_2) \) is the partition function of the log-gamma polymer with weights \( w_{i,j} = 1/w(e_{i,j}) \) as defined in Definition 1.7.

The statement of Proposition 2.11 differs slightly from [KV19] Proposition 2.3. In particular, [KV19] considers a more general directed polymer model (see Section 2.2.3 below), which in the case of present interest would be a variant of the log-gamma polymer model with edges weighted by \(-1\). This is why we included the factor \((-1)^{|T-S|_1}\) in (2.9). Although the proof of Proposition 2.11 can be found in [KV19], let us explain where it comes from. Using (2.7), one can see that (2.9) is equivalent to the recurrence relation of the partition function, which reads

\[
\frac{1}{w_{x,y}} Z(S; T) = \nabla_x Z(S; T) + \nabla_y Z(S; T), \quad \text{for } S \leq T \text{ with } S \neq T
\]

where \( \nabla_x, \nabla_y \) act on the \( T \) variable and have been defined in (2.8).

Proposition 2.11 allows us to estimate the smallest positive eigenvalue of the adjacency matrix \( A_N \) in terms of the partition function of the log-gamma polymer.

Proposition 2.12 (Special case of [KV19 Theorem 2.5]). Fix \( N \in \mathbb{N} \) and recall that \( \lambda_1 \) is the smallest positive eigenvalue of \( A_N \). We have \( F_N \leq -\log \lambda_1 \leq F_N + 4 \log N \).

Note that Proposition 2.12 may appear different from the statement of [KV19] Theorem 2.5. First, [KV19] Theorem 2.5 is more general than Proposition 2.12 (see more details in Section 2.2.3 below), and second, it provides an estimate for the largest singular value \( \sigma_1 \) of \( \tilde{A}^{-1} \) but this singular value is exactly equal to \( 1/\lambda_1 \) hence the statement of Proposition 2.12.

In the remainder of this section we use Proposition 2.12 to deduce Corollaries 2.8, 2.9 and 2.10 from Theorems 1.2, 1.3 and 1.4 respectively.

Proof of Corollary 2.8. From Proposition 2.12 we have for each \( r \in \mathbb{R} \)

\[
\mathbb{P} \left( \frac{F_N + 2N \Psi(\theta/2) + 4 \log N}{\sigma_0 N^{1/3}} \leq r \right) \leq \mathbb{P} \left( \frac{-\log \lambda_1 + 2N \Psi(\theta/2)}{\sigma_0 N^{1/3}} \leq r \right) \leq \mathbb{P} \left( \frac{F_N + 2N \Psi(\theta/2)}{\sigma_0 N^{1/3}} \leq r \right).
\]

From Theorem 1.2 and the continuity of \( F_{\text{GUE}}(x) \) we see that the first and third term above both converge to \( F_{\text{GUE}}(r) \) as \( N \to \infty \). By the squeeze theorem the same is true for the middle term. \( \square \)
Proof of Corollary 2.9. From Proposition 2.12 we have the inclusion of events

\[
\left\{ c_1 N^{1/3} (\log N)^{2/3} \leq F_N \leq c_2 N^{1/3} (\log N)^{2/3} \right\} \subseteq \\
\left\{ c_1 N^{1/3} (\log N)^{2/3} \leq -\log \lambda_1 \leq c_2 N^{1/3} (\log N)^{2/3} + 4 \log N \right\}.
\]

From Theorem 1.3 the probabilities of the events on the first line converge to 1 as \( N \to \infty \), and so the same is true for the probabilities of the events of the second line. \( \square \)

Proof of Corollary 2.10. From Proposition 2.12 we have the inclusion of events

\[
\left\{ c_1 \log N \leq F_N \leq c_2 \log N \right\} \subseteq \left\{ c_1 \log N \leq -\log \lambda_1 \leq (4 + c_2) \log N \right\}.
\]

From Theorem 1.4 the probabilities of the left events converge to 1 as \( N \to \infty \), and so the same is true for the probabilities of the events on the right. \( \square \)

2.2.2. Possible extensions. Instead of assuming that the weights of the log-gamma polymer (resp. of the operator \( A_N \)) are distributed as inverse-gamma random variables (resp. gamma random variables) of shape parameter \( \theta \), we may assume that they are distributed as inverse-gamma random variables (resp. gamma random variables) with shape parameter \( \theta \) and scale parameter \( \beta \). This extension was considered in [KV19]. The hypothesis used in [KV19, Theorem 1.1] remains that \( \mathbb{E} [\log w_{i,j}] > 0 \), so that the condition on the parameter \( \theta \) becomes \( \psi(\theta) - \log \beta < 0 \). In presence of such a scale parameter \( \beta \), we expect that the optimal condition for the statements of Theorem 1.2 and Corollary 2.8 to be valid is \( \psi(\theta/2) - \log \beta < 0 \), that is when \( \lim_{N \to \infty} \frac{\log Z(1,1;N,N)}{N} > 0 \).

Figure 4. We associate to the planar graph \( G_N \) (left) a subset of the \( \mathbb{Z}^2 \) lattice (right) as follows. The horizontal blue edges of \( G_N \) are mapped to vertices in \( \mathbb{Z}^2 \), while the red oblique edges of \( G_N \), which always join two horizontal blue edges, are mapped to the edges of \( \mathbb{Z}^2 \). For arbitrary weights, the adjacency matrix on \( G_N \) is related to a polymer model on \( \mathbb{Z}^2 \), see Section 2.2.2 and [KV19].

The statement given in [KV19, Proposition 2.3] is more general than Proposition 2.11. For any choice of weights \( w(e) \) on the edges of \( G_N \), [KV19, Proposition 2.3] shows that \( A^{-1}_{T,S} = Z(S;T)(-1)^{||T-S||_1} \), where the partition function \( Z(S;T) \) corresponds to a directed polymer model on the \( \mathbb{Z}^2 \) lattice with arbitrary weights on both vertices and edges. The correspondence between the edges of \( G_N \) and the vertices and edges of the \( \mathbb{Z}^2 \) lattice (on which the polymer model is defined) is illustrated in Figure 4. Horizontal edges of \( G_N \) are mapped to vertices in \( \mathbb{Z}^2 \) while oblique edges in \( G_N \) (the red ones) are mapped to edges of \( \mathbb{Z}^2 \). The weights on vertices of \( \mathbb{Z}^2 \) are the inverses of weights on horizontal edges in \( G_N \), as in the statement of Proposition 2.11. However, the polymer weights on edges of \( \mathbb{Z}^2 \) are the opposite of weights on the red edges of \( G_N \). The result can be further generalized to more general bipartite graphs \( G_N \) endowed with a perfect matching, when the matrix \( \tilde{A} \) defines a directed graph which must contain no cycles except for loops at every vertex (see [KV19] for more details).
This correspondence between directed polymer models and eigenvalues of random operators on the honeycomb lattice could also be applicable to other integrable polymer models such as the strict-weak polymer \cite{OOT15,CSS15} and the beta polymer \cite{BC17}, as this was already mentioned in \cite{KV19}. In the case of the strict weak polymer, only half of the oblique vertices in $G_N$ would be weighted by independent gamma random variables — say up-right red edges linking $O_{x,y}$ to $\bullet_{x,y+1}$ — while all other edges would have weight one. In the case of the beta polymer, horizontal edges would have weight 1 while the oblique edges would carry beta distributed weights: up-right red edges linking $O_{x,y}$ to $\bullet_{x,y+1}$ would carry a beta distributed weight $B_{x,y}$ and down-right red edges linking $O_{x,y}$ to $\bullet_{x+1,y}$ would carry the weight $1 - B_{x,y}$, the collection of $B_{x,y}$ being independent.

2.2.3. Other eigenvalues. The statement given in \cite[Theorem 2.5]{KV19} is more general than Proposition 2.12. It estimates the product of the first $k$ largest singular values of $\bar{A}^{-1}$ (i.e. the first $k$ smallest positive eigenvalues of $A_N$) using partition functions for $k$ non-intersecting paths in the log-gamma polymer maximized over $k$-tuples of starting and ending points. More precisely we have from \cite[Theorem 2.5]{KV19} that for all $1 \leq k \leq N^2$,

\begin{equation}
\max_{\mathbf{S},\mathbf{T}} |Z(\mathbf{S}; \mathbf{T})| \leq \prod_{i=1}^{k} \sigma_i \left( A^{-1} \right) = \prod_{i=1}^{k} \frac{1 - \sigma_i}{\sigma_i} \leq \left( \frac{N^2}{k} \right)^2 \max_{\mathbf{S},\mathbf{T}} |Z(\mathbf{S}; \mathbf{T})|,
\end{equation}

where now the maximum runs over pairs of $k$-tuples of distinct points $\mathbf{S} = (S_1, \ldots, S_k), \mathbf{T} = (T_1, \ldots, T_k)$ where $S_i, T_j \in [1, N]^2$, $Z(\mathbf{S}; \mathbf{T})$ is the sum of weights of $k$-tuples of non-intersecting paths with starting and ending points $S_i, T_i$, and $\sigma_i(\cdot)$ denote the singular values of a matrix in decreasing order. Note that we consider absolute values of partition functions $|Z(\mathbf{S}; \mathbf{T})|$ in (2.10) because the result hold for arbitrary – not necessarily positive – weights, as in Section 2.2.2. We refer to \cite{KV19} for details. The case $k = 1$ of (2.10) corresponds to Proposition 2.12.

When $\theta = 0$, that is, if we replace all gamma distributed weights by 0, then 0 is an eigenvalue of the matrix $A_N$ with multiplicity $2N$. Indeed, $A_N$ can be decomposed into $2N$ blocks corresponding to connected components in the graph $G_N$ and each of these blocks has corank 1, so that 0 is an eigenvalue of $A_N$ with multiplicity $2N$. As $\theta$ becomes positive, but smaller than $\theta_c$, we expect that some of the eigenvalues that are the closest to zero will remain exponentially small. The precise asymptotics of $\lambda_k$, for a fixed $k$, would be controlled by the asymptotics of the maximal partition function for $k$ non-intersecting paths in the log-gamma polymer model, which we expect to grow linearly as $-2\Psi(\theta/2)N$ with fluctuations distributed as the $k$-th eigenvalue in the Airy point process. When $\theta$ becomes larger than $\theta_c$, however, the smallest eigenvalues are no longer exponentially small, as proved in Corollary 2.10. It would be interesting to determine more precise asymptotics when $\theta = \theta_c$ in order to better understand the nature of this transition.

2.2.4. Eigenvectors and relation to the literature on random operators. In the context of random operators, an important question is the localization or delocalization of eigenvectors. In our case, the eigenvectors of $A_N$ are related to the singular vectors of $\bar{A}$, which are in turn related to the singular value decomposition of the matrix of partition functions $(Z(\mathbf{S}; \mathbf{T}))_{\mathbf{S}, \mathbf{T} \in [1, N]^2}$ via Proposition 2.11. It would be very interesting to determine if not only the behavior of the smallest positive eigenvalues obeys a transition at $\theta_c$ but also the localization properties of the associated eigenvectors. However, we emphasize that in the model from Definition 2.7, disorder is carried by edges of the lattice and not by vertices, so we may have a different behavior than for the more commonly studied models of the type $\Delta + V$ where $\Delta$ is a discrete or continuous Laplacian and $V$ is a noise potential on vertices. The latter is often referred to as Anderson model or random Schrödinger operator.

For operators of the type Laplacian plus noise on a two-dimensional lattice, connections to directed polymers models have been established numerically in the physics literature, see \cite{PSO05, SOP07, SLO15} and the references therein. The physics literature predicts that matrix elements of the resolvent of the operator, which are considered to be a good probe for the conductance of
the physical model described by the operator, can be expanded as a series of products [And58], see also the lecture notes [Hum08] or the book [AW15, Chapter 6]. When the disorder is strong enough, in the localized phase, this series is expected to be well-approximated by the partition function of a directed polymer model (though the weights in the auxiliary polymer model are not necessarily positive unlike usual polymer models), and hence have the same fluctuations as for models in the Kardar-Parisi-Zhang universality class, see e.g. the physics lecture notes [ST17].

3. Preliminary results

We present various technical results to be used in the proofs of Theorems 1.2, 1.3 and 1.4.

3.1. Summary of results from [BCD21a] and [BCD21b]. We recall some key results from [BCD21a] and [BCD21b]. Define the function

$$g_\theta(z) = \frac{\sum_{n=0}^{\infty} \frac{1}{(n+\theta-z)^2}}{\sum_{n=0}^{\infty} \frac{1}{(n+z)^2}},$$

and observe that $g_\theta$ is a smooth, increasing bijection from $(0, \theta)$ to $(0, \infty)$. The inverse function $g_\theta^{-1} : (0, \infty) \rightarrow (0, \theta)$ is likewise a strictly increasing smooth bijection. For $x \in (0, \infty)$, define

$$h_\theta(x) = x \cdot \Psi(g_\theta^{-1}(x)) + \Psi(\theta - g_\theta^{-1}(x)),$$

with $\Psi(x)$ from (3.5). Clearly $h_\theta$ is smooth on $(0, \infty)$. For $x \in (0, \infty)$, define

$$\sigma_\theta(x) := \left( \sum_{n=0}^{\infty} \frac{x}{(n + g_\theta^{-1}(x))^3} + \sum_{n=0}^{\infty} \frac{1}{(n + \theta - g_\theta^{-1}(x))^3} \right)^{1/3}.$$

Finally, for all $N, M \geq 1$, we define the rescaled free energy

$$\mathcal{F}(M, N) = \frac{\log Z(1, 1; M, N) + M h_\theta(N/M)}{M^{1/3} \sigma_\theta(N/M)}.$$

With the above notation in place we can state the two results we require from [BCD21a].

Proposition 3.1. [BCD21a, Theorem 1.2] Let $\theta > 0$, $\delta \in (0, 1)$ and $\mathcal{F}(M, N)$ be as in (3.4). Assume that we let $M, N$ go to infinity in such a way that the sequence $N/M \in [\delta, \delta^{-1}]$ for all sufficiently large $M$. Then for all $y \in \mathbb{R}$ we have

$$\lim_{M, N \to \infty} \mathbb{P}(\mathcal{F}(M, N) \leq y) = F_{\text{GUE}}(y).$$

Proposition 3.2. [BCD21a, Theorem 1.7] Let $\theta > 0$, $\delta \in (0, 1)$ and $\mathcal{F}(M, N)$ be as in (3.4). There exist constants $C_1, C_2, c_1, c_2 > 0$ (all depending on $\theta, \delta$) such that the following holds for all $M, N \in \mathbb{N}$ with $N/M \in [\delta, \delta^{-1}]$ and $x \geq 0$

$$\mathbb{P}(\mathcal{F}(M, N) \geq x) \leq C_1 e^{-c_1 M} + C_2 e^{-c_2 x^{3/2}}.$$

In addition to Propositions 3.1 and 3.2 we will require the following quantitative and uniform comparison of the upper tails of $\mathcal{F}(M, N)$ and $F_{\text{GUE}}$. Its proof is postponed until Appendix A.1 and relies on exact formulas from [BCD21a].

Proposition 3.3. Let $\theta > 0$, $\delta \in (0, 1]$, $\epsilon_0 \in (0, 1/3)$. There exist positive constants $M_0, C_0$ (depending on $\theta, \delta$) such that for all $M, N \in \mathbb{N}$ with $M \geq N \geq \delta M$, $M \geq M_0$

$$\sup_{x \in [1, \infty)} \left| \mathbb{P}(\mathcal{F}(M, N) \leq x) - F_{\text{GUE}}(x) \right| \leq C_0 M^{-\epsilon_0}.$$
Definition 3.4. Fix any $T > 0$, $\theta > 0$, $r \in (0, \infty)$, and assume $N$ is large enough so that $rN \geq 2 + TN^{2/3}$. For each $x \in [-T - N^{-2/3}, T + N^{-2/3}]$ with $xN^{2/3} \in \mathbb{Z}$, define $n = \lfloor rN \rfloor + xN^{2/3}$ and

$$f_N^L(x) = N^{-1/3}\left(\log Z(1, 1; n, N) + h_\theta(r)N + h'_\theta(r)xN^{2/3}\right),$$

and then extend $f_N^L$ to $x \in [-T, T]$, by linear interpolation. Clearly $f_N^L$ is a random continuous curve in $(C[-T, T], C)$ – the space of continuous functions on $[-T, T]$ with the uniform topology and Borel $\sigma$-algebra $C$ (see e.g. Chapter 7 in [Bil99]) – and we denote its law by $\mathbb{P}_N$.

Remark 3.5. One way to interpret the function $f_N^L(x)$ is that it is a rescaled and perturbed version of $\mathcal{F}(rN + xN^{2/3}, N)$ that has been shifted by a parabola as we explain here. Using the fact that $g_\theta^{-1}(1/x) = \theta - g_\theta^{-1}(x)$ we see that $Mh_\theta(N/M) = Nh_\theta(M/N)$ and $M^{1/3}\sigma_\theta(N/M) = N^{1/3}\sigma_\theta(M/N)$. To get to $f_N^L$ from $\mathcal{F}(rN + xN^{2/3}, N)$ as in (3.4) one needs to do the following:

1. Replace $Mh_\theta(N/M)$ with $Nh_\theta(M/N)$ and $M^{1/3}\sigma_\theta(N/M)$ with $N^{1/3}\sigma_\theta(M/N)$.
2. Set $M = rN + xN^{2/3}$ and Taylor expand $h_\theta(M/N)$ and $\sigma_\theta(M/N)$ to see that

$$\mathcal{F}(rN + xN^{2/3}, N) = \frac{\log Z(1, 1; rN + xN^{2/3}, N) + N h_\theta(r) + h'_\theta(r)xN^{2/3} + (1/2)h''_\theta(r)x^2N^{1/3} + O(1)}{N^{1/3}\sigma_\theta(r)}.$$

3. Subtracting the parabola $\frac{h'_\theta(r)x^2}{2\sigma_\theta(r)}$ from the last expression and multiplying the result by $\sigma_\theta(r)$ brings us to $f_N^L(x)$ up to an $O(N^{-1/3})$ error.

Since $f_N^L(x)$ and $\mathcal{F}(rN + xN^{2/3}, N)$ are up to a deterministic shift and scalar multiplication $O(N^{-1/3})$ away from each other any asymptotic statement for $f_N^L(x)$ can be rephrased in terms of $\mathcal{F}(rN + xN^{2/3}, N)$. The next result, from [BCD21b], is phrased in terms of $f_N^L(x)$ instead of $\mathcal{F}(rN + xN^{2/3}, N)$, hence the need to give this definition.

Proposition 3.6. [BCD21b, Theorem 1.10] Fix any $T, \theta, r > 0$. Then the laws $\mathbb{P}_N$ of $f_N^L$, as in Definition 3.4, form a tight sequence in $N$.

We end this section by summarizing various properties of the functions $h_\theta$ and $\sigma_\theta$ from (3.2) and (3.3), which we use later in the text. Note that for any $\theta > 0$, since $g_\theta$ is a smooth bijection from $(0, \theta)$ to $(0, \infty)$, it follows that $\sigma_\theta$ is a smooth positive function on $(0, \infty)$. For any $\theta > 0$, $g_\theta(x) = \frac{\psi_\theta(x)}{\psi_\theta'(x)}$ and so $h_\theta(x) = \psi_\theta^{-1}(x)$. Since $g_\theta^{-1}$ and $\Psi$ are strictly increasing on $(0, \infty)$ the same is true for $h'_\theta(x) = \psi_\theta^{-1}(x)$. Also note that $g_\theta^{-1}(1) = \theta/2$ and so $h'_\theta(1) = \Psi(\theta/2)$.

Observe that $f(y) = h_\theta(y^2)$ is differentiable on $(0, 1]$ with a derivative that is bounded by a constant that depends on $\theta$ alone. Indeed, by the smoothness of $g_\theta^{-1}$ on $(0, \infty)$ and $\Psi$ on $(0, \theta)$ the latter is clear for $y$ that are bounded away from 0. On the other hand, when $y \to 0^+$ we have $g_\theta^{-1}(y^2) \sim y$ and $\Psi(y) \sim -y^{-1}$, so that $h_\theta(y^2) = h_\theta(0) = y^2 \cdot \Psi(g_\theta^{-1}(y^2)) + \Psi(\theta - g_\theta^{-1}(y^2)) = O(y)$. In particular, there is a constant $b_1 > 0$ (depending on $\theta$) such that for all $0 \leq x \leq y \leq 1$

$$|h_\theta(x) - h_\theta(y)| \leq b_1 \cdot y.$$ (3.9)

We next note that there exists $\epsilon > 0$ (depending on $\theta$) so that for all $x \in (0, \epsilon]$ we have

$$-\Psi(\theta/2) \geq \epsilon - h_\theta(x).$$ (3.10)

To see why such a choice of $\epsilon$ is possible note that when $x \to 0^+$ we have $g_\theta^{-1}(x) \sim \sqrt{x}$ and $\Psi(x) \sim -x^{-1}$, so that $h_\theta(x) = x \cdot \Psi(g_\theta^{-1}(x)) + \Psi(\theta - g_\theta^{-1}(x)) = O(\sqrt{x}) + \Psi(\theta)$. The latter implies that as $\epsilon \to 0^+$ and $x \in (0, \epsilon]$ we have $\epsilon - h_\theta(x) = O(\sqrt{x}) - \Psi(\theta)$. The last equation and the fact that $\Psi(x)$ is strictly increasing together imply the existence of $\epsilon \in (0, 1]$ so that (3.10) holds.
When $\theta = \theta_c$ as in (1.6) we also have

\begin{equation}
\begin{align*}
\frac{h_{\theta_c}'(0)}{2} = h_{\theta_c}'(1) = 0 = h_{\theta_c}(1 + x) = \frac{h_{\theta_c}''(1)}{2} \cdot x^2 + O(x^3),
\end{align*}
\end{equation}

(3.11)

and in the last inequality we used that $g_{\theta_c}^{-1}$ and $\Psi$ are strictly increasing on $(0, \infty)$.

As explained above, $h_{\theta_c}'$ is strictly increasing on $(0, \infty)$ and since $h_{\theta_c}'(1) = 0$, see (3.11), we get

\begin{equation}
\begin{align*}
\frac{h_{\theta_c}(x)}{x} \geq 0 \quad \text{for all } x \leq 1, \quad \text{and } \frac{h_{\theta_c}(x)}{x} \geq h_{\theta_c}(1/2) > 0 \quad \text{for } x \in [0, 1/2].
\end{align*}
\end{equation}

(3.12)

3.2. Miscellaneous probability results. We formulate three results that will be useful later.

\textbf{Lemma 3.7.} For all $a, \theta > 0$ there is $C > 0$ (depending on $\theta$ and $a$), such that for any $n \in \mathbb{N}$

\begin{equation}
\begin{align*}
\mathbb{P} \left( \max_{1 \leq i \leq n} \log X_i \geq \theta^{-1}(1 + a) \log n \right) \leq C n^{-a},
\end{align*}
\end{equation}

(3.13)

where $\{X_i\}_{i=1}^\infty$ are i.i.d. random variables with density $f_\theta(x)$ as in (1.7).

\textbf{Proof.} Put $f(a) = \theta^{-1}(1 + a)$. Notice that

\begin{equation}
\begin{align*}
\mathbb{P} \left( \max_{1 \leq i \leq n} \log X_i \geq f(a) \log n \right) = 1 - \mathbb{P} \left( X_1 < n^{f(a)} \right)^n.
\end{align*}
\end{equation}

In addition, we have

\begin{equation}
\begin{align*}
\mathbb{P} \left( X_1 < n^{f(a)} \right) &= 1 - \int_{n^{f(a)}}^{\infty} \frac{x^{-\theta^{-1}}}{\Gamma(\theta)} \cdot \exp(-x^{-1}) \, dx \\
&\leq 1 - \int_{n^{f(a)}}^{\infty} \frac{x^{-\theta^{-1}}}{\Gamma(\theta)} \, dx = 1 - \frac{1}{\theta \Gamma(\theta)n^{\theta f(a)}}.
\end{align*}
\end{equation}

Using that $\log(1 - x) \geq -2x$ for all $x \in [0, 1/4]$ we get

\begin{equation}
\begin{align*}
\exp \left( n \cdot \log \left( 1 - \frac{1}{\theta \Gamma(\theta)n^{\theta f(a)}} \right) \right) \geq \exp \left( -\frac{2n}{\theta \Gamma(\theta)n^{\theta f(a)}} \right),
\end{align*}
\end{equation}

provided $\frac{1}{\theta \Gamma(\theta)n^{\theta f(a)}} \in [0, 1/4]$. Under this condition, the last three inequalities combine to

\begin{equation}
\begin{align*}
\mathbb{P} \left( \max_{1 \leq i \leq n} \log X_i \geq f(a) \log n \right) \leq 1 - \exp \left( -\frac{2n}{\theta \Gamma(\theta)n^{\theta f(a)}} \right).
\end{align*}
\end{equation}

Using that $1 - e^{-x} \leq 2x$ for all $x \in [0, \infty)$ and $f(a) = \theta^{-1}(1 + a)$ the last inequality implies

\begin{equation}
\begin{align*}
\mathbb{P} \left( \max_{1 \leq i \leq n} \log X_i \geq f(a) \log n \right) \leq \frac{4}{\theta \Gamma(\theta)n^a}.
\end{align*}
\end{equation}

This implies (3.13) with $C = \frac{4}{\theta \Gamma(\theta)}$ as long as $\frac{4}{\theta \Gamma(\theta)} n^{1+a} \leq n^{1+a}$. Setting $C := \frac{4}{\theta \Gamma(\theta)} + \left( \frac{4}{\theta \Gamma(\theta)} \right)^{a/(1+a)}$, we see that (3.13) holds for $\frac{4}{\theta \Gamma(\theta)} n^{1+a} \leq n^{1+a}$ by the previous argument and for $n^{1+a} \leq \frac{4}{\theta \Gamma(\theta)}$ it holds as the right side of (3.13) is bigger than 1. Thus (3.13) holds for all $n \geq 1$ with this choice of $C$. \hfill \Box

\textbf{Proposition 3.8.} [KMT76, Theorem 1] Let $F$ be a cumulative distribution function on $\mathbb{R}$ and $\lambda > 0$ be given. Suppose that for some $t_0 > 0$ and all $|t| \leq t_0$ we have

\begin{equation}
\begin{align*}
\int_{\mathbb{R}} x F(dx) = 0, \quad \int_{\mathbb{R}} x^2 F(dx) = 1, \quad \int_{\mathbb{R}} e^{tx} F(dx) < \infty.
\end{align*}
\end{equation}

There exist constants $K, C > 0$, depending on $F$ and $\lambda$, a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and random variables $\{X_i\}_{i=1}^\infty, \{Y_i\}_{i=1}^\infty$ defined on this space that satisfy the following properties.

- Under $\mathbb{P}$ the random variables $\{X_i\}_{i=1}^\infty$ are i.i.d. with distribution $F$.
- Under $\mathbb{P}$ the random variables $\{Y_i\}_{i=1}^\infty$ are i.i.d. with a standard normal distribution.
\bullet If $S_k = X_1 + \cdots + X_k$ and $T_k = Y_1 + \cdots + Y_k$ for $k \in \mathbb{N}$ then for any $n \in \mathbb{N}$ and $x > 0$
\begin{equation}
P\left( \max_{1 \leq k \leq n} |S_k - T_k| > C \log n + x \right) < Ke^{-\lambda x}.
\end{equation}

We denote by $\Phi(x)$ and $\phi(x)$ the cumulative distribution function and density of a standard normal random variable. The following lemma can be found in [MZ12, Section 4.2].

**Lemma 3.9.** There is a constant $c_0 > 1$ such that for all $x \geq 0$ we have
\begin{equation}
\frac{1}{c_0(1 + x)} \leq \frac{1 - \Phi(x)}{\phi(x)} \leq \frac{c_0}{1 + x}.
\end{equation}

### 3.3. Bounds on point-to-point free energies

In this section we derive two estimates on $\log Z(1, 1; m, n)$. The first, given in Lemma 3.10, provides estimates on this quantity that are uniform in the entries $(m, n)$ but are sub-optimal. We use this result to control $\log Z(1, 1; m, n)$ when either $m/n$ or $n/m$ are close to 0. The second, given in Lemma 3.11, provides useful bounds on $\log Z(1, 1; m, n)$ when $m, n \in [1, N]$ in terms of $N$. We mention that Lemma 3.10 is formulated in the case $m \geq n$ to ease the notation, which does not lead to any loss of generality by the distributional equality of $\log Z(1, 1; m, n)$ and $\log Z(1, 1; n, m)$.

**Lemma 3.10.** Let $\theta > 0$ be given. For any $\epsilon \in (0, 1]$, there exist $B, b > 0$ (depending on $\theta, \epsilon$) such that for $m, n \in \mathbb{N}$, $m \geq n$ one has
\begin{equation}
P\left( \log Z(1, 1; m, n) \geq m \cdot \epsilon - m \cdot h_\theta(n/m) \right) \leq B \exp (-bm).
\end{equation}

**Proof.** For clarity we split the proof into four steps.

**Step 1.** In this step we introduce a bit of notation that will be used throughout the proof.

We define $b_0 = |\Psi(\theta)| + 2 + b_1$, where $b_1$ depends on $\theta$ and is such that (3.9) holds. We also set $\epsilon_1 = \epsilon/b_0$. In the steps below we will prove that there exist $B, b > 0$ (depending on $\theta, \epsilon$) such that for $m, n \in \mathbb{N}$, $m \geq n$ one has
\begin{equation}
P\left( \log Z(1, 1; m, n) \geq m \cdot b_0 \cdot \epsilon_1 - m \cdot h_\theta(n/m) \right) \leq B \exp (-bm).
\end{equation}

Since (3.17) clearly establishes (3.16) it suffices to prove (3.17).

**Step 2.** Put $N = \lceil m \epsilon_1 \rceil$. We claim that there exist constants $M_1, M_2 \in \mathbb{N}$, $D_1, D_2, d_1, d_2 > 0$ (depending on $\theta, \epsilon$) such that for all $m \geq M_1$ we have
\begin{equation}
P\left( \log Z(1, 1; m, N) + mh_\theta(\epsilon_1) \geq m \epsilon_1 \right) \leq D_1 \exp (-d_1 m),
\end{equation}

and for all $m \geq M_2$, and $1 \leq n \leq N$ we have
\begin{equation}
P\left( - \sum_{j=n+1}^{N} \log w_{m,j} \geq m(b_0 - b_1 - 1)\epsilon_1 \right) \leq D_2 \exp (-d_2 m),
\end{equation}

where we recall that $w_{i,j}$ are i.i.d. random variables with density (1.1). We will prove equations (3.18) and (3.19) in the steps below. Here we will assume their validity and establish (3.17) by considering the cases $n/m \in (0, \epsilon_1]$ and $n/m \in [\epsilon_1, 1]$ separately.

Using (1.3) we have for all $m, n \in \mathbb{N}$ with $n/m \in (0, \epsilon_1]$ that
\begin{equation*}
Z(1, 1, m, N) \geq Z(1, 1, m, n) \cdot \prod_{j=n+1}^{N} w_{m,j},
\end{equation*}

and by taking logarithms on both sides we conclude that
\begin{equation*}
\log Z(1, 1; m, N) \geq \log Z(1, 1; m, n) + \sum_{j=n+1}^{N} \log w_{m,j}.
\end{equation*}
The last equation implies that if \( m \geq \max(M_1, M_2) \) and \( n/m \in (0, \epsilon_1] \) then
\[
\mathbb{P}\left( \log Z(1,1;m,n) \geq m \cdot b_0 \cdot \epsilon_1 - m \cdot h_\theta(n/m) \right) \leq \\
\mathbb{P}\left( \log Z(1,1;m,N) - \sum_{j=n+1}^{N} \log w_{m,j} \geq m \cdot b_0 \cdot \epsilon_1 - m \cdot h_\theta(n/m) \right) \leq \\
\mathbb{P}\left( \log Z(1,1;m,N) - \sum_{j=n+1}^{N} \log w_{m,j} \geq m \cdot (b_0 - b_1) \cdot \epsilon_1 - m \cdot h_\theta(\epsilon_1) \right) \leq \\
\mathbb{P}\left( \log Z(1,1;m,N) + mh_\theta(\epsilon_1) \geq me_1 \right) + \mathbb{P}\left( - \sum_{j=n+1}^{N} \log w_{m,j} \geq m(b_0 - b_1 - 1)\epsilon_1 \right) \leq \\
D_1 \exp(-d_1m) + D_2 \exp(-d_2m),
\]
where in going from the second to the third line we used (3.19) and in going from the third to the fourth line we used the triangle inequality. In the last inequality we used (3.18) and (3.19). The last inequality implies (3.17) for all \( n, m \in \mathbb{N}, n/m \in (0, \epsilon_1] \) and \( m \geq \max(M_1, M_2) \). By increasing \( B \) we can ensure that the RHS of (3.17) is greater than or equal to 1 when \( m < \max(M_1, M_2) \), which concludes the proof in the case of \( n/m \in (0, \epsilon_1] \).

If \( n/m \in [\epsilon_1, 1] \) we have from Proposition 3.2 with \( \delta = \epsilon_1 \) that there exist constants \( C_1, C_2, c_1, c_2 \) (depending on \( \theta \) and \( \epsilon \)) such that for all \( n/m \in [\epsilon_1, 1] \) and \( x \geq 0 \) we have
\[
\mathbb{P}\left( \log Z(1,1;m,n) + mh_\theta(n/m) \geq m^{1/3} \sigma_\theta(n/m)x \right) \leq C_1 e^{-c_1 m} + C_2 e^{-c_2 x^{3/2}}.
\]
Since \( \sigma_\theta(y) \) is bounded away from 0 and \( \infty \) for \( y \in [\epsilon_1, 1] \), we see that the last inequality with \( x = \frac{c_1 m^{2/3} b_0}{\sigma_\theta(n/m)} \) implies that there are positive constants \( C, c > 0 \) (depending on \( \theta, \epsilon \)) such that
\[
\mathbb{P}\left( \log Z(1,1;m,n) + mh_\theta(n/m) \geq m \cdot b_0 \cdot \epsilon_1 \right) \leq C e^{-cm},
\]
which implies (3.17) in the case \( n/m \in [\epsilon_1, 1] \).

**Step 3.** In this step we establish (3.18). If \( m \geq 2\epsilon_1^{-1} \) note that \( \epsilon_1^{-1} \geq N/m \geq \epsilon_1 \). Consequently, by Proposition 3.2 (applied to \( \delta = \epsilon_1, M = m \) and \( N, \theta \) as above) we can find positive constants \( C, c \) (depending on \( \theta, \epsilon \)) such that for \( m \geq 2\epsilon_1^{-1} \)
\[
\mathbb{P}\left( \log Z(1,1;m,N) + mh_\theta(N/m) \geq m^{2/3} \sigma_\theta(N/m) \right) \leq C e^{-cm}.
\]
We now observe that by definition \( N/m = \lceil me_1 \rceil/m = \epsilon_1 + O(1/m) \) and using the differentiability of \( h_\theta \) and \( \sigma_\theta \) we have \( h_\theta(\epsilon_1) = h_\theta(N/m) + O(1/m), \sigma_\theta(\epsilon_1) = \sigma_\theta(N/m) + O(1/m) \), where the constants in the big \( O \) notations depend on \( \theta \) and \( \epsilon \). In particular, we conclude that there exists \( M_1 \in \mathbb{N} \) (depending on \( \theta \) and \( \epsilon \)) such that for \( m \geq M_1 \) we have
\[
\mathbb{P}\left( \log Z(1,1;m,N) + mh_\theta(\epsilon_1) \geq me_1 \right) \leq C e^{-cm},
\]
which implies (3.18).

**Step 4.** In this final step we prove (3.19). Since \( w_{i,j} \) are i.i.d. random variables with density (1.1), \( -\log w_{i,j} \) are i.i.d. random variables with a density on \( \mathbb{R} \), given by \( \rho_\theta(x) = \frac{e^{\theta x} \exp(-e^x)}{\Gamma(\theta)} \). If \( X \) has density \( \rho_\theta \) we see by a direct computation that its moment generating function equals
\[
M_X(t) := \mathbb{E}[e^{tX}] = \int_{\mathbb{R}} e^{tx} \rho_\theta \exp(-e^x) dx = \frac{\Gamma(\theta + t)}{\Gamma(\theta)},
\]
whenever \( t \in (-\theta, \infty) \). For \( t \leq \theta \) we have \( M_X(t) = \infty \). From the definition of \( \Psi \) in (1.5)
\[
\mathbb{E}[X] = \left. \frac{dM_X(t)}{dt} \right|_{t=0} = \Psi(\theta) \text{ and } \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \left. \frac{d^2M_X(t)}{dt^2} \right|_{t=0} - \Psi(\theta)^2 = \Psi'(\theta).
\]
Let us denote $Z_i = (-\log w_{n+i,m} - \Psi(\theta))/\sqrt{\Psi'(\theta)}$ for $i = 1, \ldots, N - n$. Then $Z_i$ are i.i.d., have mean 0 and variance 1 and the moment generating function $M_{Z_i}(t)$ is finite for $t \in (-\theta \sqrt{\Psi'(\theta)}, \infty)$.

By Proposition 3.8 applied to $\lambda = 1$, there exist constants $C, K > 0$ (depending on $\theta$ alone) and a probability space $(\Omega, F_1, P_1)$ that supports random variables $\{X_i\}_{i=1}^{N-n}$ and $\{Y_i\}_{i=1}^{N-n}$ such that $(X_1, \ldots, X_{N-n})$ has the same distribution as $(Z_1, \ldots, Z_{N-n})$ and $Y_i$ are i.i.d. standard normals and also for all $x > 0$

$$P_1 \left( \left| \sum_{i=1}^{N-n} X_i - \sum_{i=1}^{N-n} Y_i \right| \geq C \log(N-n) + x \right) < Ke^{-x}. \tag{3.20}$$

Furthermore, since $\sum_{i=1}^{N-n} Y_i$ is a normal random variable with mean 0 and variance $N-n$ we have

$$P_1 \left( \left| \sum_{i=1}^{N-n} Y_i \right| > \frac{m\epsilon_1}{2\sqrt{\Psi'(\theta)}} \right) = 2 \left[ 1 - \Phi \left( \frac{m\epsilon_1}{2\sqrt{\Psi'(\theta)(N-n)}} \right) \right] \leq c_0 \exp \left( -\frac{m^2\epsilon_1^2}{8\Psi'(\theta)(N-n)} \right), \tag{3.21}$$

where in the last inequality we used Lemma 3.9 and that $2 \leq 2\pi$. Also $c_0$ is as in Lemma 3.9.

By the triangle inequality, our definition $b_0 = |\Psi(\theta)| + 2 + b_1$ and the fact that $N-n \leq \epsilon_1 m$,

$$P \left( -\sum_{j=n+1}^{N} \log w_{m,j} \geq m(b_0 - b_1 - 1)\epsilon_1 \right) = P_1 \left( \sum_{i=1}^{N-n} X_i \geq \frac{m(b_0 - b_1 - 1)\epsilon_1 - (N-n)\Psi(\theta)}{\sqrt{\Psi'(\theta)}} \right) \leq P_1 \left( \left| \sum_{i=1}^{N-n} X_i \right| \geq \frac{m\epsilon_1}{\sqrt{\Psi'(\theta)}} \right).$$

The last inequality and the triangle inequality imply that for $\frac{m\epsilon_1}{2\sqrt{\Psi'(\theta)}} > C \log(N-n)$ we have

$$P \left( -\sum_{j=n+1}^{N} \log w_{m,j} \geq m(b_0 - b_1 - 1)\epsilon_1 \right) \leq P_1 \left( \left| \sum_{i=1}^{N-n} X_i - \sum_{i=1}^{N-n} Y_i \right| \geq \frac{m\epsilon_1}{2\sqrt{\Psi'(\theta)}} \right) + P_1 \left( \left| \sum_{i=1}^{N-n} Y_i \right| \geq \frac{m\epsilon_1}{2\sqrt{\Psi'(\theta)}} \right) \leq K \exp \left( -\frac{m\epsilon_1}{2\sqrt{\Psi'(\theta)}} + C \log(N-n) \right) + c_0 \exp \left( -\frac{m^2\epsilon_1^2}{8\Psi'(\theta)(N-n)} \right),$$

where in the last inequality we used (3.20) and (3.21). Since $N-n \leq \epsilon_1 m$ we see that the last inequality implies (3.19) and hence the lemma.

**Lemma 3.11.** Let $\theta > 0$ and $\theta_c$ be as in (1.6). If $\theta \in (0, \theta_c]$ and $a > 0$ there exist $M_1, D_1 > 0$ (depending on $\theta$ and $a$) such that for all $M \geq M_1$, $M \geq N \geq 1$ and $m \in [1, M]$, $n \in [1, N]$

$$P \left( \log Z(1, 1; m, n) \geq -\Psi(\theta/2) \cdot (M + N) + D_1 M^{1/3} (\log M)^{2/3} \right) \leq M^{-a}. \tag{3.22}$$

If $\theta > \theta_c$ and $a > 0$ there exist $M_2, D_2 > 0$ (depending on $\theta$ and $a$) such that for all $M \geq M_2$, $M \geq N \geq 1$ and $m \in [1, M]$, $n \in [1, N]$

$$P \left( \log Z(1, 1; m, n) \geq D_2 \log M \right) \leq M^{-a}. \tag{3.23}$$

**Proof.** For clarity we split the proof into four steps.

**Step 1.** In this step we introduce a bit of notation that will be used throughout the proof and fix the constants $D_1, D_2$ in the statement of the lemma.

In the remainder of the proof we fix $\epsilon > 0$ to be sufficiently small (depending on $\theta$) so that (3.10) holds. Since $\sigma_\theta(x)$ is smooth on $(0, \infty)$ we see that there is a constant $d_\epsilon > 0$ such that

$$\sigma_\theta(x) \leq d_\epsilon \text{ for all } x \in [\epsilon, 1]. \tag{3.24}$$
By Proposition 3.2 (applied to $\theta$ as in the statement of the lemma and $\delta = \epsilon$ as above) we obtain for any $m,n \in N$ with $\min(m,n)/\max(m,n) \in [\epsilon,1]$ and $y \geq 0$ that
\begin{equation}
\mathbb{P} \left( \log Z(1,1;m,n) \geq -\max(m,n)\epsilon h_\theta \left( \frac{\min(m,n)}{\max(m,n)} \right) + \max(m,n)^{1/3}\sigma_\theta \left( \frac{\min(m,n)}{\max(m,n)} \right) \cdot y \right) \\
\leq C_1 e^{-c_1 \max(m,n)} + C_2 e^{-c_2 y^3/2},
\end{equation}
for some constants $C_1, C_2, c_1, c_2 > 0$. In deriving (3.25) we also used the distributional equality of $Z(1,1;m,n)$ and $Z(1,1;n,m)$.

We let $B, b > 0$ be as in Lemma 3.10 for the above choice of $\epsilon$ and $\theta$ as in the statement of the present lemma. We let $R_\epsilon > 0$ be sufficiently large (depending on $a, \theta, \epsilon$) so the following hold
\begin{equation}
b \cdot R_\epsilon \geq a + 1, \quad c_1 \cdot R_\epsilon \geq a + 1,
\end{equation}
if $\theta > \theta_c$ we also assume $R_\epsilon$ is sufficiently large so that $c_2 d_\epsilon e^{-c_2 y^3/2} R_\epsilon \Psi(\theta/2)^{3/2} \geq a + 1$.

We recall that $\Psi(\theta/2) > 0$ for $\theta > \theta_c$, cf. (1.6), so that the choice of $R_\epsilon$ in the second line of (3.26) is indeed possible.

We finally specify the constants $D_1, D_2$ in the statement of the lemma. If $d_\epsilon$ is such that (3.24) holds and $c_1, c_2$ are such that (3.25) holds we let $D_1, D_2 > 0$ be sufficiently large so that
\begin{equation}
c_2 D_1^{3/2} d_\epsilon^{-3/2} \geq a + 1 \text{ if } \theta \in (0, \theta_c] \text{ and } D_2 - 1 \{ \theta \leq 3 \} \cdot 2 R_\epsilon \cdot \log \left( \frac{2}{\theta - 1} \right) \geq a + 1 \text{ if } \theta > \theta_c.
\end{equation}

**Step 2.** In this and the next step we prove (3.22) and (3.23) when $\max(m,n) \geq R_\epsilon \log M$. The case when $\max(m,n) \leq R_\epsilon \log M$ is handled in Step 4. We split the proof into the cases when $\min(m,n)/\max(m,n) \leq \epsilon$ and $\min(m,n)/\max(m,n) \in [\epsilon,1]$, with the former handled in this step and the latter in the next.

From the distributional equality of $Z(1,1;m,n)$ and $Z(1,1;n,m)$ and Lemma 3.10 we see that
\begin{equation}
\mathbb{P} \left( \log Z(1,1;m,n) \geq \max(m,n) \cdot \epsilon - \max(m,n) \cdot h_\theta \left( \frac{\min(m,n)}{\max(m,n)} \right) \right) \leq B \exp \left( -b \max(m,n) \right).
\end{equation}

The latter inequality, (3.10) and $b \max(m,n) \geq b R_\epsilon \log M \geq (a + 1) \log M$ (see (3.26)) imply
\begin{equation}
\mathbb{P} \left( \log Z(1,1;m,n) \geq -\Psi(\theta/2) \max(m,n) \right) \leq BM^{-a-1}.
\end{equation}

The latter inequality implies (3.22) for any $D_1 > 0$ (and thus in particular $D_1$ as in (3.27)) since when $\theta \in (0, \theta_c]$ we have $\Psi(\theta/2) \leq 0$, cf. (1.6), and $\max(m,n) \leq M + N$. The latter inequality also implies (3.23) for any $D_2 > 0$ (and thus in particular $D_2$ as in (3.27)) since when $\theta > \theta_c$ we have $\Psi(\theta/2) > 0$, cf. (1.6).

**Step 3.** In this step we prove (3.22) and (3.23) when $\max(m,n) \geq R_\epsilon \log M$ and $\min(m,n)/\max(m,n) \in [\epsilon,1]$.

As shown in Section 3.1 $h_\theta'(x)$ is strictly increasing on $(0,1]$, $g_\theta^{-1}(1) = \theta/2$ and $h_\theta'(1) = \Psi(\theta/2)$ so $h_\theta'(x) \geq \Psi(\theta/2)$ for all $x \in [\epsilon,1]$. This and the mean value theorem imply that for some $c \in [\epsilon,1]
\begin{equation}
-h_\theta(x) = -h_\theta(1) + (1 - x)h_\theta'(c) \leq -h_\theta(1) + (1 - x)\Psi(\theta/2) = -(1 + x)\Psi(\theta/2).
\end{equation}

Furthermore, combining (3.24), (3.25) and (3.28) we conclude that for all $y \geq 0$
\begin{equation}
\mathbb{P} \left( \log Z(1,1;m,n) \geq -(m + n)\Psi(\theta/2) + \max(m,n)^{1/3} d_\epsilon y \right) \leq C_1 e^{-c_1 \max(m,n)} + C_2 e^{-c_2 y^3/2}.
\end{equation}

We now use (3.29) to prove (3.22) and (3.23) by considering the cases $\theta \in (0, \theta_c]$ and $\theta > \theta_c$ separately. Suppose that $\theta \in (0, \theta_c]$ and put
\begin{equation}
y = \frac{(m + n - M - N)\Psi(\theta/2) + D_1 M^{1/3}(\log M)^{2/3}}{d_\epsilon \max(m,n)^{1/3}}.
\end{equation}
Note that if \( \theta \in (0, \theta_c] \) we have \( y \geq D_1 d_\epsilon^{-1} (\log M)^{2/3} \) since \( \max(m, n)^{1/3} \leq M^{1/3}, 1 \leq m \leq M, 1 \leq n \leq N \) and \( \Psi(\theta/2) \leq 0 \). Plugging (3.30) into (3.29) and using the fact that \( c_1 \max(m, n) \geq c_1 R_\epsilon \cdot \log M \geq (a + 1) \log M \) (see (3.26)), while \( y \geq D_1 d_\epsilon^{-1} (\log M)^{2/3} \) we conclude

\[
P \left( \log Z(1, 1; m, n) \geq D_1 M^{1/3} (\log M)^{2/3} \right) \leq C_1 M^{-a-1} + C_2 e^{-c_2 d_\epsilon^{-3/2} R_\epsilon \cdot \Psi(\theta/2)^{3/2} \log M},
\]

which implies (3.22), since by (3.27) we have \( c_2 D_1^{3/2} d_\epsilon^{-3/2} \geq a + 1 \).

Suppose now that \( \theta > \theta_c \). Put

\[
y = (m + n) \Psi(\theta/2) + D_2 \log M / \log(\max(m, n))^{1/3}.
\]

Note that if \( \theta > \theta_c \) we have from (1.6) that \( \Psi(\theta/2) \geq 0 \) and so \( y \geq d_\epsilon^{-1} \max(m, n)^{2/3} \Psi(\theta/2) \). Plugging (3.31) into (3.29) and using the fact that \( c_1 \max(m, n) \geq c_1 R_\epsilon \cdot \log M \geq (a + 1) \log M \) (see (3.26)), while \( y \geq d_\epsilon^{-1} \max(m, n)^{2/3} \Psi(\theta/2) \geq d_\epsilon^{-1} R_\epsilon^{2/3} (\log M)^{2/3} \cdot \Psi(\theta/2) \) we conclude

\[
P \left( \log Z(1, 1; m, n) \geq D_2 \log M \right) \leq C_1 M^{-a-1} + C_2 e^{-c_2 d_\epsilon^{-3/2} R_\epsilon \cdot \Psi(\theta/2)^{3/2} \log M}.
\]

The latter implies (3.23), since by (3.26) we have \( c_2 \cdot d_\epsilon^{-3/2} \cdot R_\epsilon \cdot \Psi(\theta/2)^{3/2} \geq a + 1 \).

**Step 4.** In this final step we prove (3.22) and (3.23) when \( \max(m, n) \leq R_\epsilon \log M \). As before we consider the cases \( \theta \in (0, \theta_c] \) and \( \theta > \theta_c \) separately.

Suppose that \( \theta \in (0, \theta_c] \). From Lemma 3.7 applied to \( n = M^2 \), there exists \( C_3 > 0 \) such that

\[
(3.32) \quad P \left( \max_{1 \leq i, j \leq M} \log w_{i,j} \geq \theta^{-1}(3 + a) \log M \right) \leq C_3 M^{-a-1},
\]

where \( w_{i,j} \) are the i.i.d. random variables with density (1.1) underlying the polymer model.

Let \( B_M = \{ \max_{1 \leq i, j \leq M} \log w_{i,j} \leq \theta^{-1}(3 + a) \log M \} \). By the definition of \( Z(1, 1; m, n) \), on \( B_M \)

\[
Z(1, 1; m, n) = \sum_{\pi \in \Pi(1,1,m,n)} w(\pi) = \sum_{\pi \in \Pi(1,1,m,n)} \exp \left( (m + n - 1) \theta^{-1}(3 + a) \log M \right)
\]

\[
\leq 2^{m+n} \exp \left( (m + n - 1) \cdot \theta^{-1}(3 + a) \log M \right) < \exp \left( M^{1/4} \right),
\]

where in going from the first to the second line we used that \( |\Pi(1,1,m,n)| \leq 2^{m+n} \), and the last inequality holds for all large enough \( M \) (here we used our assumption \( \max(m, n) \leq R_\epsilon \log M \)).

Combining the last two equations we see that

\[
P \left( \log Z(1, 1; m, n) \geq M^{1/4} \right) \leq C_3 M^{-a-1}.
\]

The last inequality implies (3.22) with any fixed \( D_1 > 0 \) (and thus in particular \( D_1 \) as in (3.27)) since \( \Psi(\theta/2) \leq 0 \) when \( \theta \in (0, \theta_c] \).

Suppose that \( \theta > \theta_c \). From [AS65, page 259] we have that \( \theta_c \approx 2.92326 \). By the definition of \( Z(1, 1; m, n) \) and the fact that \( E[w_{1,1}] = (\theta - 1)^{-1} \) we observe that

\[
E \left[ Z(1, 1; m, n) \right] = \sum_{\pi \in \Pi(1,1,m,n)} \prod_{i,j \in \pi} w_{i,j} = \sum_{\pi \in \Pi(1,1,m,n)} \frac{1}{(\theta - 1)^{m+n-1} \prod_{i,j \in \pi} (2/\theta - 1)} \leq \exp \left( 1 \{ \theta \leq 3 \} \cdot 2R_\epsilon \log M \cdot \log \left( \frac{2}{\theta - 1} \right) \right),
\]

where in going from the first to the second line we used that \( |\Pi(1,1,m,n)| \leq 2^{m+n-1} \), and in the last inequality we used that \( \max(m, n) \leq R_\epsilon \log M \). This and Chebyshev’s inequality imply

\[
P \left( \log Z(1, 1; m, n) \geq D_2 \log M \right) \leq \exp \left( 1 \{ \theta \leq 3 \} \cdot 2R_\epsilon \log M \cdot \log \left( \frac{2}{\theta - 1} \right) - D_2 \log M \right) \leq M^{-a-1},
\]
where in the last inequality we used that \(D_2 - 1\{\theta \leq 3\} \cdot 2R_e \cdot \log \left(\frac{2}{\theta-1}\right) \geq a + 1\), cf. (3.27). The last inequality implies (3.28). This concludes the proof of the lemma. \(\square\)

4. Proof of Theorem 1.2

In this section we will assume \(\theta \in (0, \theta_c)\), where \(\theta_c\) is as in (1.6). For such values of \(\theta\) we construct a sequence of random variables \(\mathbb{Y}_N\), which dominate with high probability \(F_N\), as in (1.4), and converge to the GUE Tracy-Widom distribution under suitable shifts and scales. The precise result is given as Proposition 4.1 in Section 4.1. In the same section we prove Theorem 1.2 by combining Proposition 4.1 and the complementary bound \(F_N \geq \log Z(1, 1; N, N)\) along with Proposition 3.1. The construction of \(\mathbb{Y}_N\) is accomplished in several steps, which are the content of Sections 4.2, 4.3 and 4.4. We continue with the same notation as in Sections 1 and 3.

4.1. Upper bounds for \(F_N\). In this section we prove Theorem 1.2 using the following proposition, whose proof is postponed until Section 4.3.

**Proposition 4.1.** Suppose that \(\theta \in (0, \theta_c)\), where \(\theta_c\) is as in (1.6). There exists a sequence of random variables \(\mathbb{Y}_N\) and events \(E_N\) such that

\[
F_N \leq \mathbb{Y}_N \text{ a.s. on } E_N \text{ for all large enough } N, \lim_{N \to \infty} P(E_N) = 1, \text{ and }
\]

\[
\lim_{N \to \infty} P\left(\frac{\mathbb{Y}_N + 2N\Psi(\theta/2)}{\sigma_N N^{1/3}} \leq y\right) = F_{\text{GUE}}(y) \text{ for any } y \in \mathbb{R}.
\]

The construction of suitable random variables \(\mathbb{Y}_N\) and events \(E_N\) is performed in three steps, which we briefly explain here.

![Diagram](image)

**Figure 5.** The picture on the left depicts the sets \(\partial SW_N^\delta\) and \(NE_N^\delta\) (these are the grey dots) and the sets \(\partial SW_N^\delta\) and \(\partial NE_N^\delta\) (these are the black dots). The picture on the right depicts the points \((m_1^*, n_1^*)\) and \((m_2^*, n_2^*)\) that maximize \(Z(m_1, n_1; m_2, n_2)\) as well as an up-right path \(\pi\) that connects these two points. If \((m_1^*, n_1^*) \in SW_N^\delta\) and \((m_2^*, n_2^*) \in NE_N^\delta\), the path \(\pi\) necessarily passes through \(\partial SW_N^\delta\) and \(\partial NE_N^\delta\). The point \((m_1, n_1)\) is the first point in \(\partial SW_N^\delta\) visited by \(\pi\) and \((m_1, n_1)\) is the closest point in \(SW_N^\delta\) to \((m_1, n_1)\) (it is also visited by \(\pi\)). The point \((m_2, n_2)\) is the last point in \(\partial NE_N^\delta\) visited by \(\pi\) and \((m_2, n_2)\) is the closest point in \(NE_N^\delta\) to \((m_2, n_2)\) (it is also visited by \(\pi\)).

The first step is to show that if \((m_1^*, n_1^*) \leq (m_2^*, n_2^*)\) are maximizers of \(Z(m_1, n_1; m_2, n_2)\) then with high probability \((m_1^*, n_1^*) \in SW_N^\delta := N^2 \cap [1, N^{1/3+2\delta}]^2\) (the south-west corner of size \(N^{1/3+2\delta}\)) and \((m_2^*, n_2^*) \in NE_N^\delta := N^2 \cap [N - N^{1/3+2\delta}, N]^2\) (the north-east corner of size \(N^{1/3+2\delta}\)) for any \(\delta \in (0, 1/3)\), see Figure 5. This step is accomplished in Section 4.2, see Lemma 4.2.
The key idea in the proof is to use Lemma 3.11 which shows that with high probability
\[ \log Z(m_1, n_1; m_2, n_2) \leq -2N \Psi(\theta/2) - N^{1/3} \log N \]
when \((m_1, n_1) \notin \text{SW}_N^\delta\) or \((m_2, n_2) \notin \text{NE}_N^\delta\). By Proposition 3.1 we have that with high probability
\[ F_N \geq \log Z(1, 1; N, N) = -2N \Psi(\theta/2) + O(N^{1/3}) . \]
The latter two statements imply with high probability \(\log Z(m_1, n_1; m_2, n_2)\) cannot compete with \(F_N\) for \((m_1, n_1), (m_2, n_2)\) away from the corners, and so \((m_1^*, n_1^*) \in \text{SW}_N^\delta\), while \((m_2^*, n_2^*) \in \text{NE}_N^\delta\).

The second step is to show that with high probability
\[ F_N \leq r_N + \max_{(m_1, n_1) \in \partial \text{SW}_N^\delta} \max_{(m_2, n_2) \in \partial \text{NE}_N^\delta} \log Z(m_1, n_1; m_2, n_2) + \Psi(\theta/2) \cdot (-m_1 - n_1 + m_2 + n_2 - 2N) , \]
where \(r_N = 2N^{1/9+2\delta/3} \log(N)\) and \(\delta \in (0, 1/3)\) (all we really need is that \(r_N = o(N^{1/3})\) and \(r_N\) grows faster than \(N^{1/9+2\delta/3}\), which is the 1/3 power of the side-lengths of \(\text{NE}_N^\delta\) and \(\text{SW}_N^\delta\), and where the maximum is over \((m_1, n_1) \in \partial \text{SW}_N^\delta\) and \((m_2, n_2) \in \partial \text{NE}_N^\delta\), which we call the frames of the corners \(\text{SW}_N^\delta\) and \(\text{NE}_N^\delta\). This step is accomplished in Section 4.3, see Lemma 4.3. The frames \(\partial \text{SW}_N^\delta, \partial \text{NE}_N^\delta\) are defined in (4.10), see also Figure 5.

To prove (4.2) we use our first step above (i.e. Lemma 4.2), which says that if \((m_1^*, n_1^*) \leq (m_2^*, n_2^*)\) are the maximizers of \(Z(m_1, n_1; m_2, n_2)\), then with high probability \((m_1^*, n_1^*) \in \text{SW}_N^\delta\) and \((m_2^*, n_2^*) \in \text{NE}_N^\delta\). In addition, from the definition of \(Z(m_1^*, n_1^*; m_2^*, n_2^*)\) we have
\[ Z(m_1^*, n_1^*; m_2^*, n_2^*) = \sum_{\pi \in \Pi(m_1^*, n_1^*; m_2^*, n_2^*)} \prod_{(i,j) \in \pi} w_{i,j} . \]
If we split the above sum over the first site in \(\partial \text{SW}_N^\delta\) and the last site in \(\partial \text{NE}_N^\delta\) that the path \(\pi\) passes through we obtain
\[ Z(m_1^*, n_1^*; m_2^*, n_2^*) = \sum_{(m_1, n_1) \in \partial \text{SW}_N^\delta} \sum_{(m_2, n_2) \in \partial \text{NE}_N^\delta} Z(m_1^*, n_1^*; (m_1, n_1)^\vee) \times \]
\[ Z(m_1, n_1; m_2, n_2) Z((m_2, n_2)^\vee; m_2^*, n_2^*) , \]
where \((m_1, n_1)^\vee\) is the closest vertex in \(\text{SW}_N^\delta\) to \((m_1, n_1) \in \partial \text{SW}_N^\delta\), and \((m_2, n_2)^\vee\) is the closest vertex in \(\text{NE}_N^\delta\) to \((m_2, n_2) \in \partial \text{NE}_N^\delta\). The points \((m_1, n_1)^\vee\), \((m_2, n_2)^\vee\) are depicted in Figure 5 and a precise definition can be found in equations (4.13) and (4.14). We have the following trivial upper bounds for the terms on the right side of (4.3)
\[ Z(m_1^*, n_1^*; (m_1, n_1)^\vee) \leq Z^{\max}((m_1, n_1)^\vee) := \max_{(1,1) \leq (m,n) \leq (m_1, n_1)^\vee} Z(m, n; (m_1, n_1)^\vee) , \]
\[ Z((m_2, n_2)^\vee; m_2^*, n_2^*) \leq Z^{\max}((m_2, n_2)^\vee) := \max_{(m_2, n_2)^\vee \leq (m,n) \leq (N,N)} Z((m_2, n_2)^\vee; m, n) . \]
In particular, from (4.3) and (4.4) we obtain
\[ Z(m_1^*, n_1^*; m_2^*, n_2^*) \leq |\partial \text{SW}_N^\delta| \cdot |\partial \text{NE}_N^\delta| \times \]
\[ \max_{(m_1, n_1) \in \partial \text{SW}_N^\delta} \max_{(m_2, n_2) \in \partial \text{NE}_N^\delta} Z^{\max}((m_1, n_1)^\vee) \cdot Z(m_1, n_1; m_2, n_2) \cdot Z^{\max}((m_2, n_2)^\vee) . \]
Once we have (4.5) in place, obtaining (4.2) is relatively straightforward, since we can control the right sides of (4.4) using Lemma 3.11. The terms \(|\partial \text{SW}_N^\delta|\) and \(|\partial \text{NE}_N^\delta|\) are benign, since upon taking logarithms they become \(O(\log N)\) and get absorbed in the error term \(2N^{1/9+2\delta/3} \log N\) in (4.2).
The benefit of the above two steps is that we have reduced our problem to finding a sequence $Y_N$ that (under appropriate shifts and scales) converges to $F_{\text{GUE}}$, and such that $Y_N$ dominates the right side of (4.2) with high probability. The setup of having the arguments $(m_1, n_1), (m_2, n_2)$ of $\log Z(m_1, n_1; m_2, n_2)$ vary over essentially 1-dimensional segments is important for the application of our tightness result, Proposition 3.6, which is the crux of the third step of constructing $Y_N$. While the first two steps are fairly intuitive, the third step is a bit involved and so we will postpone its discussion to the main argument in Section 4.4 after we have introduced more notation.

In the remainder of this section we use Proposition 4.1 to prove Theorem 1.2.

**Proof of Theorem 1.2.** Let $E_N$ and $Y_N$ be as in the statement of Proposition 4.1 and $y \in \mathbb{R}$. It follows from (4.1) that

$$
\liminf_{N \to \infty} \mathbb{P} \left( \frac{F_N + 2N \Psi(\theta/2)}{\sigma \sqrt{N}} \leq y \right) \leq \liminf_{N \to \infty} \mathbb{P} \left( E_N \cap \left\{ \frac{F_N + 2N \Psi(\theta/2)}{\sigma \sqrt{N}} \leq y \right\} \right) \geq
$$

$$
\liminf_{N \to \infty} \mathbb{P} \left( E_N \cap \left\{ \frac{Y_N + 2N \Psi(\theta/2)}{\sigma \sqrt{N}} \leq y \right\} \right) \geq \liminf_{N \to \infty} \mathbb{P} \left( \frac{Y_N + 2N \Psi(\theta/2)}{\sigma \sqrt{N}} \leq y \right) - \mathbb{P}(E_N) = F_{\text{GUE}}(y).
$$

On the other hand, we have almost surely $F_N \geq \log Z(1, 1; N, N)$. The latter combined with the fact that $g_\theta^{-1}(1) = \theta/2$ and Proposition 3.1 imply

$$
\limsup_{N \to \infty} \mathbb{P} \left( \frac{F_N + 2N \Psi(\theta/2)}{\sigma \sqrt{N}} \leq y \right) \leq \limsup_{N \to \infty} \mathbb{P} \left( \frac{\log Z(1, 1; N, N) + 2N \Psi(\theta/2)}{\sigma \sqrt{N}} \leq y \right) = F_{\text{GUE}}(y).
$$

The last two equations imply the statement of the theorem. \qed

4.2. Localization to corners. For two finite sets $K_1, K_2 \subset \mathbb{N}^2$ we let

$$
F(K_1, K_2) := \max_{(m_1, n_1) \leq (m_2, n_2)} \log Z(m_1, n_1; m_2, n_2),
$$

where we set $F(K_1, K_2) = -\infty$ if the set over which the maximum is taken is empty.

The goal of this section is to show that if $(m_1^*, n_1^*) \leq (m_2^*, n_2^*)$ are maximizers of $Z(m_1, n_1; m_2, n_2)$ then $(m_1^*, n_1^*) \in \text{SW}_N^\delta$ and $(m_2^*, n_2^*) \in \text{NE}^\delta_N$ for any $\delta \in (0, 1/3)$ with very high probability, where

$$
\text{SW}_N^\delta := \mathbb{N}^2 \cap [1, N^{1/3+2\delta}]^2 \quad \text{and} \quad \text{NE}^\delta_N := \mathbb{N}^2 \cap [N - N^{1/3+2\delta}, N]^2.
$$

The precise statement is detailed in the following lemma.

**Lemma 4.2.** Suppose that $\theta \in (0, \theta_c)$, where $\theta_c$ is as in (1.6) and fix $\delta \in (0, 1/3)$. Then

$$
\lim_{N \to \infty} \mathbb{P} \left( F_N = F(\text{SW}_N^\delta, \text{NE}_N^\delta) \right) = 1.
$$

**Proof.** For clarity we split the proof into two steps.

**Step 1.** We claim that there exists a constant $N_0 > 0$ (depending on $\theta$ and $\delta$) such that for all $N \geq N_0$, $m_1, n_1, m_2, n_2 \in [1, N]$ with $(m_1, n_1) \leq (m_2, n_2)$ and $((m_1, n_1), (m_2, n_2)) \notin \text{SW}_N^\delta \times \text{NE}_N^\delta$,

$$
\mathbb{P} \left( \log Z(m_1, n_1; m_2, n_2) \geq -2N \Psi(\theta/2) - N^{1/3} \log N \right) \leq N^{-5}.
$$

We will establish (4.9) in Step 2 below. Here we assume it and conclude the proof of the lemma.

Let $A_N$ be the event that $F_N = \log Z(m_1, n_1; m_2, n_2)$ for some $m_1, m_2, n_1, n_2 \in [1, N]$ such that $(m_1, n_1) \leq (m_2, n_2)$ and $((m_1, n_1), (m_2, n_2)) \notin \text{SW}_N^\delta \times \text{NE}_N^\delta$. By taking union bound of (4.9) (note that there are at most $N^4$ choices for the quadruple $m_1, m_2, n_1, n_2$) we conclude that

$$
\mathbb{P} \left( A_N \cap \{ F_N \geq -2N \Psi(\theta/2) - N^{1/3} \log N \} \right) \leq N^{-1}.
$$
In addition, we know that \( F_N \geq \log Z(1, 1; N, N) \). The latter, combined with the fact that \( h_\theta(1) = 2\Psi(\theta/2) \) and Proposition 3.11 implies
\[
\lim_{N\to\infty} \Pr \left( F_N \geq -2N\Psi(\theta/2) - N^{1/3} \log N \right) = 1.
\]
The last two equations imply that \( \lim_{N\to\infty} \Pr(A_N) = 0 \), which in turn implies (4.8).

**Step 2.** In this step we prove (4.9). Since \( Z(m_1, n_1; m_2, n_2) \) has the same distribution as \( Z(n_1, m_1; n_2, m_2) \) we may assume that \( m_2 - m_1 \geq n_2 - n_1 \). In addition, notice that \( Z(m_1, n_1; m_2, n_2) \) has the same distribution as \( Z(1, 1; m_2 - m_1 + 1, n_2 - n_1 + 1) \). Finally, observe that since we have assumed that \( [(m_1, n_1), (m_2, n_2)] \notin SW_N^\delta \times NE_N^\delta \), we have \( \log m_2 - m_1 + 1 \geq n_2 - n_1 + 1 \), and \( m_2 - m_1 + 1 \leq N-k+1 \) where \( k = \lceil N^{1/3+2\delta} \rceil + 1 \).

The latter observations and Lemma 3.11 (applied to \( a = 5 \), \( M = N \) and \( N = N-k+1 \)) together imply that for all large \( N \)
\[
\Pr \left( \log Z(m_1, n_1; m_2, n_2) \geq -\Psi(\theta/2) \cdot (2N-k+1) + N^{1/3} \log N \right) \leq N^{-5}.
\]
The latter inequality implies (4.9) since \( k = \lceil N^{1/3+2\delta} \rceil + 1 \) and \( \Psi(\theta/2) < 0 \).

### 4.3. Comparison with the maximum over frames.
For a fixed \( \delta \in (0, 1/3) \) and \( N \in \mathbb{N} \) let \( SW_N^\delta \) and \( NE_N^\delta \) be as in (4.7). We define the frames of these two corners to be (see Figure 5)
\[
\begin{align*}
\partial SW_N^\delta &:= \{(m, n) \in \mathbb{N}^2 \setminus SW_N^\delta : (m-1, n) \in SW_N^\delta \text{ or } (m, n-1) \in SW_N^\delta\}, \\
\partial NE_N^\delta &:= \{(m, n) \in \mathbb{N}^2 \setminus NE_N^\delta : (m+1, n) \in NE_N^\delta \text{ or } (m, n+1) \in NE_N^\delta\}.
\end{align*}
\]
The main result of the section is the following lemma, which substantiates (4.2) from Section 4.1.

**Lemma 4.3.** Under the same notation and assumptions as in Lemma 4.2 we have
\[
\lim_{N\to\infty} \Pr \left( F(SW_N^\delta, NE_N^\delta) \leq r_N + \max_{(m_1, n_1) \in \partial SW_N^\delta, (m_2, n_2) \in \partial NE_N^\delta} \log Z(m_1, n_1; m_2, n_2) + \Psi(\theta/2) \cdot D_N \right) = 1,
\]
where \( r_N = 2N^{1/3+2\delta/3} \log(N) \) and \( D_N(m_1, n_1; m_2, n_2) = -m_1 - n_1 + m_2 + n_2 - 2N \).

**Proof.** For clarity we split the proof into two steps.

**Step 1.** In this step we introduce some notation, which will be required for our arguments. For simplicity we set \( k = \lceil N^{1/3+2\delta} \rceil + 1 \). Let us define the sets
\[
\begin{align*}
\partial SW_N^\delta, right &:= \{(m, n) \in \partial SW_N^\delta : m = k\}, \\
\partial SW_N^\delta, top &:= \{(m, n) \in \partial SW_N^\delta : n = k\}, \\
\partial NE_N^\delta, left &:= \{(m, n) \in \partial NE_N^\delta : m = N-k\}, \\
\partial NE_N^\delta, bot &:= \{(m, n) \in \partial NE_N^\delta : n = N-k\},
\end{align*}
\]
where \( \partial SW_N^\delta, \partial NE_N^\delta \) are as in equation (4.10). Note that \( \partial SW_N^\delta = \partial SW_N^\delta, right \cup \partial SW_N^\delta, top \) and \( \partial NE_N^\delta = \partial NE_N^\delta, left \cup \partial NE_N^\delta, bot \). Let \( (m_1^*, n_1^*) \in SW_N^\delta, (m_2^*, n_2^*) \in NE_N^\delta \) be such that
\[
F(SW_N^\delta, NE_N^\delta) = \log Z(m_1^*, n_1^*; m_2^*, n_2^*),
\]
and recall from (4.3) that we have
\[
Z(m_1^*, n_1^*; m_2^*, n_2^*) = \sum_{\pi \in \Pi(m_1^*, n_1^*; m_2^*, n_2^*)} \prod_{(i,j) \in \pi} w_{i,j}.
\]
Assuming that \( N \) is sufficiently large so that \( N \geq 2k+1 \) we have that \( SW_N^\delta \) and \( NE_N^\delta \) are well-separated and so we can split the above sum over the first site in \( \partial SW_N^\delta \) and the last site in \( \partial NE_N^\delta \).
that the path \( \pi \) passes through, see Figure 5. Consequently, we obtain the formula (4.13), where the notation \((m_1, n_1)^c\) and \((m_2, n_2)^c\) used therein is explicitly given by

\[
(4.13) \quad (m_1, n_1)^c = \begin{cases} 
(m_1 - 1, n_1) & \text{if } (m_1, n_1) \in \partial SW_N^{b, \text{right}}, \\
(m_1, n_1 - 1) & \text{if } (m_1, n_1) \in \partial SW_N^{b, \text{top}},
\end{cases}
\]

and

\[
(4.14) \quad (m_2, n_2)^c = \begin{cases} 
(m_2 + 1, n_2) & \text{if } (m_2, n_2) \in \partial NE_N^{b, \text{left}}, \\
(m_2, n_2 + 1) & \text{if } (m_2, n_2) \in \partial NE_N^{b, \text{bot}}.
\end{cases}
\]

For \((m_1, n_1) \in \partial SW_N^b\) and \((m_2, n_2) \in \partial NE_N^b\), let \(Z^{\max}(m_1, n_1)^c\) and \(Z^{\max}(m_2, n_2)^c\) be as in (4.14).

**Step 2.** Note that \(Z(m_1, n_1; m_2, n_2)\) has the same distribution as \(Z(m_1 + a, n_1 + b; m_2 + a, n_2 + b)\) for any \(a, b \in \mathbb{Z}\). The latter implies that for any \((m_2, n_2) \in \partial NE_N^{b, \text{left}}\) and \((m_2, n_2)^c \leq (m, n) \leq (N, N)\) we have that \(Z(m_2, n_2)^c; m, n\) has the same law as \(Z(1, 1; m - m_2, n - n_2 + 1)\). From Lemma 3.11 (applied to \(M = k\), \(N = N - n_2 + 1\) and \(a = 6\)) we conclude that for all large \(N\),

\[
(4.15) \quad P\left( Z((m_2, n_2)^c; m, n) \geq -\Psi(\theta/2) \cdot (k + N - n_2 + 1) + k^{1/3}\log k \right) \leq k^{-6},
\]

provided that \((m_2, n_2) \in \partial NE_N^{b, \text{left}}\) and \((N, N) \geq (m, n) \geq (m_2, n_2)^c\).

Taking a union bound of (4.15) over \((m, n)\) (note that there are at most \(2^2\) possible choices) we get for all large \(N\) and any \((m_2, n_2) \in \partial NE_N^{b, \text{left}}\) (notice that \(k = N - m_2\))

\[
(4.16) \quad P\left( \log Z^{\max}(m_2, n_2)^c \geq -\Psi(\theta/2) \cdot (2N - n_2 - m_2 + 1) + k^{1/3}\log k \right) \leq k^{-4}.
\]

Using the distributional equality of \(Z(m_1, n_1; m_2, n_2)\) and \(Z(n_1, m_1; m_2, n_2)\) we conclude that (4.16) holds for \((m_2, n_2) \in \partial NE_N^{b, \text{bot}}\) as well. An analogous argument allows us to conclude that for \((m_1, n_1) \in \partial SW_N^b\)

\[
(4.17) \quad P\left( \log Z^{\max}((m_1, n_1)^c) \geq -\Psi(\theta/2) \cdot (m_1 + n_1 - 1) + k^{1/3}\log k \right) \leq (k - 1)^{-4}.
\]

We mention that we have \(k^{-4}\) in (4.16) and \((k - 1)^{-4}\) in (4.17), because \(SW_N^b = [1, k - 1] \times [1, k - 1]\) (and hence contains \((k - 1)^2\) points), while \(NE_N^b = [N - k + 1, N] \times [N - k + 1, N]\) (and hence contains \(k^2\) points). This discrepancy can be traced to the fact that we are working with the domain \([1, N] \times [1, N]\) as opposed to \([0, N] \times [0, N]\).

Let \(A_N\) denote the event

\[
A_N := \bigcap_{(m_1, n_1) \in \partial SW_N^b} \left\{ \log Z^{\max}((m_1, n_1)^c) < -\Psi(\theta/2) \cdot (2N - n_2 - m_2 + 1) + k^{1/3}\log k \right\} \bigcap \bigcap_{(m_2, n_2) \in \partial NE_N^b} \left\{ \log Z^{\max}((m_2, n_2)^c) < -\Psi(\theta/2) \cdot (m_1 + n_1 - 1) + k^{1/3}\log k \right\},
\]

and observe that by a union bound we have \(1 - P(A_N) = O(k^{-2})\) so that \(\lim_{N \to \infty} P(A_N) = 1\).

We see that on the event \(A_N\) we have the following tower of inequalities

\[
Z(m_1^*, n_1^*; m_2^*, n_2^*) \leq \sum_{(m_1, n_1) \in \partial SW_N^b} \sum_{(m_2, n_2) \in \partial NE_N^b} Z^{\max}((m_1, n_1)^c) Z(m_1, n_1; m_2, n_2) Z^{\max}((m_2, n_2)^c) \leq 4k^2 \cdot \exp\left(2k^{1/3}\log k\right) \cdot \max_{(m_1, n_1) \in \partial SW_N^b, (m_2, n_2) \in \partial NE_N^b} Z(m_1, n_1; m_2, n_2) \cdot \exp\left(-\Psi(\theta/2)[2N - n_2 - m_2 + m_1 + n_1]\right),
\]

where in the first inequality we used (4.3) and \(Z^{\max}((m_1, n_1)^c) \geq Z(m_1^*, n_1^*; (m_1, n_1)^c)\), \(Z^{\max}((m_2, n_2)^c) \geq Z((m_2, n_2)^c; m_2^*, n_2^*)\); in the second inequality we used the definition of \(A_N\).
and the fact that $|\partial \text{SW}_N^\delta| = 2k - 2$, $|\partial \text{NE}_N^\delta| = 2k$. The last inequality, the fact that $k = \lfloor N^{1/3+2\delta} \rfloor + 1$ and $\lim_{N \to \infty} \mathbb{P}(A_N) = 1$ together imply (4.11).

4.4. Proof of Proposition 4.1. In this section we present the proof of Proposition 4.1. For clarity we split the proof into three steps.

Step 1. In this step we define $Y_N$ and $E_N$ as in the statement of the proposition.

Set $\delta = 1/12$, $K = \lfloor N^{4/5} \rfloor$ and define the random variables

\[
Y_N^{\text{SW}} = \min_{(m,n) \in \partial \text{SW}_N^\delta} \log Z(-K, -K; m, n) + \Psi(\theta/2)(2K + m + n),
\]

\[
Y_N^{\text{NE}} = \min_{(m,n) \in \partial \text{NE}_N^\delta} \log Z(m, n; K + N, K + N) + \Psi(\theta/2)(2K + 2N - m - n),
\]

where $\partial \text{SW}_N^\delta$ and $\partial \text{NE}_N^\delta$ are as in (4.10). We then define

\[
Y_N = 2N^{1/4} + \log Z(-K, -K; N + K, N + K) + 4K\Psi(\theta/2) - Y_N^{\text{SW}} - Y_N^{\text{NE}}.
\]

This specifies the variables $Y_N$. We mention that the $N^{1/4}$ in (4.19) is a convenient choice – for this term we could take any $N^\rho$ with $1/6 < \rho < 1/3$. The bound $\rho < 1/3$ is dictated from our desire to have this modification become irrelevant in the $N^{1/3}$ scale. The bound $\rho > 1/6$ is dictated from our desire to have $r_N \leq N^\rho$, where $r_N = 2N^{1/6}\log(N)$ is as in Lemma 4.3 for $\delta = 1/12$.

We next define the events $E_N = E_N^1 \cap E_N^2 \cap E_N^3$, where

\[
E_N^1 = \left\{ F_N = F(\text{SW}_N^\delta, \text{NE}_N^\delta) \right\}, \quad E_N^2 = \left\{ \max_{1 \leq i,j \leq N} \log w_{i,j} \leq 4\theta^{-1}\log N \right\},
\]

\[
E_N^3 = \left\{ F(\text{SW}_N^\delta, \text{NE}_N^\delta) \leq N^{1/4} + \max_{(m,n) \in \partial \text{SW}_N^\delta} \log Z(m, n; m_2, n_2) + \Psi(\theta/2) \cdot D_N \right\},
\]

and we recall from Lemma 4.3 that $D_N(m_1, n_1; m_2, n_2) = -m_1 - n_1 + m_2 + n_2 - 2N$.

Step 2. Recall from the statement of the proposition that we seek to prove that

(I) $F_N \leq Y_N$ a.s. on $E_N$ for all large enough $N$, \hspace{1cm} (II) $\lim_{N \to \infty} \mathbb{P}(E_N) = 1,$

(III) $\lim_{N \to \infty} \mathbb{P}\left( \frac{Y_N + 2N\Psi(\theta/2)}{\sigma_\theta N^{1/3}} \leq y \right) = F_{\text{GUE}}(y)$ for any $y \in \mathbb{R}$.

In this step we prove (III).

From the definition of $Y_N$ in (4.19) we have

\[
\frac{Y_N + 2N\Psi(\theta/2)}{\sigma_\theta N^{1/3}} = \frac{\log Z(-K, -K; N + K, N + K) + (4K + 2N)\Psi(\theta/2)}{\sigma_\theta (N + K)^{1/3}} \cdot \frac{(N + K)^{1/3}}{N^{1/3}} + \frac{2N^{1/4} - Y_N^{\text{SW}} - Y_N^{\text{NE}}}{\sigma_\theta N^{1/3}}.
\]

From Proposition 3.1 we have that the first term on the right weakly converges to $F_{\text{GUE}}$ as $N \to \infty$ (here we used that $K = \lfloor N^{4/5} \rfloor$ so that $N^{-1/3}(N + K)^{1/3} \to 1$ as $N \to \infty$). Thus all we need to show is that $N^{-1/3}(Y_N^{\text{SW}} + Y_N^{\text{NE}})$ converges weakly to zero.

Focusing on $Y_N^{\text{SW}}$, we note that $Y_N^{\text{SW}} = \min \left( Y_N^{\text{top,SW}}, Y_N^{\text{right,SW}} \right)$, where

\[
Y_N^{\text{top,SW}} = \min_{(m,n) \in \partial \text{SW}_N^{\delta,\text{top}}} \log Z(-K, -K; m, n) + \Psi(\theta/2)(2K + m + n),
\]

\[
Y_N^{\text{right,SW}} = \min_{(m,n) \in \partial \text{SW}_N^{\delta,\text{right}}} \log Z(-K, -K; m, n) + \Psi(\theta/2)(2K + m + n),
\]
and we recall that \( \partial F_N^{k,\text{top}} \), \( \partial F_N^{k,\text{right}} \) were defined in \([4,12]\).

In words, \( Y_{N}^{\text{top,SW}} \) is the (properly centered) minimal point-to-point free energy from the point \((-K, -K)\) to \( \partial F_N^{k,\text{top}} \). This set of points is drawn on Figure 6 and is a horizontal segment connecting \((1, k)\) and \((k, k)\). Similarly, \( Y_{N}^{\text{right,SW}} \) is the (properly centered) minimal point-to-point free energy from the point \((-K, -K)\) to \( \partial F_N^{k,\text{right}} \).

![Figure 6](image)

**Figure 6.** The picture depicts the geometric meaning of the quantities \( N, k = \lfloor N^{1/2} \rfloor + 1, K = \lfloor N^{4/5} \rfloor, \) and \((k, k)\) that appear in our arguments.

Let us set \( k = \lfloor N^{1/2} \rfloor + 1 \) and for each \( x \in [-1 - (k+k)^{-2/3}, 1 + (k+k)^{2/3}] \), such that \( x(k+k)^{2/3} \) is an integer, define \( s = \lfloor x(k+k)^{2/3} \rfloor \) and

\[
f_N(x) = (k+k)^{-1/3} \left( \log Z(-K, -K; s - K, k) + h_0(1)(K+k) + h_0(1)x(k+k)^{2/3} \right),
\]

and then extend \( f_N \) to all points \( x \in [-1,1] \) by linear interpolation. The function \( f_N \) records the (centered and scaled) point-to-point free energy from the point \((-K, -K)\) to the horizontal segment of length \( 2(k+k)^{2/3} \) centered at the point \((k, k)\) (half of this segment is drawn in Figure 6). We have the following relationship between \( f_N \) and \( Y_{N}^{\text{top,SW}} \)

\[
(K+k)^{-1/3} Y_{N}^{\text{top,SW}} = \min_{x \in [-k-k, (k+k)^{-2/3}, -(k+k)^{-2/3}]} f_N(x).
\]

From Proposition 3.6 we know that \( f_N \) form a tight \( N \)-indexed sequence of random continuous functions on \([-1,1]\) (here we also used that \( Z(1,1;\cdot,\cdot) \) has the same distribution as \( Z(a+1,b+1; a+\cdot, b+\cdot) \) for any \( a,b \in \mathbb{Z} \)). Consequently, \( (k+k)^{-1/3} Y_{N}^{\text{top,SW}} \) also forms a tight \( N \)-indexed sequence. Here it is important that \( K^{2/3} \) (which is the transversal scale in \( f_N \)) is larger than \( k \) so that \( k(k+k)^{-2/3} \leq 1 \). The latter is true since \( K \sim N^{4/5} \), while \( k \sim N^{1/2} \). The fact that \( (k+k)^{-1/3} Y_{N}^{\text{top,SW}} \) is tight implies that \( N^{-1/3} Y_{N}^{\text{right,SW}} \) weakly converges to zero, since \( N^{-1/3}(K+k)^{1/3} \to 0 \). Here it is important that \( K/N \to 0 \), which is true as \( K \sim N^{4/5} \).

Using that \( Z(m_1, n_1; m_2, n_2) \) has the same finite dimensional distribution as \( Z(n_1, m_1; n_2, m_2) \) we analogously conclude that \( N^{-1/3} Y_{N}^{\text{right,SW}} \) weakly converges to zero and then so does \( N^{-1/3} Y_{N}^{\text{top,SW}} \). Using that \( Z(m_1, n_1; m_2, n_2) \) has the same finite dimensional distribution as \( Z(-m_2, -n_2; -m_1, -n_1) \) we
can repeat the above argument to conclude that $N^{-1/3}Y_N^{\text{SW}}$ weakly converges to zero. This means that $N^{-1/3}(\mathcal{Y}^{\text{SW}}_N + \mathcal{Y}^{\text{NE}}_N)$ converges weakly to zero, which is what was needed to complete $(III)$.

In plain words, the consequence of Proposition 3.6 is that both $Y_N^{\text{SW}}$ and $Y_N^{\text{NE}}$ are of order $O((K + k)^{1/3})$ and thus do not affect the asymptotics of $Y_N$ in (4.22) because of the $N^{1/3}$ in the denominator. Thus the fluctuations of $Y_N$ are entirely governed by the first term on the right of (4.22), which weakly converges to $F_{\text{GUE}}$ as $N \to \infty$ by Proposition 3.4.

**Step 3.** In this step we prove $(I)$ and $(II)$ in (4.21). From Lemma 4.2 with $\delta = 1/12$ we have $\lim_{N \to \infty} P(E_N^1) = 1$; from Lemma 3.7 (with $a = 1$ and $n = N^2$) we have $\lim_{N \to \infty} P(E_N^2) = 1$; and from Lemma 4.3 with $\delta = 1/12$ we have $\lim_{N \to \infty} P(E_N^3) = 1$. This proves $(II)$.

On the event $E_N$ we know that $E_N^1 \cap E_N^3$ also holds, and thus to prove $(I)$ in (4.21) it suffices to prove that for all large enough $N$ we have

(4.23) $Y_N \geq N^{1/4} + \max_{(m_1, n_1) \in \partial \mathcal{SW}_N^5} \log Z_{n_1} + \Psi(\theta/2) \cdot (-m_1 - n_1 + m_2 + n_2 - 2N) - \Psi(\theta/2) \cdot \langle \mathbf{1}, \mathbf{1} \rangle$

Let us fix $(m_1, n_1) \in \partial \mathcal{SW}_N^5$ and $(m_2, n_2) \in \partial \mathcal{NE}^5_N$ that achieve the maximum in (4.23). By the definition of $Z_{-K; -K; N + K, N + K}$ from (1.3) we have

$$Z_{-K; -K; N + K, N + K} = \prod_{\pi \in \Pi(-K; -K; N + K, N + K)} w_{i,j},$$

and by restricting the above sum to paths $\pi$ that pass through $(m_1, n_1)$ and $(m_2, n_2)$ we obtain

(4.24) $Z_{m_1, n_1; m_2, n_2} \cdot Z_{m_2, n_2; N + K, N + K} \cdot w_{n_1}^{-1} m_1 \cdot w_{n_2}^{-1} m_2.$

We now observe the following inequalities, which hold for large enough $N$:

$$N^{1/4} + \log Z_{m_1, n_1; m_2, n_2} + \Psi(\theta/2) \cdot (-m_1 - n_1 + m_2 + n_2 - 2N) \leq N^{1/4} + \log Z_{-K; -K; N + K, N + K} + \Psi(\theta/2) \cdot (-m_1 - n_1 + m_2 + n_2 - 2N) - \log Z_{-K; -K; m_1, n_1} - \log Z_{m_2, n_2; N + K, N + K} + \log w_{n_1}^{-1} m_1 + \log w_{n_2}^{-1} m_2 \leq N^{1/4} + 8\theta^{-1} \log N + \log Z_{-K; -K; N + K, N + K} + 4K \Psi(\theta/2) - Y_N^{\text{NE}} \leq Y_N^{\text{SW}}$$

where in the first inequality we used (4.24), in the second inequality we used the definitions of $Y_N^{\text{NE}}, Y_N^{\text{SW}}$ from (1.18) and also that $\log w_{i,j} \leq 4\theta^{-1} \log N$ since on the event $E_N$, we have $E_N^2$ as well. In the last inequality we used the definition of $Y_N$ from (1.19) and the fact that $8\theta^{-1} \log N \leq N^{1/4}$ for all large $N$. This establishes (4.23) and hence the proposition.

5. Proof of Theorem 1.3

In this section we work in the case $\theta = \theta_c$, where $\theta_c$ is as in (1.6). For this choice of $\theta$ we construct a sequence of random variables $Y_N$ such that $Y_N \leq F_N$, as in (1.4), and $Y_N$ are, with high probability, bounded away from zero on scale $N^{1/3}(\log N)^{2/3}$. The precise result is given as Proposition 5.1 in Section 5.1. In the same section we use Proposition 5.1 along with Lemma 3.11 (which gives an upper bound) to prove Theorem 1.3. The construction of $Y_N$ is accomplished in Section 5.2 and relies on two key lemmas, Lemma 5.3 and Lemma 5.4 whose proofs are given in Sections 5.3 and A, respectively. We continue with the same notation as in Sections 4 and 3.

5.1. Lower bounds for $F_N$. We prove Theorem 1.3 via the following (proved in Section 5.2).

**Proposition 5.1.** Suppose that $\theta = \theta_c$. There exists a sequence of random variables $Y_N$ such that

(5.1) $Y_N \leq F_N$ a.s. for all large $N$, and $\lim_{N \to \infty} P\left(Y_N \geq 20^{-1} \cdot \sigma_{\theta_c}(1) \cdot N^{1/3}(\log N)^{2/3}\right) = 1$,

where we recall that $\sigma_{\theta_c}(x)$ is as in (2.23).
Proof of Theorem 7.3. From Proposition 5.1, \( \lim_{N \to \infty} P \left( F_N \geq 20^{-1} \cdot \sigma_{\theta_1}(1) \cdot N^{1/3} (\log N)^{2/3} \right) = 1 \), which establishes the lower bound in Theorem 1.3.

On the other hand, we have by Lemma 5.11 (applied to \( a = 5, M = N \) and \( \theta = \theta_c \)) that there exists \( D_1 > 0 \) such that for all large \( N \), \( P \left( \log Z(1,1;m,n) \geq D_1 N^{1/3} (\log N)^{2/3} \right) \leq N^{-5} \). Since \( Z(m_1,n_1;m_2,n_2) \) has the same distribution as \( Z(a + m_1,b + n_1; a + m_2, b + n_2) \) for any \( a,b \in \mathbb{Z} \), we conclude the same with \( (1,1;m,n) \) replaced by \( (m_1,n_1;m_2,n_2) \) for all \( m_1,n_1,m_2,n_2 \in [1,N] \) with \( (m_1,n_1) \leq (m_2,n_2) \). Taking a union bound of these inequalities (note that we have at most \( N^4 \) possible choices for the quadruple of \( m_1,n_1,m_2,n_2 \)) we conclude that
\[
P \left( F_N \geq D_1 N^{1/3} (\log N)^{2/3} \right) \leq N^{-1},
\]
which establishes the upper bound in Theorem 1.3 with \( c_2 = D_1 \). \( \square \)

5.2. Proof of Proposition 5.1. In this section we present the proof of Proposition 5.1. We begin this section by introducing some notation and explicitly constructing the variables \( Y_N \) in the statement of Proposition 5.1. Afterwards we explain the main ideas behind our construction, and state two lemmas, which will be used in our analysis. Finally, we show that the \( Y_N \) we construct satisfy the conditions in Proposition 5.1.

Definition 5.2. For \( P = (x_1,y_1), Q = (x_2,y_2) \in \mathbb{R}^2 \) with \( x_1 \leq x_2 \) and \( y_1 \leq y_2 \) and \( a \in [0,\infty) \) satisfying \( \min(x_2 - x_1, y_2 - y_1) \geq a \) we let \( H(P;Q;a) \) denote the hexagonal domain in \( \mathbb{R}^2 \), whose boundary is obtained by straight segments connecting the points
\[
(x_1,y_1) \to (x_1 + a,y_1) \to (x_2,y_2 - a) \to (x_2,y_2) \to (x_2 - a,y_2) \to (x_1,y_1 + a) \to (x_1,y_1).
\]
The hexagon \( H(P;Q;a) \) is illustrated in Figure 7.

![Figure 7](image-url)

**Figure 7.** The left picture represents the hexagonal domain \( H(P;Q;a) \) with \( P = (x_1,y_1) \) and \( Q = (x_2,y_2) \) and a path \( \pi \in \Pi(P;Q;a) \). The right figure depicts the hexagonal domains \( H_N^i = H(P_N^i;Q_N^i;K - 1) \) for \( i = 1,\ldots,k \), where \( M = \lfloor N/2 \rfloor \), \( K = \lfloor M^{3/4} \rfloor \) and \( k = \lfloor N^{1/8} \rfloor \).

If \( P = (m_1,n_1), Q = (m_2,n_2) \in \mathbb{Z}^2 \) are such that \( (m_1,n_1) \leq (m_2,n_2) \) and \( a \in [0,\infty) \) we let
\[
\Pi(P;Q;a) = \Pi(m_1,n_1;m_2,n_2;a) := \{ \pi \in \Pi(m_1,n_1;m_2,n_2) : V(\pi) \subset H(P;Q;a) \},
\]
where we recall that \( \Pi(m_1,n_1;m_2,n_2) \) was defined in Section 4 and was the set of up-right paths connecting \( (m_1,n_1) \) and \( (m_2,n_2) \), and we have written \( V(\pi) \) for the set of vertices in \( \mathbb{Z}^2 \) that are visited by \( \pi \). We also define the modified partition function
\[
Z(P;Q;a) = Z(m_1,n_1;m_2,n_2;a) := \sum_{\pi \in \Pi(m_1,n_1;m_2,n_2;a)} \prod_{(i,j) \in \pi} w_{i,j}.
\]
If \( \Pi(m_1, n_1; m_2, n_2; a) = 0 \) we adopt the convention that \( Z(P; Q; a) = 0 \).

With the above notation in place we can proceed with the construction of \( Y_N \). For \( N \in \mathbb{N} \) we denote for simplicity \( M = \lfloor N/2 \rfloor \), \( K = \lfloor M^{3/4} \rfloor \) and \( k = \lfloor N^{1/8} \rfloor \). We also define the points \( P_N^i, Q_N^i \) and hexagonal domains \( H_N^i \) for \( i = 1, \ldots, k \) via

\[
(5.4) \quad P_N^i = (1, 1 + 2K \cdot (i - 1)), \quad Q_N^i = (M, M + 2K \cdot (i - 1)), \quad \text{and} \quad H_N^i = H(P_N^i; Q_N^i; K - 1).
\]

The latter domains are depicted in Figure 7. Finally, we define

\[
(5.5) \quad Y_N = \max_{i=1,\ldots,k} \log Z_N^i, \quad \text{with} \quad Z_N^i = Z(P_N^i; Q_N^i; K - 1).
\]

This specifies the choice of the \( Y_N \), for which we will establish Proposition 5.1.

Let us explain some of the ideas behind the above choice of \( Y_N \). The idea is to find a large collection of \( k \) i.i.d. positive random variables \( Z_N^i \) (\( k \geq N^\alpha \) for some \( \alpha > 0 \) will suffice), such that

1. \( \log Z_N^i \) are of order \( N^{1/3} \);
2. for large \( x \) and some \( \gamma > 0 \) we have \( \mathbb{P}(N^{-1/3} \log Z_N^i \geq x) \geq e^{-\gamma x^{3/2}} \); and
3. \( F_N \geq \log Z_N \) a.s. for all \( i = 1, \ldots, k \).

Once we have found such a collection, we should be able to show that \( Y_N := \max_{i=1,\ldots,k} \log Z_N^i \) is a lower bound for \( F_N \) and \( Y_N \) is bounded away from zero by \( cN^{1/3}(\log N)^{2/3} \) with high probability where the constant \( c \) would depend on \( \alpha \) and \( \gamma \) above. The improvement from \( N^{1/3} \) to \( N^{1/3}(\log N)^{2/3} \) comes from the independence of the \( Z_N^i \), the fact that \( k \geq N^\alpha \) and the upper tail lower bound in condition (2) above.

Since \( \exp(F_N) \) is the maximal point-to-point partition function, it makes sense to have our choice of \( Z_N^i \) be related to some type of partition functions of the polymer. A naïve guess is to choose \( Z_N^i = Z(P_N^i; Q_N^i) \) for \( i = 1, \ldots, k \), where we recall that \( Z(P_N^i; Q_N^i) \) are the partition functions from \( P_N^i \) to \( Q_N^i \). Such a choice clearly satisfies condition (3). In view of Proposition 5.1, we have that \( \log Z(P_N^i; Q_N^i) / M^{1/3} \sigma_{\text{GUE}}(1) \) weakly converge to \( F_{\text{GUE}} \) and so condition (1) is also satisfied. One also expects (although it does not follow from weak convergence alone) that condition (2) is satisfied in light of the tail behavior of \( F_{\text{GUE}} \). Specifically, from [DV13] Theorem 1 we have that there exists a constant \( \lambda > 0 \) such that for all \( x \geq 1 \)

\[
(5.6) \quad \exp \left( -\frac{4}{3} x^{3/2} + \lambda \log(x + 1) \right) \geq 1 - F_{\text{GUE}}(x) \geq \exp \left( -\frac{4}{3} x^{3/2} - \lambda \log(x + 1) \right),
\]

which suggests that one might be able to establish a similar upper tail exponent of \( 3/2 \) for \( \log Z(P_N^i; Q_N^i) \), which would yield (2).

The main issue with the above proposal for \( Z_N^i \) is that the resulting variables are not independent. Indeed, the \( Z(P_N^i; Q_N^i) \) and \( Z(P_N^j; Q_N^j) \) for \( i \neq j \) are summations over paths that may pass through the same vertices and so the same \( w_{m,n} \) appear in the formulas for both partition functions. The key idea is then to restrict the summation from all up-right paths that connect \( P_N^i \) to \( Q_N^i \) to ones that belong in narrow enough diagonal strips - these are the \( H_N^i \)'s. The resulting functions are precisely \( Z(P_N^i; Q_N^i; K - 1) \) and we take \( Z_N^i = Z(P_N^i; Q_N^i; K - 1) \). Since \( Z_N^i \) involve \( w_{m,n} \) with \((m,n) \in H_N^i\) and the latter are pairwise disjoint we indeed have that \( Z_N^i \) are i.i.d.

At this point there are two main issues that need to be addressed. In order for our strategy to be successful we must show that the following conditions hold:

I. \( Z(P_N^i; Q_N^i; K - 1) \) is close to \( Z(P_N^i; Q_N^i) \);
II. \( \log Z(P_N^1; Q_N^1) \) has an upper tail exponent of \( 3/2 \).

Based on a version of the fact that the transversal fluctuations for the log-gamma polymer are of order \( N^{2/3} \), we expect that statement I is true provided that \( K \) is much bigger than \( N^{2/3} \) and that is why we choose \( K = \lfloor M^{3/4} \rfloor \sim N^{3/4} \). A precise formulation of statement I is given in Lemma 5.4 below. The form of statement II that we establish is given in Lemma 5.4.
Lemma 5.3. Let $\theta = \theta_c$ as in (1.6). There exists $M_0 \in \mathbb{N}$ such that for $M \geq M_0$ and $K = \lceil M^{3/4} \rceil$
\[ \mathbb{P} \left( Z(1,1; M, M) - Z(1,1; M, M; K - 1) \geq e^{-M^{1/2}h_{\theta_c}(1)/4} \right) \leq e^{-M^{1/6}}, \]
where we recall that $h_{\theta}(x)$ is as in (3.2).

Lemma 5.4. Let $\theta = \theta_c$ as in (1.6). For any $\alpha \in (0,1/3)$ there is $M_0 \in \mathbb{N}$ such that for $M \geq M_0$
\[ \mathbb{P} \left( \log Z(1,1; M, M) \geq \sigma_{\theta_c}(1) \cdot M^{1/3} \cdot (\alpha \log M)^{2/3} \right) \geq M^{-\alpha}, \]
where we recall that $\sigma_{\theta_c}(x)$ is as in (3.3).

Remark 5.5. Neither Lemma 5.3 nor Lemma 5.4 are optimal though they will suffice for proving Proposition 5.1. Lemma 5.3 is proved in Section 5.3 and is essentially a consequence of Proposition 3.2 and Lemma 3.10. Lemma 5.4 is proved below by utilizing Proposition 3.3.

Proof of Lemma 5.4. Let us fix $\alpha \in (0,1/3)$ and $\epsilon_0 \in (\alpha,1/3)$. From (3.4) and (3.11) we have
\[ \sigma_{\theta_c}(1) \cdot M^{1/3} \cdot F(M, M) = \log Z(1,1; M, M), \]
where we used that $\theta = \theta_c$. The latter and Proposition 3.3 applied to $\theta = \theta_c$, $\delta = 1$ and $\epsilon_0$ as above implies that there exists $M_1, C_1 > 0$ such that for $M \geq M_1$ we have
\[ \left| \mathbb{P} \left( \log Z(1,1; M, M) \geq \sigma_{\theta_c}(1) \cdot M^{1/3} \cdot (\alpha \log M)^{2/3} \right) - (1 - F_{\text{GUE}} \left( (\alpha \log M)^{2/3} \right) \right) \leq C_1 M^{-\epsilon_0}. \]
On the other hand, by (5.6) we have for some $\lambda > 0$ and all large $M$
\[ 1 - F_{\text{GUE}} \left( (\alpha \log M)^{2/3} \right) \geq \exp \left( -\frac{4\alpha}{3} \log M - \lambda \log \left( (\alpha \log M)^{2/3} + 1 \right) \right) \geq \frac{2}{M^\alpha}. \]
Combining the last two inequalities we see that for all large $M$
\[ \mathbb{P} \left( \log Z(1,1; M, M) \geq \sigma_{\theta_c}(1) \cdot M^{1/3} \cdot (\alpha \log M)^{2/3} \right) \geq 2M^{-\alpha} - C_1 M^{-\epsilon_0}, \]
which implies the lemma as $\epsilon_0 > \alpha$. \hfill \Box

The above discussion aimed to explain the origin of the choice of $Y_N$ in (5.5). The only quantity we did not discuss is $k = \lfloor N^{1/8} \rfloor$. We make this choice to ensure that $H_{1;N}^i$ all belong to $[1,N]^2$ for $i = 1, \ldots, k$. We mention that one could make different choices for $K$ and $k$, which might produce a higher constant than $20^{-1} \cdot \sigma_{\theta_c}(1)$ in (5.1) – we have not attempted to optimize over such choices.

We now proceed to show that the $Y_N$ as in (5.5) satisfy Proposition 5.1.

Proof of Proposition 5.1. We prove the proposition with $Y_N$ as in (5.5). We also continue with the same notation as in the beginning of the section. By definition, we have for all large $N$ that $P_N^i, Q_N^i \in [1,N]^2$ for $i = 1, \ldots, k$ and so
\[ F_N \geq \max_{i=1,\ldots,k} \log Z(P_N^i; Q_N^i) \geq \max_{i=1,\ldots,k} \log Z(P_N^i; Q_N^i; K - 1) = Y_N, \]
which implies the first part of the proposition.

Let us set $R = 20^{-1} \cdot \sigma_{\theta_c}(1)$ for simplicity. Since $Z(P_N^i; Q_N^i; K - 1)$ depends only on $w_{m,n}$ for $(m,n) \in H_{1;N}$ as in (5.2), which are disjoint for all large $N$, we conclude that $\log Z(P_N^i; Q_N^i; K - 1)$ are i.i.d. random variables with the same distribution as $\log Z(P_N^1; Q_N^1; K - 1)$. The latter implies
\[ \mathbb{P} \left( Y_N \geq RN^{1/3}(\log N)^{2/3} \right) = \mathbb{P} \left( \max_{i=1,\ldots,k} \log Z(P_N^i; Q_N^i; K - 1) \geq RN^{1/3}(\log N)^{2/3} \right) = \]
\[ 1 - \mathbb{P} \left( \log Z(P_N^1; Q_N^1; K - 1) < RN^{1/3}(\log N)^{2/3} \right)^k. \]
From Lemma 5.4 (applied to $\alpha = 1/10$ and $M = \lfloor N/2 \rfloor$) we have for all large $N$ that
\[ \mathbb{P} \left( \log Z(P_N^1; Q_N^1) < 2RN^{1/3}(\log N)^{2/3} \right) \leq 1 - M^{-1/10}. \]
The latter inequality and Lemma [5.3] imply that for all large $N$
\[
\mathbb{P} \left( \log Z(P_N^1; Q_N^1; K - 1) < RN^{1/3}(\log N)^{2/3} \right) \leq 1 - (1/2)M^{-1/10}.
\]

Substituting the last inequality into (5.7) we obtain
\[
\mathbb{P} \left( \mathcal{Y}_N \geq RN^{1/3}(\log N)^{2/3} \right) \geq 1 - \left( 1 - (1/2)M^{-1/10} \right)^k,
\]
which implies the second part of (5.1) since $k = \lfloor N^{1/8} \rfloor$ and $M = \lfloor N/2 \rfloor$.

\[\square\]

5.3. Proof of Lemma [5.3] For clarity we split the proof of the lemma into three steps.

**Step 1.** We begin by introducing some notation. Let $H$ denote the hexagonal domain $H(P; Q; K - 1)$ for $P = (1, 1)$ and $Q = (M, M)$ as in Definition 5.2. For $M \geq 2$ we define
\[
\partial H^{\text{top}} := \{(m, n) \in \mathbb{Z}^2 : (m, n - 1) \in H \text{ and } n - m = K\},
\]
\[
\partial H^{\text{bot}} := \{(m, n) \in \mathbb{Z}^2 : (m - 1, n) \in H \text{ and } m - n = K\}.
\]
In words, $\partial H^{\text{top}}$ is the set of points $\mathbb{Z}^2$ that lie on the segment connecting $(1, K + 1)$ and $(M - K, M)$, and $\partial H^{\text{bot}}$ is the set of points $\mathbb{Z}^2$ that lie on the segment connecting $(K + 1, 1)$ and $(M, M - K)$.

For every $(m, n) \in \partial H^{\text{top}} \cup \partial H^{\text{bot}}$ we let $(m^*, n^*)$ denote $(m, n - 1)$ if $(m, n) \in \partial H^{\text{top}}$ and $(m, n)$ if $(m, n) \in \partial H^{\text{bot}}$. We claim that there exist $M_1, D_1, d_1 > 0$ such that if $M \geq M_1$ and $(m, n) \in \partial H^{\text{top}} \cup \partial H^{\text{bot}}$ we have
\[
\mathbb{P}(E_M) \leq D_1 \exp \left( -d_1 M^{3/16} \right), \quad \text{with}
\]
\[
E_M = \left\{ Z(1, 1; m^*, n^*) \cdot Z(m, n; M, M) \geq e^{-M^{1/2}h_{\delta_c}(1)/3} \right\}.
\]
We prove (5.8) in the steps below. Here we assume its validity and conclude the proof of the lemma.

From the definition of $Z(1, 1; M, M; K - 1)$ in [5.3] we have that
\[
Z(1, 1; M, M; K - 1) = \sum_{\pi \in \Pi(m_1, n_1; m_2, n_2) \setminus \Pi(m_1, n_1; m_2, n_2; K - 1)} \prod_{(i, j) \in \pi} w_{i,j}.
\]
Notice that each of the paths $\pi \in \Pi(m_1, n_1; m_2, n_2) \setminus \Pi(m_1, n_1; m_2, n_2; K - 1)$ needs to contain at least one vertex in $\partial H^{\text{top}} \cup \partial H^{\text{bot}}$. Splitting the sum over the first vertex $(m, n)$ in $\pi$ that belongs to $\partial H^{\text{top}} \cup \partial H^{\text{bot}}$ we obtain the upper bound
\[
Z(1, 1; M, M) - Z(1, 1; M, M; K - 1) \leq \sum_{(m, n) \in \partial H^{\text{top}} \cup \partial H^{\text{bot}}} Z(1, 1; m^*, n^*) \cdot Z(m, n; M, M).
\]
The latter inequality and union bound of (5.8) gives
\[
\mathbb{P} \left( Z(1, 1; M, M) - Z(1, 1; M, M; K - 1) \geq 2Me^{-M^{1/2}h_{\delta_c}(1)/3} \right) \leq 2MD_1 \exp \left( -d_1 M^{3/16} \right),
\]
where we used that $|\partial H^{\text{top}} \cup \partial H^{\text{bot}}| \leq 2M$. The latter inequality implies the statement of the lemma, since $h_{\delta_c}(1) > 0$ by (3.11) and $3/16 > 1/6$.

**Step 2.** By the distributional equality of $Z(m_1, n_1; m_2, n_2)$ and $Z(n_1, m_1; n_2, m_2)$ it suffices to prove (5.8) when $(m, n) \in \partial H^{\text{bot}}$. Furthermore, by the distributional equality of $Z(m_1, n_1; m_2, n_2)$ and $Z(-m_2, -n_2; -m_1, -n_1)$ we only need to consider the case when $m = K + n$ and $1 \leq n \leq M/2 - K + 1$ (i.e. the left half of $\partial H^{\text{bot}}$). In particular, we have $m^* = m - 1$ and $n^* = n$.

In this step we prove (5.8) when $m = K + n$ and $1 \leq n \leq M/2 - K + 1$ and $n/(m - 1) \geq 1/2$. In this case we have from Proposition 5.2 (applied to $\delta = 1/2, x = M^{1/8}$) that there are constants
\(C_1, c_1 > 0\) such that for all large \(M\) we have
\[
P(E^1_M) \leq C_1 e^{-c_1 M^{3/16}}, \quad \text{with}
\]
\[
E^1_M = \left\{ \frac{\log Z(1,1; m^*, n^*) + (m - 1)h_{\theta_e}(n/(m - 1))}{(m - 1)^{1/3}\sigma_{\theta_e}(n/(m - 1))} \geq M^{1/8} \right\},
\]
where we used that \(m^* = m - 1, \ n^* = n, \ m \geq K\) and \(K = |M^{3/4}|\). By the same proposition we also know that there constants \(C_2, c_2 > 0\) such that for all large \(M\) we have
\[
P(E^2_M) \leq C_2 e^{-c_2 M^{3/16}}, \quad \text{with}
\]
\[
E^2_M = \left\{ \frac{\log Z(m, n; M, M) + (M - n + 1)h_{\theta_e}(M/m + 1)}{(M - n + 1)^{1/3}\sigma_{\theta_e}(M/m + 1)} \geq M^{1/8} \right\},
\]
where we also used the distributional equality of \(Z(m_1, n_1; m_2, n_2)\) and \(Z(n_1, m_1; n_2, m_2)\).

In the remainder of this step we prove that for all large \(M\)
\[
M^{1/8}(m - 1)^{1/3}\sigma_{\theta_e}(n/(m - 1)) - (m - 1)h_{\theta_e}(n/(m - 1)) +
\]
\[
M^{1/8}(M - n + 1)^{1/3}\sigma_{\theta_e}\left(\frac{M - m + 1}{M - n + 1}\right) - (M - n + 1)h_{\theta_e}\left(\frac{M - m + 1}{M - n + 1}\right) \leq -\frac{M^{1/2}h''_{\theta_e}(1)}{3}.
\]

We recall from Section 3.1 that \(\sigma_{\theta_e}(x)\) is a smooth positive function on \((0, \infty)\) and both \(\frac{n}{m-1}\) and \(\frac{M-m+1}{M-n+1}\) belong to \([1/2, 1]\). This means that we can find a constant \(A_1 > 0\) such that
\[
M^{1/8}(m - 1)^{1/3}\sigma_{\theta_e}(n/(m - 1)) \leq A_1 M^{11/24}
\]
and
\[
M^{1/8}(M - n + 1)^{1/3}\sigma_{\theta_e}\left(\frac{M - m + 1}{M - n + 1}\right) \leq A_1 M^{11/24},
\]
where we also used that \(m \leq M\) and \(M - n + 1 \leq M\).

Using that \(m = n + K\), (5.11) and (5.12) we get
\[
(M - n + 1)h_{\theta_e}\left(\frac{M - m + 1}{M - n + 1}\right) = \frac{h''_{\theta_e}(1)K^2}{2(M - n + 1)} + O(M^{1/4}) \text{ and } (m - 1)h_{\theta_e}\left(\frac{n}{m - 1}\right) \geq 0.
\]
Combining (5.12) and (5.13) we get for all large \(M\)
\[
M^{1/8}(m - 1)^{1/3}\sigma_{\theta_e}(n/(m - 1)) - (m - 1)h_{\theta_e}(n/(m - 1)) + M^{1/8}(M - n + 1)^{1/3}\sigma_{\theta_e}\left(\frac{M - m + 1}{M - n + 1}\right)
\]
\[
- (M - n + 1)h_{\theta_e}\left(\frac{M - m + 1}{M - n + 1}\right) \leq (2A_1 + 1)M^{11/24} - \frac{h''_{\theta_e}(1)K^2}{2(M - n + 1)}.
\]
Since \(K = |M^{3/4}|\) and \(h''_{\theta_e}(1) > 0\) we see that the last inequality implies (5.11).

Step 3. In this step we prove (5.8) when \(m = K + n + 1 \leq n \leq |M/2| - K + 1\) and \(n/(m - 1) \in (0, 1/2)\). Let us set \(\epsilon = \min(1, h_{\theta_e}(1/2))\), which is positive in view of (3.12). From Lemma 3.10 we have that there are positive constants \(B_1, b_1 > 0\) such that
\[
P(E^3_M) \leq B_1 e^{-b_1 M^{3/4}}, \quad \text{with}
\]
\[
E^3_M = \left\{ \log Z(1,1; m^*, n^*) + (m - 1)h_{\theta_e}\left(\frac{n}{m - 1}\right) - \epsilon \geq 0 \right\}.
\]
In addition, since \(1 \leq n \leq |M/2| - K + 1\), we have that (5.10) also holds for all large enough \(M\).

In the remainder of this step we prove that for all \(M\) sufficiently large we have that \(E_M \subset E^2_M \cup E^3_M\), which in view of (5.10) and (5.14), would imply (5.5).
We will show that $E_M \subset E_M^2 \cup E_M^3$ by demonstrating that $(E_M^2)^c \cap (E_M^3)^c \subset E_M^c$. The latter statement would follow if we can show that

$$-(m-1)(h_{\theta_\epsilon}(n/(m-1)) - \epsilon) + M^{1/8}(M-n+1)^{1/3} \sigma_{\theta_\epsilon} \left( \frac{M-m+1}{M-n+1} \right) - \frac{M^{1/2}h''_{\theta_\epsilon}(1)}{3}.$$

(5.15)

In view of (3.12) and the definition of $\epsilon$ we have

$$h_{\theta_\epsilon}(n/(m-1)) - \epsilon \geq 0.$$

Combining the latter with the second line of (5.12) and the first equality in (5.13) we conclude

$$-(m-1)(h_{\theta_\epsilon}(n/(m-1)) - \epsilon) + M^{1/8}(M-n+1)^{1/3} \sigma_{\theta_\epsilon} \left( \frac{M-m+1}{M-n+1} \right) - \frac{M^{1/2}h''_{\theta_\epsilon}(1)}{2} \cdot \frac{K^2}{(M-n+1)},$$

which implies (5.15) as $K = [M^{3/4}]$ and $h''_{\theta_\epsilon}(1) > 0$. This concludes the proof of the lemma.

6. Proof of Theorem 1.4

In this section we prove Theorem 1.4. The theorem follows immediately from Lemma 6.1 (which provides a lower bound) and Lemma 6.2 (which provides an upper bound). The lower bound relies on the simple observation that $F_N$, as in (1.4), cannot be smaller than the maximum of the variables $\log w_{i,j}$ for $i, j \in [1, N]$. To prove the upper bound, we will use Lemma 3.11. In this section we continue with the notation of Sections 1 and 3.

6.1. Lower bound. The following result proves the lower bound in Theorem 1.4 for any $c_1 < 2/\theta$.

**Lemma 6.1.** Fix $\theta > 0$. For any $\epsilon > 0$,

$$\lim_{N \to \infty} \mathbb{P}(F_N \geq (2/\theta - \epsilon) \log N) = 1.$$

**Proof.** Recall the definition of $F_N$ in (1.4). By restricting the maximum over $(m_1, n_1) = (m_2, n_2)$, we have that $F_N \geq \log w_N^{\text{max}}$ where $w_N^{\text{max}} = \max_{1 \leq m,n \leq N} w_{m,n}$. The $w_{m,n}$ form a collection of $N^2$ i.i.d. inverse-gamma random variables. Thus, for any $c > 0$, we have

$$\mathbb{P}(F_N \geq c \log N) \geq \mathbb{P}(\log w_{N}^{\text{max}} \geq c \log N) = 1 - \mathbb{P}(w_{1,1} \leq N^c)^N.$$

We have

$$\mathbb{P}(w_{1,1} \leq N^c) = 1 - \frac{1}{\Gamma(\theta)} \int_{N^c}^{\infty} x^{-\theta-1} \exp(-x^{-1})dx \leq 1 - \frac{\exp(-N^{-c}) N^{-c\theta}}{\theta \Gamma(\theta)}.$$

Using the bound $\log(1-x) \leq -x$ for $x \in (0, 1)$, we find that for any $0 < c < 2/\theta$,

$$\mathbb{P}(w_{1,1} \leq N^c)^N \leq \exp \left( - \frac{N^2-c\theta e^{-N^{-c}}}{\theta \Gamma(\theta)} \right) \xrightarrow{N \to \infty} 0.$$

Combining the last statement with (6.1), we find that for any $0 < c < 2/\theta$,

$$\lim_{N \to \infty} \mathbb{P}(F_N \geq c \log N) = 1,$$

which proves the lemma. \qed
6.2. Upper bound. The following result proves the upper bound in Theorem 1.4.

**Lemma 6.2.** Let \( \theta_c \) be as in [1.6] and fix \( \theta > \theta_c \). There exists a constant \( c_2 \) (depending on \( \theta \)) with
\[
\lim_{N \to \infty} \mathbb{P} ( F_N \leq c_2 \log N ) = 1.
\]

**Proof.** From Lemma 3.11 with \( a = 5 \), \( M = N \) and \( \theta \) as in the statement of the present lemma, we know that there are constants \( M_2, D_2 > 0 \) such that for \( N \geq M_2 \) and \( m, n \in [1, N] \) we have
\[
\mathbb{P}(\log Z(1, 1; m, n) \geq D_2 \log N) \leq N^{-5}.
\]
Using the distributional equality of \( Z(m_1, n_1; m_2, n_2) \) and \( Z(m_1 + a, n_1 + b; m_2 + a, n_2 + b) \) for any \( a, b \in \mathbb{Z} \) we conclude that for any \( (m_1, n_1), (m_2, n_2) \in [1, N]^2 \) with \( (m_1, n_1) \leq (m_2, n_2) \) we have
\[
\mathbb{P}(\log Z(m_1, n_1; m_2, n_2) \geq D_2 \log N) \leq N^{-5}.
\]
Taking a union bound of the above (there are at most \( N^2 \) choices for \( m_1, n_1, m_2, n_2 \)) we get
\[
\mathbb{P}(F_N \geq D_2 \log N) \leq N^{-1}.
\]
The latter implies the statement of the lemma with \( c_2 = D_2 \). \( \square \)

**Remark 6.3.** In the case \( \theta > 3 \), the constant \( c_2 \) in Lemma 6.2 can be chosen as any constant \( c_2 > 2 \). Indeed, from the definition of \( F_N \) and (3.33) we have
\[
\mathbb{E} \left[ e^{F_N} \right] \leq \sum_{m_1, n_1, m_2, n_2 \in [1, N]} \mathbb{E} \left[ Z(m_1, n_1; m_2, n_2) \right] \leq \sum_{m_1, n_1, m_2, n_2 \in [1, N]} \left( \frac{2}{\theta - 1} \right)^{m_2 + n_2 - m_1 - n_1 + 1}.
\]
The last inequality implies
\[
\mathbb{E} \left[ e^{F_N} \right] \leq \sum_{m_1 = 1}^{N} \sum_{n_1 = 1}^{N} \sum_{m_2 = m_1}^{\infty} \sum_{n_2 = n_1}^{\infty} \left( \frac{2}{\theta - 1} \right)^{m_2 + n_2 - m_1 - n_1 + 1} = \frac{2(\theta - 1)N^2}{(\theta - 3)^2},
\]
where the last equality used the formula for a geometric series, and it is here we use our assumption that \( \theta > 3 \). Thus, by Chebyshev’s inequality, for any \( \varepsilon > 0 \) we have
\[
P ( F_N \geq (2 + \varepsilon) \log N ) \leq \mathbb{P} \left( e^{F_N} \geq N^{2+\varepsilon} \right) \leq N^{-2-\varepsilon} \cdot \mathbb{E} \left[ e^{F_N} \right] \xrightarrow{N \to \infty} 0,
\]
which proves the Lemma 6.2 for any \( c_2 > 2 \).

7. Proofs of Propositions 2.2, 2.3 and 2.4

In this section we prove Propositions 2.2, 2.3 and 2.4 about the log-gamma polymer measure from Section 2.1. Throughout this section we use freely notation from Sections 1.1, 2.1 and 3 above.

7.1. Proof of Proposition 2.2. Throughout this section we assume that \( \theta \in (0, \theta_c) \).

We recall from [1.7] for \( \delta \in (0, 1/3) \) the following corner sets
\[
SW_N^\delta := \mathbb{N}^2 \cap [1, N^{1/3+2\delta}]^2 \quad \text{and} \quad NE_N^\delta := \mathbb{N}^2 \cap [N - N^{1/3+2\delta}, N]^2.
\]
We claim that for each \( \delta \in (0, 1/3) \) we have that
\[
\lim_{N \to \infty} \mathbb{P} \left( (X^\delta_N, Y^\delta_N) \in SW_N^\delta \right) = 1 \quad \text{and} \quad \lim_{N \to \infty} \mathbb{P} \left( (X^\delta_N, Y^\delta_N) \in NE_N^\delta \right) = 1.
\]
We prove (7.2) below. For now, we assume its validity and conclude the proof of Proposition 2.2.

Fix \( \varepsilon > 0 \) as in the statement of the proposition. From (7.2) with \( \delta = \min(1/6, \varepsilon/3) \) we immediately conclude that the first four sequences of variables in (2.3) converge to 0 in probability as
\( N \to \infty \). To illustrate our point, let us focus on \( \frac{X_N^e}{N^{1/3+\epsilon}} \) and notice that for any \( a > 0 \) and \( N \) large enough so that \( aN^\epsilon > N^{2\delta} \)

\[
P\left( \frac{|X_N^e|}{N^{1/3+\epsilon}} > a \right) = P\left( X_N^e > aN^{1/3+\epsilon} \right) \leq P\left( X_N^e > N^{1/3+2\delta} \right) \leq 1 - P\left( (X_N^e, Y_N^e) \in SW_N^\delta \right).
\]

Combining the latter with (7.2) and taking \( N \to \infty \) we conclude that \( \frac{X_N^e}{N^{1/3+\epsilon}} \) converges to 0 in probability. An analogous argument shows that \( \frac{Y_N^e - N}{N^{1/3+\epsilon}}, \frac{Y_N^e - N}{N^{1/3+\epsilon}} \) all converge to 0 in probability. Since \( L_N = X_N^e + Y_N^e - Y_N^e - Y_N^e \) we also conclude that

\[
\frac{L_N - 2N}{N^{1/3+\epsilon}} = \frac{X_N^e - N}{N^{1/3+\epsilon}} + \frac{Y_N^e - N}{N^{1/3+\epsilon}} - \frac{Y_N^e}{N^{1/3+\epsilon}}
\]

also converges to 0 in probability. Finally, as \( S_N = \frac{Y_N^e - Y_N^e + 1}{X_N^e - X_N^e + 1} \), we have

\[
\frac{N(S_N - 1)}{N^{1/3+\epsilon}} = \frac{(Y_N^e - N) - Y_N^e - (X_N^e - N) + X_N^e}{N^{1/3+\epsilon}} \cdot \frac{N^{2/3-\delta}}{N^{2/3-\delta} + N^{-1/3-\delta} ((X_N^e - N) - X_N^e + 1)}.
\]

From our work so far, the first term in the above product converges to 0 in probability and the second one (from our work with \( \epsilon \) replaced with \( \delta \)) converges to 1, and so the product converges to 0 in probability as desired. This proves that all sequences of variables in (2.3) converge to 0 in probability as \( N \to \infty \).

In the remainder of this section we prove (7.2). From Definitions 1.1 and 2.1 we know that for each \((m_1, n_1), (m_2, n_2) \in [1, N]^2\) we have

\[
P(X_N^e = m_1, Y_N^e = n_1, X_N^e = m_2, Y_N^e = n_2) = E\left[ Z(m_1, n_1; m_2, n_2) \right].
\]

Let \( S_N^\delta \) be the set of quadruples \((m_1, n_1, m_2, n_2) \in [1, N]\) such that \((m_1, n_1) \leq (m_2, n_2)\) and \(((m_1, n_1), (m_2, n_2)) \notin SW_N^\delta \times EW_N^\delta \) and note \(|S_N^\delta| \leq N^4\). To prove (7.2) it suffices to show that

\[
\lim_{N \to \infty} P\left( (X_N^e, Y_N^e, X_N^e, Y_N^e) \in S_N^\delta \right) = 0.
\]

From (4.9) we know that there exists \( N_0 > 0 \) (depending on \( \theta \) and \( \delta \)) such that for all \( N \geq N_0 \) and \((m_1, n_1, m_2, n_2) \in S_N^\delta \) we have

\[
P\left( \log Z(m_1, n_1; m_2, n_2) \geq -2N\Psi(\theta/2) - N^{1/3} \log N \right) \leq N^{-5}.
\]

In addition, from Proposition 3.1 we know that

\[
\lim_{N \to \infty} P\left( E_N \right) = 0, \text{ where } E_N = \{ \log Z(1, 1; N, N) \leq -2N\Psi(\theta/2) - (1/2)N^{1/3} \log N \}.
\]

Let us introduce the event

\[
F_N = \cup_{(m_1, n_1, m_2, n_2) \in S_N^\delta} \{ \log Z(m_1, n_1; m_2, n_2) \geq -2N\Psi(\theta/2) - N^{1/3} \log N \}
\]

and note that by union bound and (7.5) we have for all \( N \geq N_0 \)

\[
P(F_N) \leq |S_N^\delta| \cdot N^{-5} \leq N^{-1}.
\]
Combining \((7.3)\), \((7.6)\) and \((7.7)\), we get

\[
\limsup_{N \to \infty} \mathbb{P}(X_N^*, Y_N^*, X_N^c, Y_N^c) \in S_N^\delta) \leq \\
\limsup_{N \to \infty} \mathbb{P}\left( (X_N^*, Y_N^*, X_N^c, Y_N^c) \in S_N^\delta \cap E_N^c \cap F_N^c \right) + \limsup_{N \to \infty} \mathbb{P}(E_N) + \limsup_{N \to \infty} \mathbb{P}(F_N) = \\
\limsup_{N \to \infty} \mathbb{P}\left( (X_N^*, Y_N^*, X_N^c, Y_N^c) \in S_N^\delta \cap E_N^c \cap F_N^c \right) = \\
\limsup_{N \to \infty} \mathbb{E}\left[ \frac{1}{Z_N^w} \sum_{(m_1, n_1, m_2, n_2) \in S_N^\delta} \frac{Z(m_1, n_1; m_2; n_2)}{Z_N^w} \cdot 1_{E_N^c} \cdot 1_{F_N^c} \right].
\]

(7.8)

On the event \(E_N^c\) we have

\[
Z_N^w \geq Z(1, 1; N, N) \geq \exp(-2N \Psi(\theta/2) - (1/2)N^{1/3} \log N).
\]

and on the event \(F_N^c\) we have for \((m_1, n_1, m_2, n_2) \in S_N^\delta\)

\[
Z(m_1, n_1; m_2; n_2) \leq \exp(-2N \Psi(\theta/2) - N^{1/3} \log N).
\]

Combining the latter two statements, we see

\[
\limsup_{N \to \infty} \mathbb{E}\left[ \frac{1}{Z_N^w} \sum_{(m_1, n_1, m_2, n_2) \in S_N^\delta} \frac{Z(m_1, n_1; m_2; n_2)}{Z_N^w} \cdot 1_{E_N^c} \cdot 1_{F_N^c} \right] \leq \\
\limsup_{N \to \infty} \mathbb{E}\left[ \frac{1}{Z_N^w} \cdot 1_{E_N^c} \cdot 1_{F_N^c} \right] \leq \limsup_{N \to \infty} N^4 \cdot e^{-(1/2)N^{1/3} \log N} = 0,
\]

which together with \((7.3)\) establishes \((7.4)\).

7.2. Proof of Proposition 2.3. Throughout this section we assume that \(\theta = \theta_c\).

From Theorem 1.3 we know that there exists a constant \(\rho > 0\) such that

\[
\lim_{N \to \infty} \mathbb{P}(A_N) = 0 \quad \text{where} \quad A_N = \{F_N \leq \rho N^{1/3}(\log N)^{2/3}\},
\]

where we recall that \(F_N\) is the maximal point-to-point free energy as in \((1.4)\).

Let \(M_1, D_1 > 0\) be as in Lemma 3.11 for \(\theta = \theta_c\) and \(a = 5\). In addition, suppose that \(\delta \in (0, 1)\) is small enough so that \(\delta^{1/3} D_1 < \rho/2\). We first show that

\[
\limsup_{N \to \infty} \mathbb{P}(L_N \leq \delta N) = 0,
\]

which would imply \((2.4)\) with \(c_1 = \delta\) once we recall that \(L_N \leq 2N - 2\) by definition.

Let \(M_2 \in \mathbb{N}\) be sufficiently large so that \(N_{N_\delta} := \lfloor \delta N \rfloor + 1 \geq M_1\) for all \(N \geq M_2\). From Lemma 3.11 we have for \(N \geq M_2\) and \((m, n) \in [1, N_{\delta}]^2\) that

\[
\mathbb{P}\left( \log Z(1, 1; m, n) \geq D_1 N_{\delta}^{1/3} \cdot (\log N_{\delta})^{2/3} \right) \leq N_{\delta}^{-5}.
\]

The latter and the tower of inequalities

\[
D_1 N_{\delta}^{1/3} \cdot (\log N_{\delta})^{2/3} \leq D_1 \delta^{1/3} N^{1/3} \cdot (\log N_{\delta})^{2/3} \leq (\rho/2) N^{1/3} (\log N_{\delta})^{2/3} \leq (\rho/2) N^{1/3} (\log N)^{2/3},
\]

together imply that for \(N \geq M_2\) and \((m, n) \in [1, N_{\delta}]^2\)

\[
\mathbb{P}\left( \log Z(1, 1; m, n) \geq (\rho/2) N^{1/3} \cdot (\log N)^{2/3} \right) \leq N_{\delta}^{-5}.
\]

(7.11)

Let \(T_N^\delta\) be the set of quadruples \((m_1, n_1, m_2, n_2) \in [1, N]\) such that \((m_1, n_1) \leq (m_2, n_2)\) and \(m_2 - m_1 + n_2 - n_1 \leq \lfloor \delta N \rfloor\). Note that \(Z(m_1, n_1; m_2, n_2)\) has the same distribution as \(Z(m_1 + a, n_1 + b; m_2 + a, n_2 + b)\)
Let \( \epsilon > 0 \) find positive constants \( \delta \) and \( N \). Notice that the last two statements with \( \epsilon \), \( \delta \), and \( N \) as in equations (7.13) and (7.14) establish (7.10).

Step 1. In this step we prove that equations (7.15) and (7.20), which we state and prove in the following two steps.

Combining the latter two statements, we see
\[
\limsup_{N \to \infty} \mathbb{E} \left[ \sum_{(m_1, n_1, m_2, n_2) \in T^\delta_N} Z(m_1, n_1; m_2, n_2) \cdot 1_{A_N^c} \cdot 1_{B_N^c} \right] \leq \limsup_{N \to \infty} N^4 \cdot \exp^{-N^1/3} \cdot N^{1/3} = 0,
\]
which together with (7.13) establishes (7.10).

In the remainder of this section we prove that \( \frac{N(S_N-1)}{N^2/3} \) converges to 0 in probability as \( N \to \infty \). We will prove that for any \( \delta_1 \in (0, 1/3) \)
\[
\lim_{N \to \infty} \mathbb{P} \left( S_N \leq \exp(-N^{-1/3+\delta_1}) \right) = 0 \quad \text{and} \quad \lim_{N \to \infty} \mathbb{P} \left( S_N \geq \exp(N^{-1/3+\delta_1}) \right) = 0.
\]

Notice that the last two statements with \( \delta_1 = \min(\epsilon/2, 1/4) \) imply the weak convergence of \( \frac{N(S_N-1)}{N^2/3} \) to 0 as \( N \to \infty \). Moreover, since \( S_N \) and \( 1/S_N \) have the same distribution (this follows from the distributional invariance of the log-gamma polymer upon exchanging the horizontal and vertical axes) it suffices to show the first inequality in (7.14).

Step 1. In this step we prove that
\[
\lim_{N \to \infty} \mathbb{P} \left( S_N \leq 1/2 \right) = 0.
\]

Let \( \epsilon = \min(1, h_{\theta_c}(1/2)) \), which in view of (3.12) is a positive number. From Lemma 3.10 we can find positive constants \( B, b > 0 \) such that for \( m, n \in \mathbb{N} \), \( m \geq n \)
\[
\mathbb{P}(\log Z(1, 1; m, n) \geq m \cdot \epsilon - m \cdot h_{\theta_c}(n/m)) \leq B \exp(-bm).
\]
The last equation combined with the definition of ε, and (3.12) allows us to conclude that for $m \geq n$ and $n/m \in (0, 1/2]$

\begin{equation}
B \exp(-bm) \geq \mathbb{P}(\log Z(1, 1; m, n) \geq m \cdot \epsilon - m \cdot \bar{h}_\epsilon(n/m)) \geq \mathbb{P}(\log Z(1, 1; m, n) \geq m \cdot \epsilon - m \cdot \bar{h}_\epsilon(1/2)) \geq \mathbb{P}(Z(1, 1; m, n) \geq 1).
\end{equation}

Let $U_N$ be the set of quadruples $m_1, n_1, m_2, n_2 \in [1, N]$ such that $(m_1, n_1) \leq (m_2, n_2)$ and

\[
\frac{m_2 - n_1 + 1}{m_2 - m_1 + 1} \in (0, 1/2). \quad \text{Let } \delta > 0 \text{ such that (7.10) holds (recall that we proved (7.10) earlier in this section), and let } T_p^\delta \text{ be as in the paragraph above (7.12). Finally, let } \rho > 0 \text{ and the events } A_N \text{ be as in (7.9).}
\]

Note that $Z(m_1, n_1; m_2, n_2)$ has the same distribution as $Z(m_1 + a, n_1 + b; m_2 + a, n_2 + b)$ for any $a, b \in \mathbb{Z}$ and also that if $(m_1, n_1, m_2, n_2) \in U_N$, then

\[
m_2 - m_1 + 1 \geq n_2 - n_1 + 1, \quad (n_2 - n_1 + 1)/(m_2 - m_1 + 1) \in (0, 1/2].
\]

The latter and a union bound of (7.16) implies that

\begin{equation}
\mathbb{P}(C_N) \leq B N^4 \exp(-\delta bN/2), \quad \text{and consequently } \lim_{N \to \infty} \mathbb{P}(C_N) = 0,
\end{equation}

where $C_N = \cup (m_1, n_1, m_2, n_2) \in U_N \setminus T_p^\delta \{ Z(m_1, m_1; m_2, n_2) \geq 1 \}$. We mention that in the first inequality in (7.17) we used that $|U_N \setminus T_p^\delta| \leq N^4$ and that for $(m_1, n_1, m_2, n_2) \in U_N \setminus T_p^\delta$ we have $m_2 - m_1 + 1 \geq \delta N/2$.

Using the equality of events $\{S_N \leq 1/2\} = \{(X_N^\delta, Y_N^\delta, X_N^\varepsilon, Y_N^\varepsilon) \in U_N\}$ and (7.3), we see that

\begin{equation}
\limsup_{N \to \infty} \mathbb{P}(S_N \leq 1/2) = \limsup_{N \to \infty} \mathbb{P}(\{S_N \leq 1/2\} \cap A_N \cap C_N^c) = \limsup_{N \to \infty} \mathbb{E} \left[ \frac{\sum_{(m_1, n_1, m_2, n_2) \in U_N \setminus T_p^\delta} Z(m_1, n_1; m_2, n_2)}{Z_N^w} \right] \cdot 1_{A_N^c} \cdot 1_{C_N^c}.
\end{equation}

In the first equality we used $\lim_{N \to \infty} \mathbb{P}(A_N) + \mathbb{P}(C_N) = 0$, which holds from (7.9) and (7.17).

We also know from (7.3) and (7.10) that

\[
\lim_{N \to \infty} \mathbb{E} \left[ \frac{\sum_{(m_1, n_1, m_2, n_2) \in U_N \setminus T_p^\delta} Z(m_1, n_1; m_2, n_2)}{Z_N^w} \right] = \lim_{N \to \infty} \mathbb{P}(L_N \leq \delta N) = 0.
\]

Combining the above two statements, we get

\begin{equation}
\limsup_{N \to \infty} \mathbb{P}(S_N \leq 1/2) = \limsup_{N \to \infty} \mathbb{E} \left[ \frac{\sum_{(m_1, n_1, m_2, n_2) \in U_N \setminus T_p^\delta} Z(m_1, n_1; m_2, n_2)}{Z_N^w} \right] \cdot 1_{A_N^c} \cdot 1_{C_N^c}.
\end{equation}

On the event $A_N^c$ we have

\[
Z_N^w \geq Z(1, 1; N, N) \geq \exp(F_N) \geq \exp \left( \rho N^{1/3} (\log N)^{2/3} \right).
\]

and on the event $C_N^c$ we have for $(m_1, n_1, m_2, n_2) \in U_N \setminus T_p^\delta$

\[
Z(m_1, n_1; m_2, n_2) \leq 1.
\]

Combining the latter statements, we see

\[
\limsup_{N \to \infty} \mathbb{E} \left[ \frac{\sum_{(m_1, n_1, m_2, n_2) \in U_N \setminus T_p^\delta} Z(m_1, n_1; m_2, n_2)}{Z_N^w} \right] \cdot 1_{A_N^c} \cdot 1_{C_N^c} \leq \limsup_{N \to \infty} \mathbb{P}(A_N^c) \cdot 1_{C_N^c} \leq \exp \left( \rho N^{1/3} (\log N)^{2/3} \right) = 0,
\]

which together with (7.19) establishes (7.15).

**Step 2.** In this step we prove that for $\delta_1 \in (0, 1/3)$ we have

\begin{equation}
\lim_{N \to \infty} \mathbb{P} \left( \frac{1}{2} < S_N \leq \exp(-N^{-1/3} + \delta_1) \right) = 0.
\end{equation}
From Proposition 3.2 (applied to \( \delta = 1/2 \)) there exist constants \( D_1, d_1, D_2, d_2 > 0 \) such that for all \( m, n \in \mathbb{N} \) with \( m \geq n, n/m \in [1/2, 1] \) and \( x \geq 0 \)

\[
(7.21) \quad \mathbb{P} \left( \log Z(1, 1; m, n) + mh_{\theta_c}(n/m) \geq x(m^{1/3} + \alpha(n/m)) \right) \leq D_1 e^{-d_1 m} + D_2 e^{-d_2 x^{3/2}}.
\]

As explained in the line above (3.12), we know that \( h_{\theta_c} \) is strictly decreasing on \([0, 1]\) and \( h_{\theta_c}(1) = 0 \). In addition, as explained in Section 3.1 (see the discussion after Proposition 3.6), we know that \( \sigma_{\theta_c}(y) \) is positive and smooth on \((0, \infty)\). In particular, we can find \( \alpha \in (0, 1) \) such that

\[
\alpha^{-1} \geq \sigma_{\theta_c}(y) \geq \alpha \text{ if } y \in [1/2, 1].
\]

Setting \( x = m^{2/3} h_{\theta_c}(n/m) \sigma_{\theta_c}(n/m) \) in (7.21), we conclude that for all \( m, n \in \mathbb{N} \) with \( m \geq n, n/m \in [1/2, 1] \)

\[
(7.22) \quad \mathbb{P}(Z(1, 1; m, n) \geq 1) \leq D_1 e^{-d_1 m} + D_2 e^{-d_2 (m^{1/3} + \alpha(n/m))^{3/2}}.
\]

Let \( \delta > 0 \) be such that (7.10) holds (recall that we proved (7.10) earlier in this section), and let \( T_N^\delta \) be as in the paragraph above (7.12). In addition, let \( U_N^\delta \) be the set of quadruples \((m_1, n_1, m_2, n_2) \in [1, N] \) such that \((m_1, n_1) \leq (m_2, n_2)\) and \( \frac{n_2 - n_1 + 1}{m_2 - m_1 + 1} \in \left(1/2, \exp \left(-N^{-1/3 + \delta_1}\right)\right) \). Notice that if \((m_1, n_1, m_2, n_2) \in U_N^\delta \) we have that \( m = m_2 - m_1 + 1, n = n_2 - n_1 + 1 \) satisfy the conditions above (7.22). Since \( Z(m_1, n_1; m_2, n_2) \) has the same distribution as \( Z(1, 1; m, n) \) we conclude that (7.22) holds with \( Z(1, 1; m, n) \) replaced with \( Z(m_1, n_1; m_2, n_2) \), and \( m = m_2 - m_1 + 1, n = n_2 - n_1 + 1 \). Furthermore as \( h_{\theta_c} \) is strictly decreasing on \([0, 1]\) we have

\[
h_{\theta_c}(n/m) \geq h_{\theta_c} \left( \exp \left(-N^{-1/3 + \delta_1}\right) \right) = \frac{h_{\theta_c}^{(1)}}{2} \cdot N^{-2/3 + 2\delta_1} + O(N^{-1 + 3\delta_1}),
\]

where in the last equality we used (3.11). This gives for \((m_1, n_2, m_2, n_2) \in U_N^\delta \)

\[
(7.23) \quad \mathbb{P}(Z(m_1, n_1; m_2, n_2) \geq 1) \leq D_1 e^{-d_1 (m_2 - m_1 + 1)} + D_2 e^{-d_2 \alpha^{3/2} (m_2 - m_1 + 1) \left(h_{\theta_c}^{(1)}/2\right)^{3/2} \cdot N^{-1 + 3\delta_1} + O((m_2 - m_1 + 1)N^{-5/3 + 5\delta_1})}.
\]

From (7.23) we can conclude the existence of constants \( D, d > 0 \) such that if \((m_1, n_1, m_2, n_2) \in U_N^\delta \setminus T_N^\delta \) (which implies \( N \geq m_2 - m_1 + 1 \geq \delta N/2 \)) we have

\[
(7.24) \quad \mathbb{P}(Z(m_1, n_1; m_2, n_2) \geq 1) \leq D e^{-d N^{3\delta_1}}.
\]

The latter and a union bound gives

\[
(7.25) \quad \mathbb{P}(C_N^\delta) \leq DN^4 \exp(-d N^{3\delta_1}), \text{ and consequently } \lim_{N \to \infty} \mathbb{P}(C_N^\delta) = 0,
\]

where \( C_N^\delta = \bigcup_{(m_1, n_1, m_2, n_2) \in U_N^\delta \setminus T_N^\delta} \{ Z(m_1, n_1; m_2, n_2) \geq 1 \} \).

From here we can repeat verbatim the argument from Step 1 above, starting from (7.18), to conclude (7.20). In that argument we only need to replace the events \{ \( S_N \leq 1/2 \) \} with \{ \( 1/2 < S_N \leq \exp(-N^{-1/3 + \delta_1}) \} = \{(X^*_N, Y^*_N, X^*_N, Y^*_N) \in U_N^\delta \} \), the sets \( U_N \) with \( U_N^\delta \), and the events \( C_N \) with \( C_N^\delta \).

7.3. Proof of Proposition 2.4. In this section we prove Proposition 2.4. Proposition 2.4 contains two types of statements: about the asymptotic distribution of the length \( L_N \) of a path sampled according to the log-gamma polymer measure, and about the asymptotic distribution of the starting point \((X^*_N, Y^*_N)\) of such a path. We will split the proofs of these two types of statements into Sections 7.3.1 and 7.3.2. Throughout this section we assume that \( \theta > \theta_c \).
7.3.1. Asymptotics of $L_N$. In this section we prove \((7.25)\) for all $\theta > \theta_c$ and that $L_N$ is tight for $\theta > 3$. For clarity we split the proof into three steps.

**Step 1.** In this step we summarize the technical inputs we will need in the other steps and specify the value of $c_2$ in the statement of the proposition.

From our discussion at the end of Section 3.1 we know that $g_\theta^{-1}$ is an increasing bijection from $(0,\infty)$ to $(0,\theta)$. Since $\theta > \theta_c$, we can find a (unique) $\phi_c$ (depending on $\theta$) such that $g_\theta(\phi_c) = \theta_c/2$. Moreover, since $g_\theta^{-1}(1) = \theta/2 > \theta_c/2$, we can conclude that $\phi_c \in (0,1)$. If $h_\theta$ is as in \((7.22)\), we know from Section 3.1 that $h_\theta'(x) = \Psi(g_\theta^{-1}(x))$ and is a strictly increasing function on $(0,\infty)$. We also know that $h_\theta'(\phi_c) = \Psi(\theta_c/2) = 0$ and so $h_\theta(x)$ is decreasing on $[0,\phi_c]$ and increasing on $[\phi_c,1]$ so that it attains its minimum on $[0,1]$ at $\phi_c$. We also know that
\[
\kappa := \min_{[0,1]} h_\theta(x) = h_\theta(\phi_c) = (\phi_c) \cdot \Psi(\theta_c/2) = \Psi(\theta - \theta_c/2) > 0,
\]
where we used \((3.2)\) and the fact that $\Psi(\theta_c/2) = 0$, while $\Psi(x) > 0$ for $x > \theta_c/2$.

Let $B, b > 0$ be as in Lemma 3.10 for $\epsilon = \min(\kappa/2,1)$. With this choice of parameters, we define
\[
c_2 := \max(8/\kappa, 10/b).
\]
This is the choice of $c_2$ with which we will prove \((2.5)\) in the next step.

The last thing we do in this step is record the following statement
\[
\lim_{N \to \infty} \mathbb{P}(A_N) = 0, \quad \text{where} \quad A_N = \left\{ \frac{1}{N^2} \sum_{i,j=1}^{N} w_{i,j} \leq \frac{1}{2(\theta - 1)} \right\};
\]
and $w_{i,j}$ are as in Definition 1.1. Equation \((7.28)\) follows from the Law of large numbers and the fact that $\mathbb{E}[w_{1,1}] = (\theta - 1)^{-1}$.

**Step 2.** In this step we prove \((7.25)\). Recall from Step 1, that $B, b > 0$ are as in Lemma 3.10 for $\epsilon = \min(\kappa/2,1)$. Suppose that $m \geq n \geq 1$ and $m \geq (c_2/2) \cdot \log N$. From Lemma 3.10 we know
\[
B \exp(-bn) \geq \mathbb{P}(\theta(1; m, n) \geq m \cdot \epsilon - m \cdot h_\theta(n/m)).
\]
Since $h_\theta(n/m) \geq \kappa$ as in \((7.26)\) and $\epsilon \leq \kappa/2$, we conclude that
\[
B \exp(-bn) \geq \mathbb{P}((\log Z(1,1; m, n) \geq -(\kappa/2)m).
\]
Since $m \geq (c_2/2) \cdot \log N$ and $c_2/2 \geq 4/\kappa$ and $c_2/2 \geq 5/b$ by \((7.27)\) we conclude that
\[
BN^{-5} \geq \mathbb{P}(Z(1,1; m, n) \geq N^{-4}).
\]
Let $T_N$ be the set of quadruples $m_1, m_1, m_2, n_2 \in [1, N]$ such that $(m_1, n_1) \leq (m_2, n_2)$ and $m_2 - m_1 + n_2 - n_1 \geq c_2 \log N$. Using that $Z(m_1, n_1; m_2, n_2)$ has the same distribution as $Z(m_1 + a, n_1 + b; m_2 + a, n_2 + b)$ for any $a, b \in \mathbb{Z}$ and also that $Z(1,1; m, n)$ has the same distribution as $Z(1,1; m, n)$, we conclude from \((7.29)\) that for $(m_1, n_1, m_2, n_2) \in T_N$ we have
\[
BN^{-5} \geq \mathbb{P}(Z(m_1, n_1; m_2, n_2) \geq N^{-4}).
\]
In particular, taking a union bound and using that $|T_N| \leq N^4$ we conclude
\[
\lim_{N \to \infty} \mathbb{P}(B_N) = 0, \quad \text{where} \quad B_N = \cup_{(m_1, m_1, m_2, n_2) \in T_N} \{ Z(m_1, n_1; m_2, n_2) \geq N^{-4} \}.
\]

Using the equality of events $\{ L_N \geq c_2 \log N \} = \{ (X_N^1, Y_N^1, X_N^2, Y_N^2) \in T_N \}$ and \((7.30)\) we see that
\[
\lim_{N \to \infty} \mathbb{P}(L_N \geq c_2 \log N) = \lim_{N \to \infty} \mathbb{P}(\{ L_N \geq c_2 \log N \} \cap A_N^c \cap B_N^c) = \lim_{N \to \infty} \mathbb{E} \left[ \sum_{(m_1, n_1, m_2, n_2) \in T_N} \frac{Z(m_1, n_1; m_2, n_2)}{Z_N} \cdot 1_{A_N^c} \cdot 1_{B_N^c} \right].
\]
In the first equality we used \( \lim_{N \to \infty} \mathbb{P}(A_N) \) \( + \mathbb{P}(B_N) = 0 \), which holds from (7.28) and (7.30).

On the event \( A_N^c \) we have
\[
Z_N^w \geq \sum_{i,j=1}^N w_{i,j} \geq \frac{N^2}{2(\theta - 1)},
\]
and on the event \( B_N^c \) we have for \( (m_1, n_1, m_2, n_2) \in T_N \)
\[
Z(m_1, n_1; m_2; n_2) \leq N^{-4}.
\]
Combining the latter statements, we see
\[
\limsup_{N \to \infty} \mathbb{E} \left[ \frac{\sum_{(m_1, n_1, m_2, n_2) \in T_N} Z(m_1, n_1; m_2; n_2)}{Z_N^w} \cdot 1_{A_N^c} \cdot 1_{B_N^c} \right] \leq \limsup_{N \to \infty} \frac{2(\theta - 1)}{N^2} = 0,
\]
which together with (7.31) establishes (2.5).

**Step 3.** In this last step we show that \( L_N \) is tight when \( \theta > 3 \).

Fix \( (m_1, n_1) \leq (m_2, n_2) \) with \( m_1, n_1, m_2, n_2 \in \mathbb{N} \). Using that \( Z(m_1, n_1; m_2, n_2) \) has the same distribution as \( Z(m_1 + a, n_1 + b; m_2 + a, n_2 + b) \) for any \( a, b \in \mathbb{Z} \) and (3.33) we conclude
\[
\mathbb{E} [Z(m_1, n_1; m_2, n_2)] \leq \left( \frac{2}{\theta - 1} \right)^{m_2-m_1+n_2-n_1+1}.
\]
The last equation implies that
\[
\sum_{(m_1, n_1) \leq (m_2, n_2) \atop m_1, n_1, m_2, n_2 \in [1, N]} \mathbb{E} [(m_2 - m_1 + n_2 - n_1)Z(m_1, n_1; m_2, n_2)] \leq C_\theta N^2, \tag{7.32}
\]
where
\[
C_\theta = \sum_{a+b=0}^\infty (a+b) \cdot \left( \frac{2}{\theta - 1} \right)^{a+b+1} \in (0, \infty).
\]
We mention that \( C_\theta \in (0, \infty) \), since \( \theta > 3 \) and hence \( 2(\theta - 1)^{-1} \in (0, 1) \).

Let \( \epsilon > 0 \) be given. We can find \( M \in \mathbb{N} \) such that
\[
\frac{2C_\theta(\theta - 1)}{M} < \epsilon, \tag{7.33}
\]
Let \( T_N^M \) be the set of quadruples \( m_1, n_1, m_2, n_2 \in [1, N] \) such that \( (m_1, n_1) \leq (m_2, n_2) \) and \( m_2 - m_1 + n_2 - n_1 \geq M \). Using the equality of events \( \{L_N \geq M\} = \{X_N^*, Y_N^*, X_N^c, Y_N^c\} \in T_N^M \) and (7.33) we see that
\[
\limsup_{N \to \infty} \mathbb{P}(L_N \geq M) = \limsup_{N \to \infty} \mathbb{P}(\{L_N \geq M\} \cap A_N^c) = \limsup_{N \to \infty} \mathbb{E} \left[ \frac{\sum_{(m_1, n_1, m_2, n_2) \in T_N^M} Z(m_1, n_1; m_2; n_2)}{Z_N^w} \cdot 1_{A_N^c} \right], \tag{7.34}
\]
where \( A_N \) is as in (7.28) and we used from the same equation that \( \lim_{N \to \infty} \mathbb{P}(A_N) \).
On the event $A_N^+$, we have

$$Z_N^w \geq \sum_{i,j=1}^{N} w_{i,j} \geq \frac{N^2}{2(\theta - 1)},$$

and so from (7.34) we get

$$\limsup_{N \to \infty} \mathbb{P}(L_N \geq M) \leq \limsup_{N \to \infty} \frac{2(\theta - 1)}{N^2} \cdot \mathbb{E}\left[ \sum_{(m_1, n_1, m_2, n_2) \in T_N^M} Z(m_1, n_1; m_2; n_2) \right].$$

For $(m_1, n_1, m_2, n_2) \in T_N^M$ we have $m_2 - m_1 + n_2 - n_1 \geq M$ and so

$$\limsup_{N \to \infty} \mathbb{P}(L_N \geq M) \leq \limsup_{N \to \infty} \frac{2(\theta - 1)}{N^2 M} \cdot \sum_{(m_1, n_1, m_2, n_2) \in T_N^M} \mathbb{E}[(m_2 - m_1 + n_2 - n_1)Z(m_1, n_1; m_2; n_2)] \leq \limsup_{N \to \infty} \frac{2(\theta - 1)}{N^2 M} \cdot C_\theta N^2 = \frac{2C_\theta(\theta - 1)}{M} < \epsilon.$$ 

We mention that the second inequality in (7.35) used (7.32) and in the last equality we used (7.33). As $\epsilon > 0$ was arbitrary we see that (7.35) implies that $L_N$ is tight.

### 7.3.2. Asymptotics of $(X_N^a, Y_N^b)$

In this section we prove that for $\theta > \theta_c$ the sequence of random vectors $(N^{-1}X_N^a, N^{-1}Y_N^b)$ converges weakly to the uniform measure on $(0, 1)^2$. For clarity we split the proof into two steps.

**Step 1.** We claim that if $x, y \in [0, 1]$ and $\Delta_x, \Delta_y \in (0, 1/4]$ are such that $[x, x + \Delta_x] \times [y, y + \Delta_y] \subseteq [0, 1]^2$, then

$$\limsup_{N \to \infty} \mathbb{P}\left( x \leq N^{-1}X_N^a \leq x + \Delta_x, y \leq N^{-1}Y_N^b \leq y + \Delta_y \right) \leq \frac{1}{(|\Delta_x^{-1}| - 3) \cdot (|\Delta_y^{-1}| - 3)}.$$

We prove (7.36) in the second step. Here we assume its validity and conclude the proof of the proposition.

Suppose that $[a, a + \Delta_a] \times [b, b + \Delta_b] \subseteq [0, 1]$ and $\Delta_a, \Delta_b > 0$. We fix $A, B \in \mathbb{N}$ such that $A, B \geq 2$ and divide $[a, a + \Delta_a] \times [b, b + \Delta_b]$ into $2^{A+B}$ equal rectangles of side lengths $2^{-A}\Delta_a$ and $2^{-B}\Delta_b$. Explicitly, these rectangles are given by

$$R_{i,j} = [x_i, x_i + 2^{-A}\Delta_a] \times [y_j, y_j + 2^{-B}\Delta_b],$$

where $x_i = a + (i - 1) \cdot 2^{-A}\Delta_a,$ $y_j = b + (j - 1) \cdot 2^{-B}\Delta_b$ for $i = 1, \ldots, 2^A$ and $j = 1, \ldots, 2^B$. We then have

$$\limsup_{N \to \infty} \mathbb{P}\left((N^{-1}X_N^a, N^{-1}Y_N^b) \in [a, a + \Delta_a] \times [b, b + \Delta_b]\right) \leq$$

$$\sum_{i=1}^{2^A} \sum_{j=1}^{2^B} \limsup_{N \to \infty} \mathbb{P}\left((N^{-1}X_N^a, N^{-1}Y_N^b) \in R_{i,j}\right) \leq \frac{2^{A+B}}{(2^A\Delta_a^{-1} - 3) \cdot (2^B\Delta_b^{-1} - 3)},$$

where in the first inequality we used subadditivity and in the second we used (7.36) for the rectangles $R_{i,j}$ (which by construction have side-lengths in $(0, 1/4)$).

As (7.37) holds for all $A, B \geq 2$ we can let $A, B \to \infty$ to conclude that

$$\limsup_{N \to \infty} \mathbb{P}\left(((N^{-1}X_N^a, N^{-1}Y_N^b) \in [a, a + \Delta_a] \times [b, b + \Delta_b]\right) \leq \Delta_a \cdot \Delta_b.$$
To prove that \((N^{-1}X_N^s, N^{-1}Y_N^s)\) converges weakly to the uniform measure on \((0,1)^2\) it suffices to show that for all \((s,t) \in (0,1)^2\) we have

\[
\lim_{N \to \infty} \mathbb{P} \left( N^{-1}X_N^s \leq s, N^{-1}Y_N^s \leq t \right) = s \cdot t.
\]

From (7.38) applied to \(a = 0, \Delta_a = s, \ b = 0, \Delta_b = t\) we see that

\[
\limsup_{N \to \infty} \mathbb{P} \left( N^{-1}X_N^s \leq s, N^{-1}Y_N^s \leq t \right) \leq s \cdot t.
\]

On the other hand,

\[
1 - \mathbb{P} \left( N^{-1}X_N^s \leq s, N^{-1}Y_N^s \leq t \right) \leq \mathbb{P} \left( (N^{-1}X_N^s, N^{-1}Y_N^s) \in [s,1] \times [0,1] \right) + \mathbb{P} \left( (N^{-1}X_N^s, N^{-1}Y_N^s) \in [0,s] \times [t,1] \right).
\]

Taking \(\limsup_{N \to \infty}\) on both sides and using (7.40) we conclude

\[
1 - \liminf_{N \to \infty} \mathbb{P} \left( N^{-1}X_N^s \leq s, N^{-1}Y_N^s \leq t \right) \leq (1-s) + (1-t)s \iff \liminf_{N \to \infty} \mathbb{P} \left( N^{-1}X_N^s \leq s, N^{-1}X_N^s \leq t \right) \geq s \cdot t.
\]

Equations (7.40) and (7.41) together imply (7.39), which concludes the proof of the proposition.

**Step 2.** In this step we prove (7.36). We begin by introducing some useful notation. We define

\[
A_N^x = [Nx], \quad B_N^x = [N(x + \Delta_x)], \quad A_N^y = [Ny], \quad B_N^y = [N(y + \Delta_y)],
\]

\[
\bar{B}_N^x = \text{min}(B_N^x + [N^{1/4}], N), \quad \bar{B}_N^y = \text{min}(B_N^y + [N^{1/4}], N), \quad D_N^x = \bar{B}_N^x - A_N^x + 1, \quad D_N^y = \bar{B}_N^y - A_N^y + 1.
\]

We further define \(S_N^{x/y}\) through

\[
S_N^{x/y} = \{ z \in \mathbb{Z} : N \geq A_N^{x/y} + z \cdot D_N^{x/y} \geq 1 \text{ and } N \geq \bar{B}_N^{x/y} + z \cdot D_N^{x/y} \geq 1 \}.
\]

Observe that both \(S_N^x\) and \(S_N^y\) contain 0 and constitute a finite sequence of consecutive integers. Denoting the smallest and largest elements in \(S_N^{x/y}\) by \(P_N^{x/y}\) and \(Q_N^{x/y}\), respectively, we have

\[
S_N^{x/y} = [P_N^{x/y}, Q_N^{x/y}].
\]

For \((p, q) \in S_N^x \times S_N^y\) we let

\[
R_N^{p,q} = \{(m,n) \in \mathbb{N} : \bar{B}_N^x + pD_N^x \geq m \geq A_N^x + pD_N^x \text{ and } \bar{B}_N^y + qD_N^y \geq n \geq A_N^y + qD_N^y\}.
\]

We summarize the following observations, which immediately follow from the above definition

1. We have the equality of events:

\[
\{ (N^{-1}X_N^s, N^{-1}Y_N^s) \in [x,x+\Delta_x] \times [y,y+\Delta_y] \} = \{ (X_N^s, Y_N^s) \in [A_N^x, B_N^x] \times [A_N^y, B_N^y] \}.
\]

2. \(R_N^{p,q}\) for \((p, q) \in S_N^x \times S_N^y\) are pairwise disjoint.

3. We have the inclusion of events:

\[
\{ (X_N^s, Y_N^s) \in [A_N^x, B_N^x] \times [A_N^y, B_N^y], (X_N^e, Y_N^e) \not\in R_N^{0,0} \} \subseteq \{ L_N \geq [N^{1/4}] \}.
\]

4. We have the inequalities

\[
(Q_N^{x/y} + 1) \cdot D_N^{x/y} + \bar{B}_N^{x/y} \geq N + 1 \text{ and } (P_N^{x/y} - 1) \cdot D_N^{x/y} + A_N^{x/y} \leq 0.
\]
We also note that for all large $N$ we have
\begin{equation}
|S_N^x| \geq |\Delta_x^{-1}| - 3 \quad \text{and} \quad |S_N^y| \geq |\Delta_y^{-1}| - 3.
\end{equation}

Let us prove the latter briefly, and as the two inequalities are analogous focus only on the first. By definition, we have $|S_N^x| = Q_N^x - P_N^x + 1$ and so the fourth observation above gives
\begin{equation}
|S_N^x| \geq \left(\frac{N + 1 - \bar{B}_N^x}{D_N^x} - 1\right) - \left(1 - \frac{A_N^x}{D_N^x}\right) + 1 = \frac{N + 1 - \bar{B}_N^x + A_N^x}{D_N^x} - 1.
\end{equation}

By definition, we have that
\begin{equation}
\lim_{N \to \infty} \frac{N + 1 - \bar{B}_N^x + A_N^x}{D_N^x} = 1 - \frac{(x + \Delta_x) + x}{\Delta_x} = 1 - 1,
\end{equation}

which together with the last inequality implies the first inequality in (7.42) for all large $N$.

We now turn to proving (7.36). Using the first and third observations above we have
\begin{align}
&\limsup_{N \to \infty} \mathbb{P}\left(x \leq N^{-1}X_N^x \leq x + \Delta_x, y \leq N^{-1}Y_N^y \leq y + \Delta_y\right) = \\
&\limsup_{N \to \infty} \mathbb{P}\left(\mathbb{I}(X_N^x, Y_N^y) \in [A_N^x, B_N^x] \times [A_N^y, B_N^y]\right) = \\
&\limsup_{N \to \infty} \mathbb{P}\left(\mathbb{I}(X_N^x, Y_N^y) \in [A_N^x, B_N^x] \times [A_N^y, B_N^y], (X_N^x, Y_N^y) \in R_N^{0,0}\right) = \\
&\limsup_{N \to \infty} \mathbb{P}\left(\mathbb{I}(X_N^x, Y_N^y) \in R_N^{0,0}, (X_N^x, Y_N^y) \in R_N^{0,0}\right).
\end{align}

We mention that in going from the second to the third line we used the third observation above and the fact that
\begin{equation}
\lim_{N \to \infty} \mathbb{P}(L_N \geq \lfloor N^{1/4} \rfloor) = 0
\end{equation}
as follows from (2.3), which we proved in the previous section.

We next have from (7.3) that
\begin{equation}
\mathbb{P}\left((X_N^x, Y_N^y) \in R_N^{0,0}, (X_N^e, Y_N^e) \in R_N^{0,0}\right) = \mathbb{E}\left[\frac{1}{Z_N^w} \sum_{(m_1, n_1) \leq (m_2, n_2) \in R_N^{0,0}} Z(m_1, n_1; m_2, n_2)\right].
\end{equation}

Let us set for $p \in S_N^x$ and $q \in S_N^y$
\begin{equation}
Z_N^{p,q} = \sum_{(m_1, n_1) \leq (m_2, n_2) \in R_N^{0,0}} Z(m_1, n_1; m_2, n_2).
\end{equation}

From the second observation above we know that $R_N^{p,q}$ are pairwise disjoint and so $Z_N^{p,q}$ are independent, but they are also identically distributed as $R_N^{0,p}$ are translates of the same rectangle $R_N^{0,0}$. In addition, by the disjointness of $R_N^{p,q}$ we have that
\begin{equation}
Z_N^w \geq \sum_{p \in S_N^x, q \in S_N^y} Z_N^{p,q}.
\end{equation}

Combining the last two observations and (7.44) gives
\begin{equation}
\mathbb{P}\left((X_N^x, Y_N^y) \in R_N^{0,0}, (X_N^e, Y_N^e) \in R_N^{0,0}\right) = \mathbb{E}\left[\frac{Z_N^{0,0}}{Z_N^w} \sum_{p \in S_N^x, q \in S_N^y} Z_N^{p,q}\right] = \frac{1}{|S_N^x| - |S_N^y|}.
\end{equation}

Combining (7.42), (7.43) and (7.45) proves (7.36).
APPENDIX A. UPPER TAIL ESTIMATES

Here we prove Proposition 3.3. We continue with the same notation as in Sections 1 and 3.1.

A.1. Estimates on Fredholm determinants. In this section we recall a result from [BCD21a], see Proposition A.3, which expresses the Laplace transform of \(Z(1, 1; M, N)\) from Section 1 in terms of a Fredholm determinant. In Proposition A.6 we show that this Laplace transform is close to the GUE Tracy-Widom distribution and we use this statement to establish Proposition 3.3. We record here that for a measurable function \(K(x, y)\) on \(X \times X\), where \((X, \mu)\) is a measure space, we define

\[
\det(I + K)_{L^2(X)} := 1 + \sum_{n=1}^{\infty} \frac{1}{n!} \int_X \cdots \int_X \det[K(x_i, x_j)]_{i,j=1}^n \prod_{i=1}^n d\mu(x_i),
\]

provided that the above integrals and corresponding sum are absolutely convergent. In this paper we will exclusively work in the case when \(X\) is a piecewise smooth contour in \(\mathbb{C}\) and \(\mu(dz) = \frac{dz}{2\pi i}\) with \(dz\) denoting the usual complex integration. We refer to [BCD21a, Section 2.1] for some basic background on Fredholm determinants.

We begin by introducing some useful notation, see also Figure 8.

\[
\text{Definition A.1.} \quad \text{For } a \in \mathbb{C} \text{ and } \phi \in (0, \pi) \text{ we define the contour } C_{a,\phi} \text{ to be the union of } \{a + ye^{-i\phi}\}_{y \in \mathbb{R}^+} \text{ and } \{a + ye^{i\phi}\}_{y \in \mathbb{R}^+} \text{ oriented to have increasing imaginary part.}
\]

\[
\text{Definition A.2.} \quad \text{Suppose that } a \in \mathbb{C}, \phi \in (0, \pi), d > 0 \text{ and } v \in \mathbb{C}. \text{ From this data we construct a contour that consists of two parts. The first, called } D^1_v, \text{ is } C_{a,\phi} \setminus \{z : \text{Im}(z) \in [\text{Im}(v) - d, \text{Im}(v) + d]\}. \text{ Let } z_-, z_+ \text{ be the points on } C_{a,\phi} \text{ that have imaginary parts } \text{Im}(v) - d \text{ and } \text{Im}(v) + d \text{ respectively. The second part of the contour, called } D^2_v, \text{ consists of straight oriented segments that connect } z_- \text{ to } v + 2d - id \text{ to } v + 2d + id \text{ to } z_+. \text{ We will denote the resulting contour by } D_v(a, \phi, d) \text{ or just } D_v \text{ when the other parameters are clear from the context.}
\]

The first major result we require from [BCD21a] is the following Laplace transform formula for the partition function \(Z(1, 1; M, N)\).

\[
\text{Proposition A.3.} \quad \text{[BCD21a, Theorem 2.12]} \quad \text{Fix integers } N \geq 9, M \geq 1 \text{ and } \theta > 0. \text{ Let } \theta > b > a > 0, d \in (0, \min(1/4, (b - a)/4)) \text{ and } u \in \mathbb{C} \text{ with Re}(u) > 0. \text{ Then}
\]

\[
\mathbb{E} \left[e^{-uZ(1, 1; M, N)}\right] = \det(I + K_u)_{L^2(C_{a,3\pi/4})},
\]

\[
\text{(A.2)}
\]
where the operator \( K_u \) is defined in terms of its integral kernel
\[
K_u(v, v') = \frac{1}{2\pi i} \int_{D_v} \frac{\pi}{\sin(\pi(v - w))} \left( \frac{\Gamma(v)}{\Gamma(w)} \right)^N \left( \frac{\Gamma(\theta - w)}{\Gamma(\theta - v)} \right)^M u^{w-v} dw / (w - v'),
\]
where \( C_{a,3\pi/4} \) and \( D_v = D_v(b, \pi/4, d) \) are as in Definitions A.1 and A.2.

Remark A.4. Part of the statement of Proposition A.3 is that \( \det (I + K_u)_{L^2(C_{a,3\pi/4})} \) is well-defined in the sense that the integral in (A.3) is well-defined for each \( v, v' \in C_{a,3\pi/4} \), the resulting function is measurable and the series in (A.1) is absolutely convergent for \( X = C_{a,3\pi/4} \) and \( \mu(dz) = dz / (2\pi i) \) with \( dz \) denoting the usual complex integration.

We next summarize a certain scaling of \( M, N \) and introduce several quantities of interest, which will appear in our analysis.

Definition A.5. Recall that \( \Psi(x) \) denotes the digamma function, see (1.3). Suppose that \( M, N \geq 1 \) and \( \theta > 0 \) are given. We let \( z_c(M, N) \) denote the maximizer of
\[
W_{M,N}(x) := N\Psi(x) + M\Psi(\theta - x)
\]
on the interval \((0, \theta)\). Notice that the above expression converges to \(-\infty\) as \( x \to 0^+ \) or \( x \to \theta^- \) and also the function \( W_{M,N}(x) \) is strictly concave, hence the maximum exists and is unique. We let \( W_{M,N} \) denote \( W_{M,N}(z_c) \). It is also worth noting that \( z_c = g^{-1}_\delta(N/M) \).

For any \( \alpha > 0 \) we define
\[
\tilde{\sigma}_\alpha := \left( \sum_{n=0}^{\infty} \frac{\alpha}{(n + z_c)^3} + \sum_{n=0}^{\infty} \frac{1}{(n + \theta - z_c)^3} \right)^{1/3}.
\]

The way we scale the parameters is as follows. We will let \( M, N \to \infty \) while \( M \geq N \) and \( \alpha = \alpha(M, N) = N/M \geq \delta \), where \( \delta > 0 \) is fixed. For a given \( x \in \mathbb{R} \) we set
\[
u = u(x, M, N) := e^{W_{M,N}-M^{1/3}\tilde{\sigma}_\alpha x}.
\]
Notice that if \( \alpha \in [\delta, 1] \) and \( \theta > 0 \) is fixed then \( \sigma_\alpha \) is positive and bounded away from 0 and \( \infty \), see e.g. [BCD21a, (5.4)]. We finally let \( K_u \) be as in Proposition A.3 for the parameter \( u \) as above, \( a = z_c \) and \( b, d \) arbitrarily chosen so as to satisfy the conditions of that proposition.

The main technical result of this section is contained in the following strong comparison of the Fredholm determinant in Proposition A.3 and \( F_{\text{GUE}} \).

Proposition A.6. Assume the same notation and scaling as in Definition A.5. Then for any \( \epsilon_0 \in (0, 1/3) \) we have
\[
\lim_{M \to \infty} \sup_{x \in [0, \infty)} M^{\epsilon_0} \left| \det (I + K_u)_{L^2(C_{z_c,3\pi/4})} - F_{\text{GUE}}(x) \right| = 0.
\]

In the remainder of this section we use Proposition A.6 to prove Proposition 3.3.

Proof of Proposition A.6. Assume the same notation as in Proposition A.3 where \( a, b, d, u \) are as in Definition A.5. It follows from Proposition A.6 that there exists \( M_0 \) sufficiently large so that for \( M \geq M_0 \) and \( M \geq N \geq \delta M \) and \( y \in [0, \infty) \) we have
\[
\left| \det (I + K_u)_{L^2(C_{z_c,3\pi/4})} - F_{\text{GUE}}(y) \right| \leq M^{-\epsilon_0},
\]
where \( u = e^{W_{M,N}-M^{1/3}\tilde{\sigma}_\alpha y} \). In particular, if for any \( x \in [1, \infty) \) we set \( x^\pm = x \pm M^{-1/3}(\log M)^2 \) and \( u^\pm = e^{W_{M,N}-M^{1/3}\tilde{\sigma}_\alpha x^\pm} \) we see that by possibly enlarging \( M_0 \) (so that \( x^- \geq 0 \)) we obtain
\[
\left| \det (I + K_{u^\pm})_{L^2(C_{z_c,3\pi/4})} - F_{\text{GUE}}(x^\pm) \right| \leq M^{-\epsilon_0}.
\]
If we set $f_M(z) = \exp(-\delta z^{1/3})$, then by Proposition A.3 we know that for all large enough $M$

$$E[f_M(F(M,N) - x^\pm)] = E[e^{-\delta z^{1/3}(1,1;M,N)}] = \det(I + K_{u^\pm})L^2(C_{x,3\pi/4}),$$

We note that by (A.5) and (A.6) we have for all large $M$

$$P(F(M,N) \leq x) = E[1\{F(M,N) \leq x\}] \geq E\left[e^{-\delta z^{1/3}(1,1;M,N)-x^-} \cdot 1\{F(M,N) \leq x\}\right] \geq \det(I + K_{u^-})L^2(C_{x,3\pi/4}) - M^{-\epsilon_0} \geq \log \Gamma(\theta) \cdot \log \Gamma(\theta - z) - M^{1/3}(\log M)^2.$$

By the same equations we also have

$$\log \Gamma(\theta) \cdot \log \Gamma(\theta - z) - M^{1/3}(\log M)^2 \geq 1 - \log \Gamma(\theta) \cdot \log \Gamma(\theta - z) - M^{-\epsilon_0} \geq \log \Gamma(\theta) \cdot \log \Gamma(\theta - z) - M^{-\epsilon_0}.$$

Combining the last two inequalities we see that

$$\log \Gamma(\theta) \cdot \log \Gamma(\theta - z) - M^{1/3}(\log M)^2 \leq \log \Gamma(\theta) \cdot \log \Gamma(\theta - z) - M^{-\epsilon_0},$$

which implies (3.1) once we utilize the fact that $\log \Gamma(\theta) \cdot \log \Gamma(\theta - z) - M^{1/3}(\log M)^2$. The latter equality follows from the boundedness of $\log \Gamma(\theta) \cdot \log \Gamma(\theta - z)$ on $[-1, \infty)$, which for example can be deduced by the log-concavity of $\log \Gamma(\theta)$ on $[0, \infty)$, see [BLS17, Theorem 5.1].

A.2. Preliminary results. In this section we summarize some results from [BCD21a] and prove Lemma A.10 which will be used in the proof of Proposition A.6 in the next section. We begin with some useful notation.

**Definition A.7.** Define

$$G_{M,N}(z) := N \log \Gamma(z) - M \log \Gamma(\theta - z) - W_{M,N}z - C_{M,N},$$

where $C_{M,N} = N \log \Gamma(z_c) - M \log \Gamma(\theta - z_c) - W_{M,N}z_c$, and $z_c, W_{M,N}$ are as in Definition A.5. We also define

$$G_\alpha(z) = M^{-1}G_{M,N}(z) = \alpha \log \Gamma(\theta) - M^{-1}W_{M,N}z - M^{-1}C_{M,N},$$

where we recall that $\alpha = N/M \in [\delta, 1]$ as in Definition A.5. We define the contour $D$ to be the contour that consists of the vertical segment connecting $1 - i$ and $1 + i$ and for $\delta > 0$ we agree that $C_{0,3\pi/4}$ from Definition A.1, see Figure 9. We also define the contours $D_M$ through $D_M = z_c + \sigma_\alpha^{-1}M^{-1/3} \cdot D$. The contours $D$ and $D_M$ are oriented to have increasing imaginary part.

We next state several results from [BCD21a].

**Lemma A.8.** [BCD21a, Lemma 3.5] Fix $\theta > 0$, $\delta \in (0,1)$ and assume that $M \geq N \geq 1$, $N/M \in [\delta, 1]$. There exist constants $C > 0$ and $r > 0$ depending on $\theta$ and $\delta$ such that $G_\alpha$ is analytic in the disc $|z - z_c| < r$ and the following hold for $|z - z_c| \leq r$:

$$|G_\alpha(z) + (z - z_c)^3\delta_\alpha^3/3| \leq C|z - z_c|^4,$$

$$\Re[G_\alpha(z)] \geq (\sqrt{2}/2)|z - z_c|^3\delta_\alpha^3/6$$

when $z \in C_{z_c, \phi}$ with $\phi = \pi/4$;

$$\Re[G_\alpha(z)] \leq -(\sqrt{2}/2)|z - z_c|^3\delta_\alpha^3/6$$

when $z \in C_{z_c, \phi}$ with $\phi = 3\pi/4$.

In the above equations $z_c, \delta_\alpha$ are as in Definition A.5, and $C_{z_c, \phi}$ is as in Definition A.7.

**Lemma A.9.** [BCD21a, Lemma 3.10] For any $\delta, \Delta > 0$ there exists $M_0 > 0$ such that if $M \geq M_0$

$$\sum_{n=1}^{\infty} n^{n/2} \Delta^n M^{n/3} \leq e^{\delta M}.$$
Lemma A.10. Suppose that \( g(v,v',w) \) is a bounded complex-valued measurable function on \( C_{0,\phi} \times C_{0,\phi} \times D \), where \( \phi \in [\pi/2, \pi) \), \( C_{0,\phi} \) is as in Definition A.1 and \( D \) is as in Definition A.3. Assume that there are constants \( A,B > 0 \) such that for some \( a,b \geq A \) we have
\[
|g(v,v',w)| \leq B \cdot \exp \left( -a|v|^3 - b|w|^3 \right).
\]
Then the kernel \( K(v,v') = \frac{1}{2\pi i} \int_D g(v,v',w) dw \) is well-defined on \( C_{0,\phi} \times C_{0,\phi} \) and also the Fredholm determinant \( \det (I + K)_{L^2(C_{0,\phi})} \) is well-defined. Furthermore, there is a constant \( N_0 \geq 1 \) (depending on \( A,B \)) such that for all \( N \geq N_0 \) we have
\[
\det (I + K)_{L^2(C_{0,\phi})} - 1 - \sum_{n=1}^{N} \frac{1}{n!} \int_{C_{0,\phi}} \cdots \int_{C_{0,\phi}} \det [K^n(v_i,v_j)]_{i,j=1}^{n} \prod_{i=1}^{n} \frac{dv_i}{2\pi i} \leq e^{- (N/3) \log N},
\]
where \( K^n(v,v') = \frac{1}{2\pi i} \int_D g^n(v,v',w) dw \) with \( g^n(v,v',w) = 1\{\max(|v|,|w|,|v'|) \leq N\} \cdot g(v,v',w) \).

Proof. The fact that all the Fredholm determinants in the statement of the lemma are well-defined in terms of absolutely convergent series as in (A.1) follows from Lemmas 2.3 and 2.4 in [BCD21a], whose proofs can be found in [Dim17 Section 2.5]. In the remainder we prove (A.11).

In view of (A.10) we can find a constant \( D_1 > 0 \) (depending on \( A,B \)) such that for all \( v,v' \in C_{0,\phi} \)
\[
\int_D |g(v,v',w)||dw| \leq D_1 \cdot e^{-A|v|^3}
\]
and so \( |K(v,v')| \leq D_1 e^{-A|v|^3} \), where \( |dw| \) denotes integration with respect to arc length. This and Hadamard’s inequality imply
\[
\frac{\sum_{n=N+1}^{\infty} \frac{1}{n!} \int_{C_{0,\phi}} \cdots \int_{C_{0,\phi}} \det [K^n(v_i,v_j)]_{i,j=1}^{n} \prod_{i=1}^{n} \frac{dv_i}{2\pi i}}{D_1^n} \leq \frac{\sum_{n=N+1}^{\infty} \frac{(D_1 D_2)^n n^{n/2}}{n!}}{D_1^n},
\]
where \( D_2 = \int_{C_{0,\phi}} e^{-A|v|^3} |dv| \). From [Rob55] (1) and (2) we have for each \( n \geq 1 \) that
\[
n! = \sqrt{2\pi n^{n+1/2}} e^{-n} r_n, \text{ where } (12n+1)^{-1} < r_n < (12n)^{-1}.
\]
Equation (A.13) implies that if we set \( a_n = \frac{(D_1 D_2)^n n^{n/2}}{n!} \) we have that \( \frac{a_{n+1}}{a_n} \sim \frac{D_1 D_2 \sqrt{e}}{\sqrt{n}} \) as \( n \to \infty \) and so for all large \( N \) we have
\[
\sum_{n=N+1}^{\infty} \frac{(D_1 D_2)^n n^{n/2}}{n!} \leq \frac{\sum_{n=N+1}^{\infty} \frac{(D_1 D_2)^n N^{N/2}}{N!}}{2} \leq \frac{e^{-(1/3)N \log N}}{\sqrt{N}}.
\]
where in the last inequality we used (A.13) again.

For \( \sigma \in S_n \) we set \( \chi_n^\sigma = \prod_{i=1}^{n} 1\{\max(|v_i|, |w_i|, |v_{\sigma(i)}|) \leq N \} \), \( C_1 = \int_{C_{0,0}} e^{-(A/2)|v|^3} dv \) and \( C_2 = \int_D e^{-(A/2)|w|^3} dw \). By expanding determinants we get

\[
\left| \sum_{n=1}^{N} \frac{1}{n!} \int_{C_{0,0}} \cdots \int_{C_{0,0}} [K(v_i, v_j)]_{i,j=1}^{n} \prod_{i=1}^{n} \frac{dv_i}{2\pi i} - \sum_{n=1}^{N} \frac{1}{n!} \int_{C_{0,0}} \cdots \int_{C_{0,0}} \det \left[ K^N(v_i, v_j) \right]_{i,j=1}^{n} \prod_{i=1}^{n} \frac{dv_i}{2\pi i} \right| \leq \\
\sum_{n=1}^{N} \sum_{\sigma \in S_n} \frac{1}{n!} \int_{C_{0,0}} \cdots \int_{C_{0,0}} \int_D \cdots \int_D (1 - \chi_n^\sigma) \prod_{i=1}^{n} |g(v_i, v_{\sigma(i)}, w_i)| \prod_{i=1}^{n} |dw_i| \prod_{i=1}^{n} |dv_i| \leq \\
e^{-(A/2)N^3} \sum_{n=1}^{N} \int_{C_{0,0}} \cdots \int_{C_{0,0}} \int_D \cdots \int_D B^n \prod_{i=1}^{n} \exp \left( -(A/2)|v_i|^3 - (A/2)|w_i|^3 \right) \prod_{i=1}^{n} |dw_i| \prod_{i=1}^{n} |dv_i| \leq \\
e^{-(A/2)N^3} \sum_{n=1}^{N} B^n C_1^n C_2^n \leq (1/2)e^{-(1/3)N \log N}.
\]

The second inequality above requires some explanation that we give momentarily and the final inequality holds for all large enough \( N \). Combining the last inequality with (A.12) and (A.14) we obtain (A.11).

For the second inequality above we note that in order for \((1 - \chi_n^\sigma)\) to be non-zero (and hence equal to 1) we must have some \( i \) for which either \(|v_i|\) or \(|w_i|\) exceeds \( N \). For that \( i \), we can use (A.10) to show that \(|g(v_i, v_{\sigma(i)}, w_i)| \leq B e^{-a|v_i|^3 - b|w_i|^3} \leq B e^{-(A/2)N^3} e^{-(A/2)|v_i|^3 - (A/2)|w_i|^3}\). This last inequality follows since \(-a|v_i|^3 - b|w_i|^3 \leq -a|v_i|^3 - A|w_i|^3 \leq -A|v_i|^3 - (A/2)|v_i|^3 - (A/2)|w_i|^3\), as follows from the fact that \(a, b \geq A\) and at least one of \(|v_i|\) or \(|w_i|\) exceeds \( N \).

**A.3. Proof of Proposition A.6.** In this section we present the proof of Proposition A.6. For clarity, we split the proof into three steps. In the first step we show that we can truncate and deform the contours in the definition of \( \det(I + K_u)_{L^2(C_{\pi, 3\pi/4})} \) to more suitable ones without affecting the value of this Fredholm determinant significantly. In the second step we claim a certain finite sum estimate, see (A.21), and deduce the statement of the proposition by combining this statement and Lemma A.10. Equation (A.21) is proved in the third and final step.

**Step 1.** Let \( r \) (depending on \( \theta \) and \( \delta \)) be as in Lemma A.8 and fix \( \epsilon > 0 \) such that \( \epsilon < \min(r/2, 1/4) \). Let \( C^e_{\pi, 3\pi/4} \) denote the portion of \( C_{\pi, 3\pi/4} \) inside \( B_\epsilon(z_c) \) – the disc of radius \( \epsilon \), centered at \( z_c \). Repeating verbatim the argument in Step 2 of the proof of [BCD21a Theorem 1.7] (see (4.4) in that paper) we get that there is a constant \( c_1 > 0 \) (depending on \( \theta, \delta, \epsilon \)) such that for all large \( M \)

\[
|\det(I + K_u)_{L^2(C_{\pi, 3\pi/4})} - \det(I + K_u)_{L^2(C^e_{\pi, 3\pi/4})}| \leq e^{-c_1 M}.
\]

By Cauchy’s theorem we may deform the \( D_v(b, d, \pi/4) \) contour in the definition of \( K_u(v, v') \) to \( D_M \) from Definition A.7 without affecting the value of the kernel as long as \( M \) is sufficiently large. Indeed, notice that by our choice of \( \epsilon \leq 1/4 \) we have that \( C^e_{\pi, 3\pi/4} +1 \) lies to the right of \( D_M \) and so we do not cross any poles while deforming \( D_v(b, d, \pi/4) \) to \( D_M \), see Figure 9. The decay estimates necessary to deform the contour near infinity come from [BCD21a Proposition 2.15] applied to \( K = [0, \theta], T = 0 \) and \( \alpha_i = \theta \), \( \alpha_j = 0 \) and \( u \) as in Definition A.5.

We write \( D_M = D^0_M \cup D^1_M \), where \( D^0_M \) is the portion of \( D_M \) inside \( B_\epsilon(z_c) \) and \( D^1_M \) is the portion outside. Define for \( \beta \in \{0, 1\} \) and \( v, v' \in C^e_{\pi, 3\pi/4} \) the kernel

\[
K^\beta(v, v') = \frac{1}{2\pi i} \int_{D^\beta_M} \frac{F(v, w)}{w - v'} dw, \quad \text{where } F(v, w) = \frac{\pi}{\sin(\pi(v - w))} e^{G_M.N(v) - G_M.N(w)} e^{M^{1/3}(w-v)\partial_u x}.
\]
From [BCD21a, equation (4.11)] we have that there is a constant $C_2, c_2 > 0$ (depending on $\theta, \delta, \epsilon$) such that for all large enough $M$

$$\left| \int_{C_{z_c, 3\pi/4}} \cdots \int_{C_{z_c, 3\pi/4}} \det \left[ K_u^\beta(v_i, v_j) \right]_{i,j=1}^n dv_i \right| \leq C_2^n n^{n/2} M^{n/3} e^{-2c_2 M \sum_{i=1}^n \beta_i}.$$  

Using the last inequality, the fact that $K_u(v, v') = K_u^0(v, v') + K_u^1(v, v')$, the definition of a Fredholm determinant [A.1], and the linearity of the determinant function we obtain

$$\left| \det(I + K_u)_{L^2(C_{z_c, 3\pi/4})} - \det(I + K_u^0)_{L^2(C_{z_c, 3\pi/4})} \right| \leq e^{-2c_2 M \sum_{n=1}^{2n} \frac{C_2^n n^{n/2} M^{n/3}}{n!}} \leq e^{-c_2 M},$$

where the last inequality used Lemma [A.9] with $\delta = c_2$ and $\Delta = 2C_2$ and holds for large $M$.

Equations (A.15) and (A.16) show that $\det(I + K_u)_{L^2(C_{z_c, 3\pi/4})}$ is very close to $\det(I + K_u^0)_{L^2(C_{z_c, 3\pi/4})}$ and in the remainder of this step we rewrite $\det(I + K_u^0)_{L^2(C_{z_c, 3\pi/4})}$ in a more suitable form for the application of Lemma [A.10] in the next step.

Applying the change of variables $v_i = \tilde{\sigma}^{-1} \alpha^{-1} M^{-1/3} \tilde{v}_i + z_c$ and $w_i = \tilde{\sigma}^{-1} \alpha^{-1} M^{-1/3} \tilde{w}_i + z_c$ we can rewrite (A.17)

$$\det(I + K_u^0)_{L^2(C_{z_c, 3\pi/4})} = \det(I + K_x, M)_{L^2(C_{0, 3\pi/4})},$$  

where $K_x, M(\tilde{v}, \tilde{v}') = \frac{1}{2\pi i} \int_D g_{x, M}(\tilde{v}, \tilde{v}') d\tilde{w}.$

$D$ is as in Definition [A.7] and

$$g_{x, M}(\tilde{v}, \tilde{v}', \tilde{w}) = \frac{1}{\{ \max(\tilde{v}, |\tilde{w}|, |\tilde{v}'|) \leq \epsilon \tilde{\sigma}^{-1} M^{1/3} \}} \cdot \pi M^{-1/3} \tilde{\sigma}^{-1} \frac{e^{MG_u(\tilde{v}\tilde{\sigma}^{-1} M^{-1/3} + z_c) + x\tilde{w}}}{e^{MG_u(\tilde{v}\tilde{\sigma}^{-1} M^{-1/3} + z_c) + x\tilde{w}}}.$$

In view of Lemma [A.8] we have that there exists a constant $B > 0$ such that

$$|g_{x, M}(\tilde{v}, \tilde{v}', \tilde{w})| \leq B \cdot \exp \left( -\left(\frac{\sqrt{2}}{2} \right)^3 |\tilde{v}|^3 / 6 - \left(\sqrt{2}/2 \right)^3 |\tilde{w}|^3 / 6 \right),$$

where we used the fact that $|e^{x\tilde{v} - x\tilde{w}}| \leq 1$ (as $x \geq 0$) and $rac{\pi M^{-1/3} \tilde{\sigma}^{-1}}{\sin(\pi M^{-1/3} \tilde{\sigma}^{-1} |\tilde{v} - \tilde{w}|)}$ is bounded, cf. [BCD21a, Lemma 3.9]. Finally, we combine (A.15), (A.16) and (A.17) to get

$$\left| \det(I + K_u)_{L^2(C_{z_c, 3\pi/4})} - \det(I + K_x, M)_{L^2(C_{0, 3\pi/4})} \right| \leq e^{-c_1 M} + e^{-c_2 M}.$$  

**Step 2.** We first note that by repeating verbatim the proof of [BCF14, Lemma C.1] we have that

$$F_{GUE}(x) = \det(I + \hat{K}_x)_{L^2(C_{0, 3\pi/4})},$$

where $\hat{K}_x(\tilde{v}, \tilde{v}') = \frac{1}{2\pi i} \int_D e^{-\tilde{\theta} \tilde{v} + \tilde{\theta} \tilde{v}' / 3} e^{3/3 - x\tilde{w} + x\tilde{w}'} d\tilde{w}.$

Set $M_1 = \left[ \frac{\log M}{\log \log M} \right]$ and $\chi_n = \prod_{i=1}^n 1 \{ \max(|\tilde{v}_i|, |\tilde{w}_i|, |\tilde{v}_{\sigma(i)}|) \leq M_1 \} \text{ for } \sigma \in S_n.$ We claim that

$$0 = \lim_{M \to \infty} \sup_{x \geq 0} \sum_{n=1}^{M_1} \sum_{\sigma \in S_n} \frac{1}{n!} \int_{C_{0, 3\pi/4}} \cdots \int_{C_{0, 3\pi/4}} \int_D \cdots \int_D \chi_n \cdot \prod_{i=1}^n g_{x, M}(\tilde{v}_i, \tilde{w}_{\sigma(i)}, \tilde{w}_i) - \prod_{i=1}^n e^{-\tilde{\theta}_i / 3 + \tilde{\theta}_i / 3 - x\tilde{w}_i + x\tilde{w}_i} \int_{\tilde{v}_i - \tilde{w}_i} (\tilde{v}_i - \tilde{w}_{\sigma(i)}) \prod_{i=1}^n |d\tilde{v}_i| \prod_{i=1}^n |d\tilde{w}_i|.$$

We will prove (A.21) in the next step. Here we assume its validity and conclude the proof of (A.4).

In view of (A.18) and Lemma [A.10] we have for all large $M$ that

$$\left| \det(I + K_x, M)_{L^2(C_{0, 3\pi/4})} - 1 - \sum_{n=1}^{M_1} \sum_{\sigma \in S_n} \frac{1}{n!} \int_{C_{0, 3\pi/4}} \cdots \int_{C_{0, 3\pi/4}} \int_D \cdots \int_D \chi_n \cdot \prod_{i=1}^n g_{x, M}(\tilde{v}_i, \tilde{v}_{\sigma(i)}, \tilde{w}_i) \prod_{i=1}^n |d\tilde{v}_i| \prod_{i=1}^n |d\tilde{w}_i| \right| \leq e^{-(M_1/3) \log M_1},$$  

where
where we expanded the determinants. From (A.20) and Lemma A.10 we have for all large $M$ that

$$\left| F_{\text{GUE}}(x) - 1 - \sum_{n=1}^{M_1} \sum_{\sigma \in S_n} \frac{1}{n!} \int_{C_{0,3\pi/4}} \int_{C_{0,3\pi/4}} \int_{D} \cdots \int_{D} \frac{d\tilde{w}_1}{2\pi i} \cdots \frac{d\tilde{w}_n}{2\pi i} \right| \leq e^{-(M_1/3)\log M_1}. \tag{A.23}$$

Combining (A.22), (A.23) with (A.21) and (A.19) we conclude (A.4).

**Step 3.** In this step we prove (A.21). We first note that for all large enough $M$ (so that $M_1 \leq \epsilon\sigma \alpha M^{1/3}$), each $n \in \{1, \ldots, M_1\}$, $\sigma \in S_n$, $\tilde{v}_1 \in C_{0,3\pi/4}$ and $\tilde{w}_i \in D$ for $i = 1, \ldots, n$ we have

$$\chi_n \prod_{i=1}^{n} g_{x,M}(\tilde{v}_1, \tilde{v}_{\sigma(i)}), \tilde{w}_i \right) - \prod_{i=1}^{n} e^{-\tilde{v}_i^3/3 + \tilde{w}_i^3/3 - x\tilde{w}_i + x\tilde{v}_i} \prod_{i=1}^{n} \frac{d\tilde{w}_i}{2\pi i} \leq \chi_n \prod_{i=1}^{n} e^{-\tilde{v}_i^3/3 + \tilde{w}_i^3/3} \prod_{i=1}^{n} (\tilde{v}_i - \tilde{w}_i)(\tilde{w}_i - \tilde{v}_{\sigma(i)}) \right| \leq \chi_n \prod_{i=1}^{n} e^{-\tilde{v}_i^3/3 + \tilde{w}_i^3/3} $$

where we also used that $x \geq 0$. In addition, from (A.8) we know that for some $D_1 > 0$ we have

$$|MG_\alpha(z\tilde{\sigma}_\alpha^{-1} M^{-1/3} + z_c) + z^3/3| \leq D_1 M_1^4 M^{-1/3},$$

for all $z \in \mathbb{C}$ such that $|z| \leq M_1$. In addition, for some $D_2 > 0$ and all $\tilde{v} \in C_{0,3\pi/4}$ and $\tilde{w} \in D$

$$\left| 1 - \frac{(\tilde{v} - \tilde{w}) M^{-1/3} \tilde{\sigma}_\alpha^{-1}}{\sin(\pi M^{-1/3} \tilde{\sigma}_\alpha^{-1} (\tilde{v} - \tilde{w}))} \right| \leq D_2 M_1 M^{-1/3}, \text{ provided that } \max(|\tilde{v}|, |\tilde{w}|) \leq M_1.$$

Combining the latter inequalities, the fact that $n \leq M_1$ and the fact that $M_1 = \lceil \log M / \log \log M \rceil$ we conclude that there is a constant $D_3 > 0$ such that for all large $M$

$$\chi_n \prod_{i=1}^{n} g_{x,M}(\tilde{v}_1, \tilde{v}_{\sigma(i)}, \tilde{w}_i) \right) - \prod_{i=1}^{n} e^{-\tilde{v}_i^3/3 + \tilde{w}_i^3/3 - x\tilde{w}_i + x\tilde{v}_i} \prod_{i=1}^{n} \frac{d\tilde{w}_i}{2\pi i} \leq D_3 \cdot \chi_n \cdot M_1^5 M^{-1/3},$$

where in the last inequality we used that $|\tilde{v} - \tilde{w}| \geq 1$ for $\tilde{v} \in C_{0,3\pi/4}$ and $\tilde{w} \in D$. Combining the above estimates we conclude that the right side of (A.21) is upper bounded by

$$\limsup_{M \to \infty} M^{\epsilon_0} \sum_{n=1}^{M_1} D_3 M_1^5 M^{-1/3} D_4^n,$$

where $D_4 = \int_{C_{0,3\pi/4}} \int_D |e^{-\tilde{v}^3/3 + \tilde{w}^3/3}| |d\tilde{w}| |d\tilde{v}|$. Since the above sum is at most $D_3 (D_4+1)^{M_1} M_1^5 M^{\epsilon_0-1/3}$ and $M_1 = \lceil \log M / \log \log M \rceil$ we see that the limit is 0, which proves (A.21).

**References**

[And58] P. W. Anderson, *Absence of diffusion in certain random lattices*, Phys. Rev. 109 (1958), no. 5, 1492.

[AS65] M. Abramowitz and I.A. Stegun, *Handbook of mathematical functions*, Dover, New York, 1965.

[AW15] M. Aizenman and S. Warzel, *Random operators*, vol. 168, American Mathematical Soc., 2015.

[BC14] A. Borodin and I. Corwin, *Macdonald processes*, Probab. Theory Relat. Fields 158 (2014), 225–400.

[BC17] G. Barraquand and I. Corwin, *Random-walk in beta-distributed random environment*, Probab. Theory Relat. Fields 167 (2017), no. 3, 1057–1116.

[BCD21a] G. Barraquand, I. Corwin, and E. Dimitrov, *Fluctuations of the log-gamma polymer free energy with general parameters and slopes*, Probab. Theo. Rel. Fields 181 (2021), 113–195.

[BCD21b] ______, *Spatial tightness at the edge of Gibbsian line ensembles*, arXiv:2101.03045.
A. Borodin, I. Corwin, and P. L. Ferrari, *Free energy fluctuations for directed polymers in random media in 1 + 1 dimension*, Commun. Pure Appl. Math. **67** (2014), 1129–1214.

A. Borodin, I. Corwin, and D. Remenik, *Log-gamma polymer free energy fluctuations via a Fredholm determinant identity*, Comm. Math. Phys. **324** (2013), 215–232.

P. Billingsley, *Convergence of probability measures, Second edition*, Academic Press, New York, 1999.

M. Bóna, M.-L. Lackner, and B. E. Sagan, *Longest increasing subsequences and log concavity*, Ann. Comb. **21** (2017), 535–549.

I. Corwin and P. Ghosal, *Lower tail of the KPZ equation*, Duke Math. J. **169** (2020), no. 7, 1329 – 1395.

I. Corwin, P. Ghosal, and A. Hammond, *KPZ equation correlations in time*, Ann. Probab. **49** (2021), 832–876.

I. Corwin, *Invariance of polymer partition functions under the geometric RSK correspondence*, Adv. Stud. Pure Math. **87** (2021), 89–136.

I. Corwin, N. O’Connell, T. Seppäläinen, and N. Zygouras, *Tropical combinatorics and Whittaker functions*, Duke Math. J. **163** (2014), 513–563.

E. Dimitrov, *KPZ and Airy limits of Hall-Littlewood random plane partitions*, Ann. Inst. Henri Poincaré Probab. Stat. **54** (2017), no. 2, 640–693.

D. Dauvergne, J. Ortmann, and B. Virág, *The directed landscape*, arXiv:1812.00309.

L. Dumaz and B. Virág, *The scaling limit of the longest increasing subsequence*, arXiv:2104.08210.

Y. Luke, *The special functions and their approximations, Volume I*, Academic press, London, 1969.

D. Mason and H. Zhou, *Quantile coupling inequalities and their applications*, Probab. Surveys **9** (2012), 439–479.

A. M. Somoz, M. Ortuno, and J. Prior, *Unbinding transition in semi-infinite two-dimensional localized systems*, Phys. Rev. B **91** (2015), no. 15, 155413.

A.M. Somoz, M. Ortuno, and J. Prior, *Universal distribution functions in two-dimensional localized systems*, Phys. Rev. Lett. **99** (2007), no. 11, 116602.

A. Scardicchio and T. Thiery, *Perturbation theory approaches to Anderson and many-body localization: some lecture notes*, arXiv:1710.01234 (2017).

C. Tracy and H. Widom, *Level-spacing distributions and the Airy kernel*, Commun. Math. Phys. **159** (1994), 151–174.

X. Wu, *Discrete Gibbsian line ensembles and weak noise scaling for directed polymers*, PhD Thesis, https://academiccommons.columbia.edu/doi/10.7916/d8-6re1-k703.