Electron. Commun. Probab. **23** (2018), no. 101, 1–12.
https://doi.org/10.1214/18-ECP201
ISSN: 1083-589X

ELECTRONIC COMMUNICATIONS in PROBABILITY

Quenched central limit theorem in a corner growth setting

H. Christian Gromoll* Mark W. Meckes† Leonid Petrov‡

Abstract

We consider point-to-point directed paths in a random environment on the two-dimensional integer lattice. For a general independent environment under mild assumptions we show that the quenched energy of a typical path satisfies a central limit theorem as the mesh of the lattice goes to zero. Our proofs rely on concentration of measure techniques and some combinatorial bounds on families of paths.

Keywords: last passage percolation; central limit theorem; concentration of measure.
AMS MSC 2010: Primary 60F05, Secondary 60K35.
Submitted to ECP on August 3, 2018, final version accepted on November 29, 2018.

1 Introduction and main results

A number of well-known probabilistic models derive their underlying complexity from a variant of the following simple setup. Put independent and identically distributed weights at each vertex of the two-dimensional integer lattice. Given a lattice point in the first quadrant, consider all paths in the lattice from the origin to this point that only move up or to the right at each step. Each such path has a random energy given by the sum of the weights along the path, and so the collection of random energies indexed by the up-right paths exhibits a complicated dependence structure. This dependence is at the heart of the difficulty in understanding such models as the totally asymmetric simple exclusion process (TASEP), the infinite tandem queue, the random directed polymer, or the corner growth model.

For example in the corner growth model, or directed nearest neighbor last-passage percolation on the 2d lattice, the fundamental issue is to understand the distribution of the maximum-energy path to a given point. This maximal energy represents the last passage time to the point, or time at which it joins the growing corner shape. In directed polymer models one assigns a Boltzmann weight to each path according to its random energy, and one is concerned with the sum of all polymer weights, or partition function,

*University of Virginia, Department of Mathematics, 141 Cabell Drive, Kerchof Hall, P.O. Box 400137, Charlottesville, VA 22904, USA
E-mail: gromoll@virginia.edu
†epartment of Mathematics, Applied Mathematics, and Statistics, Case Western Reserve University, 10900 Euclid Ave., Cleveland, OH 44106, USA
E-mail: mark.meckes@case.edu
‡University of Virginia, Department of Mathematics, 141 Cabell Drive, Kerchof Hall, P.O. Box 400137, Charlottesville, VA 22904, USA, and Institute for Information Transmission Problems, Bolshoy Karetny per. 19, Moscow, 127994, Russia
E-mail: lenia.petrov@gmail.com

which normalizes the polymer weights into probabilities. (See e.g. [5] for a description of the model as well as its relationship to equivalent models). Both are difficult models because of the high degree of dependence in the joint distribution of the collection of path energies. A more detailed discussion of these models as they relate to our result appears further below.

In this paper we derive a result about the joint distribution of the path energies which to our knowledge has not been observed previously. Namely, we show that conditional on the environment of weights, the empirical distribution of the family of path energies is approximately Gaussian. More precisely, for almost every environment, the energy of an up-right path selected uniformly at random is asymptotically normally distributed as the mesh gets small.

This is the content of our main theorem, which is proved for generally distributed weights with nonzero variance and under further moment assumptions (in fact, in our setting the weights need not even be identically distributed). We now introduce some notation, describe our results, and then discuss connections to the corner growth and directed random polymer models.

1.1 Up-right paths in a random environment

We work inside the positive quadrant $\mathbb{Z}^2_{\geq 1}$ of the two-dimensional integer lattice. By convention, coordinates $(i, j) \in \mathbb{Z}^2_{\geq 1}$ refer to squares, see Figure 1. A (fixed) *environment* w is an assignment of a real number $w_{ij} \in \mathbb{R}$ to every square of $\mathbb{Z}^2_{\geq 1}$. We call the w_{ij} *weights*.

Fix integers $M, N \geq 1$ and consider the rectangle $1 \leq i \leq M$, $1 \leq j \leq N$ inside the quadrant. Denote it by $\square_{M,N}$. The environment w restricted to $\square_{M,N}$ is a vector in $\mathbb{R}^{MN}$.

An *up-right path* σ from $(1, 1)$ to (M, N) is a collection $\{(i_k, j_k) \colon k = 1, \ldots, M + N - 1\}$ such that $(i_{k+1} - i_k, j_{k+1} - j_k)$ is either $(1, 0)$ (horizontal step) or $(0, 1)$ (vertical step), $(i_1, j_1) = (1, 1)$, and $(i_{M+N-1}, j_{M+N-1}) = (M, N)$. To each up-right path σ we associate a vector

$$Y^\sigma \in \mathbb{R}^{MN}, \qquad Y^\sigma_{ij} := \mathbf{1}_{(i,j)\in\sigma}, \qquad 1 \leq i \leq M,\ 1 \leq j \leq N.$$

Here and below $\mathbf{1}_A$ is the indicator of A.

For any up-right path $\sigma = \{(i_k, j_k)\}$ define its *energy* with respect to an environment w as

$$\langle Y^\sigma, w\rangle = \sum_{k=1}^{M+N-1} w_{i_k j_k}, \tag{1.1}$$

where $\langle \cdot, \cdot \rangle$ is the standard inner product in $\mathbb{R}^{MN}$.

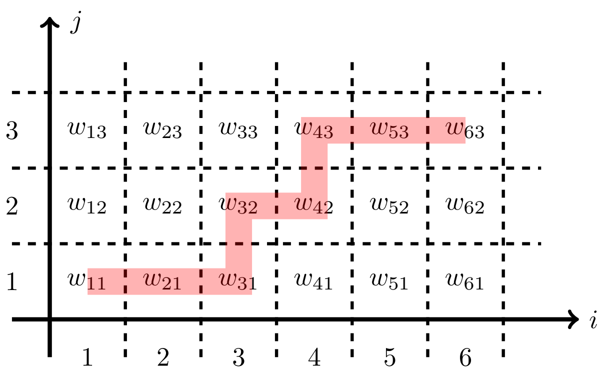

Figure 1: Environment and an up-right path from $(1, 1)$ to $(M, N) = (6, 3)$.

Now suppose both the environment and the path are chosen at random, independently of each other. Assume the random environment w consists of independent random variables w_{ij} defined on a probability space $(\Omega_w, \mathcal{F}_w, \mathbb{P}_w)$ with

$$\mathbb{E}_w w_{ij} = 0, \qquad \mathbb{E}_w w_{ij}^2 = 1, \qquad \mathbb{E}_w |w_{ij}|^{\mathsf{p}} \leq \mathsf{K}, \qquad \text{for all } i, j, \tag{1.2}$$

for some p (to be specified later) and $\mathsf{K} > 0$. Here $\mathbb{E}_w$ is expectation with respect to $\mathbb{P}_w$.

Assume that σ is a random up-right path chosen (according to some distribution) from all paths inside a given subset $\Sigma_{M,N} \subseteq \Box_{M,N}$. A simple example is when $\Sigma_{M,N} = \Box_{M,N}$ and σ is chosen uniformly from all $\binom{M+N-2}{M-1}$ possible paths. In any case, σ is a random path defined on some $(\Omega_\sigma, \mathcal{F}_\sigma, \mathbb{P}_\sigma)$, and expectation with respect to $\mathbb{P}_\sigma$ will be denoted $\mathbb{E}_\sigma$.

We thus think of the energy of a path $\langle Y^\sigma, w\rangle$ in (1.1) as a random variable defined on $(\Omega_w \times \Omega_\sigma, \mathcal{F}_w \times \mathcal{F}_\sigma, \mathbb{P}_w \times \mathbb{P}_\sigma)$ which depends on both the randomness in the environment and in the path (the path is independent from the environment). The goal of this paper is to show that for certain natural distributions of σ and $\mathbb{P}_w$-almost every environment, the (quenched) random variable $\langle Y^\sigma, w\rangle$ depending on the random path σ is asymptotically Gaussian as $M, N \to \infty$. A precise formulation is given next.

1.2 Quenched central limit theorems

Let the environment $w = \{w_{ij}\}$ consist of independent random variables satisfying (1.2). Let $M = \lfloor \xi N \rfloor$, where $\xi > 0$ is fixed.

Theorem 1.1. *If $\mathsf{p} > 12$ and σ is chosen uniformly from all up-right paths in the rectangle $\Box_{M,N}$, then $\mathbb{P}_w$-almost surely,*

$$\frac{1}{\sqrt{M+N-1}} \sum_{(i,j)\in\sigma} w_{ij} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1), \qquad N \to \infty, \tag{1.3}$$

where the convergence in distribution to the standard normal is with respect to the marginal $\mathbb{P}_\sigma$.

In other words, for large N and almost every environment w, the empirical distribution of the family of path energies will be approximately Gaussian. The assumption $\mathsf{p} > 12$ may seem unexpected. As will be seen in the proofs, it arises from a combination of the moment assumptions needed for our concentration inequality and our path counting estimates.

Corollary 1.2. *The statement of Theorem 1.1 remains valid (still with $\mathsf{p} > 12$) when σ is chosen uniformly from all up-right paths passing through each of the points $(\lfloor \xi_i N\rfloor, \lfloor \zeta_i N\rfloor)$, $i = 1, \dots, \ell$ (for finite ℓ), where $0 < \xi_1 < \dots < \xi_\ell < \xi$ and $0 < \zeta_1 < \dots < \zeta_\ell < 1$ are fixed. See Figure 2, (a).*

Theorem 1.1 and Theorem 1.2 are proven in Section 3.1.

To illustrate that our approach can yield similar results for other path families, as long as suitable path counting arguments are available, we will also prove the analogous result for the family of paths avoiding a hole of fixed proportion in the center of $\Box_{M,N}$. For simplicity in Theorem 1.3 below we assume that $M = N$ (though a suitably modified statement can be established for $M = \lfloor \xi N \rfloor$ as well). Let $B = \lfloor \beta N \rfloor$, where $\beta \in (0,1)$ is fixed, and define the subset $\Sigma_{N,N} = \{(i,j) \in \Box_{N,N} \colon \max(|i - N/2|, |j - N/2|) \geq B/2\}$; see Figure 2, (b).

Theorem 1.3. *If $\mathsf{p} > 12$, and σ be chosen uniformly from the set of up-right paths that remain in $\Sigma_{N,N}$, then $\mathbb{P}_w$-almost surely,*

$$\frac{1}{\sqrt{2N-1}} \sum_{(i,j)\in\sigma} w_{ij} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1), \qquad N \to \infty. \tag{1.4}$$

This theorem is proven in Section 3.2.

Remark 1.4. In (1.2) we assumed that the environment random variables w_{ij} have mean 0 and variance 1. By shifting and scaling the energies $\langle Y^\sigma, w\rangle$ one readily sees that all our results formulated above can be extended to independent environments with arbitrary constant mean and nonzero variance.

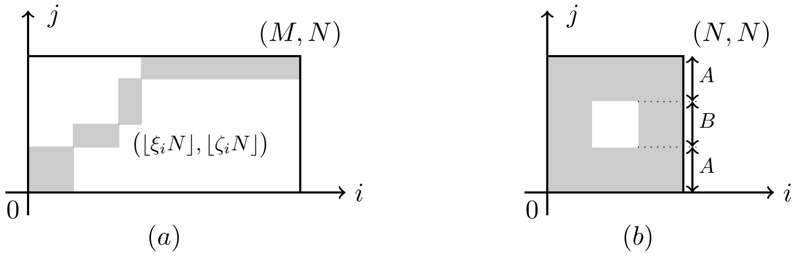

Figure 2: Subsets of the rectangle from which the up-right path σ is chosen uniformly in (a) Theorem 1.2, (b) Theorem 1.3.

1.3 Relation to other models

We now briefly compare our setting with other models based on directed up-right paths in a random environment. Consider first the directed polymer models introduced in statistical physics in [9]; see also e.g. [10], [16]. In the lattice setting, the directed polymer partition function is defined as (we continue to assume that $M = N$)

$$Z_N(w) = \sum_\sigma \exp\Big\{\beta \sum_{(i,j)\in\sigma} w_{ij}\Big\} = \sum_\sigma \exp\big\{\beta\,\langle Y^\sigma, w\rangle\big\},$$

where the outer sum is taken over all up-right paths σ inside $\square_{N,N}$, and $\beta > 0$ is the inverse temperature. The polymer weight of a path σ is defined as

$$Q_N(\sigma; w) = \frac{1}{Z_N(w)} \exp\big\{\beta\,\langle Y^\sigma, w\rangle\big\}.$$

The study of the asymptotic behavior of Z_N and Q_N as $N \to \infty$ has received a lot of attention in the past 30 years. Of particular interest are the asymptotic fluctuations of the free energy $\log Z_N(w)$. These fluctuations are expected to grow as $N^{1/3}$ under mild assumptions. However, this scaling behavior is currently known only for a number of integrable cases (that is for special choices of the distribution of w_{ij} leading to exact formulae for the Laplace transform of Z_N); see [14], [3], [4]. In integrable cases, the fluctuations themselves are governed by one of the Tracy-Widom distributions [19], which is characteristic for the Kardar-Parisi-Zhang universality [5]. Study of the asymptotic fluctuations of $\log Z_N(w)$ when the integrability is not known presents a major open problem in the field.

Passing to the zero temperature limit $\beta \to \infty$ turns the free energy $\log Z_N(w)$ into the last passage percolation time:

$$G_N(w) = \max_\sigma \langle Y^\sigma, w\rangle, \tag{1.5}$$

where the maximum is taken over all up-right paths inside $\square_{N,N}$. Assume that the environment variables w_{ij} are nonnegative. This does not significantly restrict generality since if the distribution of w_{ij} is bounded from below, one can achieve nonnegativity by adding a fixed constant to all the w_{ij}. We can then interpret the nonnegative w_{ij}

as random waiting times in the corner growth model so that (1.5) becomes the time at which the growing interface covers (N, N). For further details on corner growth we refer to [11], [15], [5], [2], [17].

Asymptotic fluctuations of $G_N(w)$ in integrable cases (when the w_{ij}'s have exponential or geometric distribution) have been shown to converge on scale $N^{1/3}$ to the Tracy-Widom distribution [11]. Again, the problem of asymptotic fluctuations in the corner growth model with other distributions of w_{ij} (for which exact formulae are not known to exist) is open.

Our results (in particular, (1.3)) mean that that the path energies $\langle Y^\sigma, w\rangle$ asymptotically behave as $(2N-1)\mathbb{E}w_{ij} + \zeta\sqrt{2N-1}$ (with ζ Gaussian). Let us compare this with the order of $G_N(w)$ known exactly for special distributions of w_{ij} from, e.g., [11]. (A similar comparison may be performed in the polymer case, but we omit it.) When w_{ij} are geometric, $\mathbb{P}(w_{ij} = k) = (1-q)q^k$, $k \in \mathbb{Z}_{\geq 0}$, we have (ignoring fluctuations) $\mathbb{E}G_N(w) \sim N\frac{2\sqrt{q}}{1-\sqrt{q}}$, whereas typical values of $\langle Y^\sigma, w\rangle$ are of order $(2N-1)\mathbb{E}w_{ij} \sim N\frac{2q}{1-q}$, which is smaller (however, the difference goes to zero as $q \searrow 0$). Similarly, for w_{ij} exponential with mean 1, the last passage time behaves as $\mathbb{E}G_N(w) \sim 4N$, and typical values of $\langle Y^\sigma, w\rangle$ are of order $(2N-1)\mathbb{E}w_{ij} \sim 2N$.

Therefore, while our results indicate that the asymptotic quenched behavior of the typical values of the path energies $\langle Y^\sigma, w\rangle$ is universal (i.e., does not depend on the distribution of the w_{ij} under mild assumptions), this conclusion does not extend to extreme values of $\langle Y^\sigma, w\rangle$ responsible for the asymptotics of the last passage time $G_N(w)$.

In Section 2 we employ Talagrand's concentration inequality to establish a quenched central limit theorem (Theorem 2.2) modulo an estimate on the distribution of the up-right path σ. In Section 3 we obtain the needed combinatorial estimates for natural ensembles of up-right paths described in Section 1.2 above, and complete the proofs of Theorem 1.1, Theorem 1.3, and Theorem 1.2.

2 Gaussian concentration and quenched central limit theorems

In this section we focus on general concentration estimates, and establish our quenched CLT modulo combinatorial estimates which are postponed until Section 3.

2.1 General concentration lemma

Let us work in a more general setting. Suppose Σ is a set with n elements equipped with independent weights $\{w_a \mid a \in \Sigma\}$ satisfying conditions (1.2) for $\mathsf{p} \geq 3$. For a fixed $R > 0$ we define the truncations

$$w_a^{(R)} := w_a \mathbf{1}_{|w_a| \leq R}. \tag{2.1}$$

Let σ be a random subset of Σ having almost surely m elements. As before, $\mathbb{P}_w$ and $\mathbb{P}_\sigma$ stand for the marginal probability measures corresponding to $\{w_a\}$ and σ, respectively, and similarly for expectations. Let $Y_a^\sigma := \mathbf{1}_{a\in\sigma}$, and let Y^σ denote the corresponding random vector in $\mathbb{R}^n$. Define

$$L := \left\|\mathbb{E}_\sigma Y^\sigma\right\|_2 = \left(\sum_{a\in\Sigma} \mathbb{P}_\sigma\left(a \in \sigma\right)^2\right)^{\frac{1}{2}}. \tag{2.2}$$

Next, let $\gamma \sim \mathcal{N}(0,1)$ denote the standard Gaussian measure on $\mathbb{R}$, and let μ_w and $\mu_w^{(R)}$ be the quenched distributions (conditioned on w) of the random variables

$$\frac{\langle Y^\sigma, w\rangle}{\sqrt{m}} = \frac{1}{\sqrt{m}}\sum_{a\in\sigma} w_a, \qquad \frac{\langle Y^\sigma, w^{(R)}\rangle}{\sqrt{m}} = \frac{1}{\sqrt{m}}\sum_{a\in\sigma} w_a^{(R)}, \tag{2.3}$$

respectively. In particular, for a test function $f\colon \mathbb{R} \to \mathbb{R}$ we have

$$\int f\, d\mu_w = \mathbb{E}_\sigma f\left(\frac{\langle Y^\sigma, w\rangle}{\sqrt{m}}\right) = \mathbb{E}_\sigma f\left(\frac{1}{\sqrt{m}} \sum_{a\in\sigma} w_a\right).$$

Lemma 2.1. *Under the above assumptions, there exist absolute constants $C, c > 0$ (not depending on parameters of the model) and a constant $\kappa > 0$ depending only on K and p in (1.2) such that for any $s, t, R > 0$ and any convex 1-Lipschitz*[1] *function $f\colon \mathbb{R} \to \mathbb{R}$ we have*

$$\mathbb{P}_w\left[\left|\int f\, d\mu_w - \int f\, d\gamma\right| \geq \frac{\kappa}{\sqrt{m}} + \sqrt{\frac{\mathsf{K}nL^2}{mR^{\mathsf{p}-2}}} + s + t\right] \leq \frac{\mathsf{K}nL^2}{mR^{\mathsf{p}-2}s^2} + C\exp\left[-c\,\frac{mt^2}{L^2R^2}\right]. \tag{2.4}$$

Proof. We have

$$\begin{aligned}
\left|\int f\, d\mu_w - \int f\, d\gamma\right| \leq & \left|\int f\, d\mu_w - \int f\, d\mu_w^{(R)}\right| && (2.5)\\
&+ \left|\int f\, d\mu_w^{(R)} - \mathbb{E}_w \int f\, d\mu_w^{(R)}\right| && (2.6)\\
&+ \left|\mathbb{E}_w \int f\, d\mu_w^{(R)} - \mathbb{E}_w \int f\, d\mu_w\right| && (2.7)\\
&+ \left|\mathbb{E}_w \int f\, d\mu_w - \int f\, d\gamma\right|. && (2.8)
\end{aligned}$$

The terms (2.5) and (2.7) on the right will be estimated by elementary methods, (2.6) via Talagrand's concentration inequality for independent bounded random variables, and (2.8) via an appropriate version of the central limit theorem. Note that the terms (2.7) and (2.8) are deterministic.

We start with the last term (2.8). As a consequence of results of Esseen [6], we have a bound on Wasserstein-1 distance that is independent of the path σ; see for example [7, Proposition 2.2] for a statement of this bound that depends only on finite third moments (as we assume here). Note also that the L_1-distance of distribution functions as used in [7] is equivalent to the Wassertein-1 distance in our setting since we use 1-Lipschitz test functions. For all σ (recall that σ has m elements a.s.) we have

$$\left|\mathbb{E}_w f\left(\frac{1}{\sqrt{m}} \sum_{a\in\sigma} w_a\right) - \int f\, d\gamma\right| \leq \frac{\kappa}{\sqrt{m}}, \tag{2.9}$$

where $\kappa > 0$ depends only on the third absolute moment of w_a and can therefore be bounded in terms of the constants p and K from (1.2). Since (2.9) holds for all σ, an application of Fubini's theorem implies that

$$\left|\mathbb{E}_w \int f\, d\mu_w - \int f\, d\gamma\right| = \left|\mathbb{E}_w \mathbb{E}_\sigma f\left(\frac{1}{\sqrt{m}} \sum_{a\in\sigma} w_a\right) - \int f\, d\gamma\right| \leq \frac{\kappa}{\sqrt{m}}. \tag{2.10}$$

(In particular, this implies that the expectation $\mathbb{E}_w \int f\, d\mu_w$ in (2.8) is finite.)

Let us turn to (2.6). Consider the function $F\colon \mathbb{R}^\Sigma \to \mathbb{R}$ defined by

$$F(w) := \int f\, d\mu_w = \mathbb{E}_\sigma f\left(\frac{\langle Y^\sigma, w\rangle}{\sqrt{m}}\right).$$

[1] Recall that a function $f\colon \mathbb{R} \to \mathbb{R}$ is called 1-Lipschitz if $|f(x) - f(y)| \leq |x - y|$ for all $x, y \in \mathbb{R}$.

Since f is convex, F is an average of convex functions on $\mathbb{R}^\Sigma$, and so is also convex. Let us now estimate the Lipschitz constant of F. We have for $w, w' \in \mathbb{R}^\Sigma$:

$$\begin{aligned}
|F(w) - F(w')| &\le \mathbb{E}_\sigma \left| f\left(\frac{\langle Y^\sigma, w\rangle}{\sqrt{m}}\right) - f\left(\frac{\langle Y^\sigma, w'\rangle}{\sqrt{m}}\right)\right| \\
&\le \frac{1}{\sqrt{m}} \mathbb{E}_\sigma \left| \langle Y^\sigma, w\rangle - \langle Y^\sigma, w'\rangle \right| \\
&\le \frac{1}{\sqrt{m}} \mathbb{E}_\sigma \langle Y^\sigma, |w - w'|\rangle \\
&= \frac{1}{\sqrt{m}} \langle \mathbb{E}_\sigma Y^\sigma, |w - w'|\rangle \\
&\le \frac{1}{\sqrt{m}} L\|w - w'\|_2. \qquad (2.11)
\end{aligned}$$

Here the third line follows because Y^σ has nonnegative components, $|w - w'|$ denotes componentwise absolute value, and the last line follows from the Cauchy–Schwarz inequality. Thus, we see that F has Lipschitz constant at most $L/\sqrt{m}$.

Using this and applying Talagrand's inequality for bounded independent random variables [18][2] to $F(w^{(R)}) = \int f\, d\mu_w^{(R)}$ we obtain for any $t > 0$ and for some absolute constants $C, c > 0$,

$$\mathbb{P}_w \left[\left| \int f\, d\mu_w^{(R)} - \mathbb{E}_w \int f\, d\mu_w^{(R)} \right| \ge t \right] \le C \exp\left[-c \frac{mt^2}{L^2 R^2} \right]. \qquad (2.12)$$

It remains to consider terms (2.5) and (2.7). From the Lipschitz estimate (2.11) we have

$$\left| \int f\, d\mu_w - \int f\, d\mu_w^{(R)} \right| = \left| F(w) - F(w^{(R)}) \right| \le \frac{L}{\sqrt{m}} \left\| w - w^{(R)} \right\|_2 = \frac{L}{\sqrt{m}} \left(\sum_{a \in \Sigma} w_a^2 \mathbf{1}_{|w_a| > R} \right)^{\frac{1}{2}}.$$

Utilizing Hölder and Chebyshev inequalities, we can write

$$\mathbb{E}_w \left(w_a^2 \mathbf{1}_{|w_a| > R} \right) \le \left(\mathbb{E}_w |w_a|^{\mathsf{p}} \right)^{\frac{2}{\mathsf{p}}} \left(\mathbb{P}_w \left[|w_a| > R \right] \right)^{1 - \frac{2}{\mathsf{p}}} \le \frac{\mathbb{E}_w |w_a|^{\mathsf{p}}}{R^{\mathsf{p}(1-2/\mathsf{p})}} \le \frac{\mathsf{K}}{R^{\mathsf{p}-2}},$$

and so for any $u > 0$ we have

$$\mathbb{P}_w \left[\sum_{a \in \Sigma} w_a^2 \mathbf{1}_{|w_a| > R} \ge u \right] \le \frac{n\mathsf{K}}{uR^{\mathsf{p}-2}}.$$

Therefore, we have

$$\left| \mathbb{E}_w \int f\, d\mu_w - \mathbb{E}_w \int f\, d\mu_w^{(R)} \right| \le L \sqrt{\frac{\mathsf{K}n}{mR^{\mathsf{p}-2}}}, \qquad (2.13)$$

and

$$\mathbb{P}_w \left[\left| \int f\, d\mu_w - \int f\, d\mu_w^{(R)} \right| \ge s \right] \le \frac{n\mathsf{K}L^2}{mR^{\mathsf{p}-2} s^2}. \qquad (2.14)$$

The lemma now follows by combining (2.10), (2.12), (2.13), and (2.14). $\square$

[2] See also [12, Corollary 1.3] and references and discussion therein.

2.2 Quenched central limit theorem

Our aim now is to apply the general Lemma 2.1 to obtain the quenched central limit theorem in the corner growth setting described in Section 1. Recall that we choose N (the vertical dimension of the rectangle in Figure 1) to be the main parameter going to infinity, and we let $M = \lfloor \xi N \rfloor$ for some $\xi > 0$. The parameters in Lemma 2.1 are instantiated as follows:

$$n(N) = |\Sigma_{M,N}| \le MN \sim \xi N^2, \qquad m(N) = M + N - 1 \sim (\xi+1)N, \qquad L = L(N). \quad (2.15)$$

Here we assume that for each N, $\Sigma_{M,N} \subseteq \square_{M,N}$ is a given subset and that σ is chosen according to some distribution such that $\sigma \in \Sigma_{M,N}$ almost surely (it replaces the set Σ in Lemma 2.1). We further assume that as $N \to \infty$,

$$\begin{aligned} n(N) &= |\Sigma_{M,N}| = O(N^\eta), \\ L(N) &= \Bigg(\sum_{a \in \Sigma_{M,N}} \mathbb{P}_\sigma \left(a \in \sigma \right)^2 \Bigg)^{\frac{1}{2}} = O(N^\lambda), \end{aligned} \quad (2.16)$$

for some $0 < \eta \le 2$ and $0 < \lambda \le \eta/2$.[3] For the specific subsets $\Sigma_{M,N}$ considered in this paper, the parameter $L(N)$ will be estimated separately in Section 3 below. Note that the constants $C, c, \mathsf{K}, \mathsf{p}, \kappa$ in Lemma 2.1 are independent of N.

Theorem 2.2. *Under the above assumptions, if $\lambda < \frac{1}{2}$ and $\mathsf{p} > \frac{6}{1-2\lambda}$ then $\mathbb{P}_w$-almost surely,*

$$\frac{1}{\sqrt{m(N)}} \sum_{a \in \sigma} w_a \xrightarrow{\mathcal{D}} \mathcal{N}(0,1), \qquad N \to \infty,$$

where the convergence in distribution to the standard normal is with respect to the marginal $\mathbb{P}_\sigma$.

Proof. To get the desired $\mathbb{P}_w$-almost sure convergence in distribution, we will choose the parameters R, s, t in Lemma 2.1 depending on N and apply the Borel-Cantelli lemma.[4] That is, from the left side of (2.4) we see that we must have

$$\lim_{N\to\infty} \left[\frac{\kappa}{\sqrt{\xi+1}} \frac{1}{\sqrt{N}} + \sqrt{\frac{\mathsf{K}}{(\xi+1)} \frac{\sqrt{n(N)} L(N)}{\sqrt{N} R(N)^{\frac{\mathsf{p}-2}{2}}}} + s(N) + t(N) \right] = 0,$$

which is equivalent to

$$\lim_{N\to\infty} \frac{n(N) L(N)^2}{N R(N)^{\mathsf{p}-2}} = 0, \qquad \lim_{N\to\infty} s(N) = \lim_{N\to\infty} t(N) = 0. \quad (2.17)$$

Moreover, to use Borel-Cantelli the right side of (2.4) must be summable, which is equivalent to

$$\sum_{N=N_0}^{\infty} \frac{n(N) L(N)^2}{N R(N)^{\mathsf{p}-2} s(N)^2} < \infty, \qquad \sum_{N=N_0}^{\infty} \exp\left[-c(\xi+1) \frac{N t(N)^2}{L(N)^2 R(N)^2} \right] < \infty, \quad (2.18)$$

for some absolute constant N_0.

[3] The fact that we must have $\lambda \le \eta/2$ follows by taking the trivial estimate $\mathbb{P}_\sigma(a \in \sigma) \le 1$ for all a.

[4] Convex 1-Lipschitz test functions are enough to conclude convergence in distribution. First, we have tightness since the first moments converge. By the Weierstrass theorem, compactly supported test functions can be approximated by polynomials (on a compact set), and polynomials are linear combinations of convex 1-Lipschitz functions. See also, e.g., [8], [13] for slightly different approaches.

Let $R(N) \sim N^{\rho}$ for some $\rho > 0$. Then for (2.17) and (2.18) to hold under our assumption (2.16) it is necessary and sufficient that

$$\eta + 2\lambda - 1 - \rho(\mathsf{p} - 2) < -1, \qquad 1 - 2\lambda - 2\rho > 0, \tag{2.19}$$

and that $s(N)$ and $t(N)$ tend to zero sufficiently slowly as negative powers of N. Setting $\rho = \frac{1}{2} - \lambda - \varepsilon$ for small enough $\varepsilon > 0$, one can check that condition (2.19) (together with our assumptions $\eta \le 2$ and $\lambda \le \eta/2$ coming from (2.2) and (2.15)) is equivalent to

$$0 < \lambda < \frac{1}{2}, \qquad \mathsf{p} > \frac{2 + 2\eta}{1 - 2\lambda}.$$

The latter inequality holds if $\mathsf{p} > \frac{6}{1-2\lambda}$, which completes the proof. □

3 Path counting

Our goal in this section is to obtain estimates of $L(N)$ of the form (2.16) with $\lambda < \frac{1}{2}$ for the concrete families of up-right paths described in Section 1.2. These estimates lead to quenched central limit theorems. Similar estimates have appeared in the context of random polymers before, for example, see [1, Appendix A].

3.1 All possible paths – proofs of Theorem 1.1 and Theorem 1.2

We start with the case when σ is chosen uniformly at random from the set of all possible up-right paths in the rectangle $\square_{M,N}$, so $\Sigma_{M,N} = \square_{M,N}$. Let us slice the rectangle as follows:

$$\square_{M,N} = \bigsqcup_{k=2}^{M+N} \square_{M,N}^{(k)}, \qquad \square_{M,N}^{(k)} := \{a = (i,j)\colon 1 \le i \le M,\, 1 \le j \le N,\, i + j = k\}. \tag{3.1}$$

We can write

$$L(N)^2 = \sum_{k=2}^{M+N} \sum_{a \in \square_{M,N}^{(k)}} \mathbb{P}_{\sigma}\left(a \in \sigma\right)^2 \le \sum_{k=2}^{M+N} \max_{a \in \square_{M,N}^{(k)}} \mathbb{P}_{\sigma}\left(a \in \sigma\right) = \sum_{k=2}^{M+N} \mathcal{M}_k, \tag{3.2}$$

where we have denoted $\mathcal{M}_k := \max_{a \in \square_{M,N}^{(k)}} \mathbb{P}_{\sigma}\left(a \in \sigma\right)$.

Remark 3.1. The estimate (3.2) holds for any distribution of the up-right path σ. Moreover, since the maximum probability over $\square_{M,N}^{(k)}$ can always be bounded by 1, we have the trivial estimate $L(N)^2 \le M + N - 1$. In our regime (M proportional to N) this estimate leads to $L(N) = O(N^{\frac{1}{2}})$ and so $\lambda = \frac{1}{2}$, which is not quite good enough for our purpose.

We will show however that one can in fact take $\lambda = \frac{1}{4}$ using a better estimate of $\mathcal{M}_k$. Recall that $M = \lfloor \xi N \rfloor$, $\xi > 0$.

Lemma 3.2. *Let σ be chosen uniformly from all possible paths in $\square_{M,N}$. There exists $\mathsf{C} > 0$ such that for all N large enough, $\sum_{k=2}^{M+N} \mathcal{M}_k \le \mathsf{C}\sqrt{N}$.*

Proof. Fix $k = 2, \ldots, M + N$ and $a = (i,j) \in \square_{M,N}^{(k)}$, that is, $i + j = k$. Then

$$\mathbb{P}_{\sigma}\left(a \in \sigma\right) = \frac{\binom{i+j-2}{i-1}\binom{M+N-i-j}{M-i}}{\binom{M+N-2}{M-1}} = \frac{\binom{k-2}{i-1}\binom{M+N-k}{M-i}}{\binom{M+N-2}{M-1}}, \qquad i = 1, \ldots, k-1, \tag{3.3}$$

which is the hypergeometric distribution. Indeed, the numerator counts pairs of paths from $(1,1)$ to (i,j) and from (i,j) to (M,N), and the denominator counts all possible paths. Let us denote the probability (3.3) by $p_i^{(k)}$. By looking at ratios $p_i^{(k)}/p_{i+1}^{(k)}$ and comparing this to 1 one can readily see that the mode of the distribution (3.3) (that is the $m \in \mathbb{R}$ such that $p_i^{(k)}$ achieves its maximum for an integer i neighboring m) is

$$m = \frac{(k-1)M}{M+N}.$$

Thus, plugging $i = m$ into (3.3) leads to an upper bound, up to a constant, on $\mathcal{M}_k$. We now consider three ranges of k. First, if $k \le \sqrt{N}$ or $k \ge M + N - \sqrt{N}$, then the number of summands in (3.2) is of order $\sqrt{N}$, and we estimate each of them by 1. Next, let k be from $\sqrt{N}$ to ϵN, where $\epsilon > 0$ is fixed (in the corresponding interval close to $M+N$ the estimate will be similar). Then using Stirling's formula one can readily see that $p_m^{(k)} = O(1/\sqrt{k})$. Summing $O(1/\sqrt{k})$ over k from $\sqrt{N}$ to εN we get a term of order $\sqrt{N}$ as well. Finally, if k is from ϵN to $(1-\epsilon)(M+N)$, then Stirling's formula gives

$$p_m^{(k)} = \frac{O(1)}{\sqrt{N}} \frac{1}{\sqrt{\frac{k}{N}(1 + \xi - \frac{k}{N})}},$$

where $O(1)$ is independent of k. The sum of these expressions over our range of k is $\sqrt{N}$ times a Riemann sum of a convergent integral, and thus also has order $\sqrt{N}$. This completes the proof. □

Proof of Theorem 1.1. This theorem follows from Lemma 3.2 and Theorem 2.2 with $\lambda = \frac{1}{4}$, which leads to the moment condition $\mathsf{p} > \frac{6}{1-2\lambda} = 12$. □

Proof of Theorem 1.2. Let us now discuss the case when σ is chosen uniformly from all up-right paths within $\square_{M,N}$ that pass through the points $(\lfloor \xi_i N \rfloor, \lfloor \zeta_i N \rfloor)$, $i = 1, \ldots, \ell$, where $0 < \xi_1 < \ldots < \xi_\ell < \xi$ and $0 < \zeta_1 < \ldots < \zeta_\ell < 1$ are fixed. In this case the sum $\sum_{k=2}^{M+N} \mathcal{M}_k$ splits into $\ell + 1$ sums. Each of these sums has $O(N)$ terms and similarly to Lemma 3.2 one can show that each sum behaves as $O(\sqrt{N})$. Thus, Theorem 1.2 holds with the same moment condition $\mathsf{p} > 12$ as in Theorem 1.1. □

Remark 3.3. The statement of Theorem 1.2 continues to be valid if the number of points $\ell = \ell(N)$ through which the path σ must pass goes to infinity, say, as $\ell(N) = O(N^\alpha)$, $0 < \alpha < 1$. Indeed, assume in addition that the smaller rectangles as in Figure 2, (a) have asymptotically equivalent sides, and that the sides of all $O(N^\alpha)$ rectangles are also asymptotically equivalent. Then $\sum_{k=2}^{M+N} \mathcal{M}_k$ is bounded by $\mathsf{C} N^\alpha N^{\frac{1-\alpha}{2}}$, and so Theorem 1.2 holds with $\lambda = \frac{1+\alpha}{4}$, which leads to the moment condition $\mathsf{p} > \frac{12}{1-\alpha}$.

3.2 Paths around a hole – proof of Theorem 1.3

Define $\Sigma_{N,N}^{(k)} := \Sigma_{N,N} \cap \square_{N,N}^{(k)}$, where $\Sigma_{N,N}$ is the set of vertices with the hole removed, and $\square_{N,N}^{(k)}$ is given by (3.1). Our goal is to estimate $\mathcal{M}_k = \max_{a \in \Sigma_{N,N}^{(k)}} \mathbb{P}_\sigma(a \in \sigma)$ to argue similarly to Section 3.1.

Lemma 3.4. *In the regime $B = \lfloor \beta N \rfloor$, $\beta \in (0,1)$, there exists $\mathsf{C} > 0$ such that for all N large enough we have $\sum_{k=2}^{2N} \mathcal{M}_k \le \mathsf{C}\sqrt{N}$.*

Proof. Let us denote $A := (N - B)/2$, so $N = 2A + B$, see Figure 2, (b). For simpler notation we will omit integer parts as this does not affect our up-to-constant estimates. In particular, we can and will assume that $A = \lfloor \frac{1-\beta}{2} N \rfloor$.

By symmetry it suffices to assume that $j \ge i$ and $i + j \le N + 1$; the estimates for other $(i,j) \in \Sigma_{N,N}$ would be the same. We have

$$\mathbb{P}_\sigma((i,j) \in \sigma) = \frac{1}{Z_N} \sum_{y=1}^{A} \binom{k-2}{i-1} \binom{N+1-k}{y-i} \binom{N-1}{y-1}, \qquad (i,j) \in \Sigma_{N,N}^{(k)}, \tag{3.4}$$

where $Z_N = 2\sum_{y=1}^{A}\binom{N-1}{y-1}^2$ is the number of up-right paths in $\Sigma_{N,N}$ (the factor 2 comes from symmetry). Here $(y, N+1-y)$, $y = 1, \dots, A$, is the point where the up-right path intersects the line $i + j = N + 1$. In (3.4) we also used the convention that $\binom{N+1-k}{y-i} = 0$ for $y < i$ since the first coordinate increases along up-right paths.

Let us first maximize the quantity under the sum in (3.4) in $i = 1, \dots, k-1$ for fixed y. By considering the ratios of the terms with i and $i+1$ we see that the mode in i is at

$$m(y) = \frac{y(k-1)}{N+1}.$$

Therefore, an up-to-constant upper bound for $\mathcal{M}_k$ following from (3.4) is (for some $\mathsf{C}_1 > 0$)

$$\mathcal{M}_k \le \frac{\mathsf{C}_1}{Z_N}\sum_{y=1}^{A}\binom{k-2}{m(y)-1}\binom{N+1-k}{y-m(y)}\binom{N-1}{y-1}. \tag{3.5}$$

The sums over y in both (3.5) and Z_N are dominated by the behavior around $y = A$ because $\frac{A}{N} < \frac{1}{2}$. Indeed, this follows from standard large deviations type equivalences for the binomial coefficients:

$$\begin{aligned}
\binom{N-1}{y-1} &= O(N^{-\frac{1}{2}})\exp\left\{-N\Big(\big(1-\frac{y}{N}\big)\log\big(1-\frac{y}{N}\big) + \frac{y}{N}\log\big(\frac{y}{N}\big)\Big) + O(N^{-1})\right\},\\
\binom{k-2}{m(y)-1} &= O(N^{-\frac{1}{2}})\exp\left\{-k\Big(\big(1-\frac{y}{N}\big)\log\big(1-\frac{y}{N}\big) + \frac{y}{N}\log\big(\frac{y}{N}\big)\Big) + O(N^{-1})\right\},\\
\binom{N+1-k}{y-m(y)} &= O(N^{-\frac{1}{2}})\exp\left\{-(N-k)\Big(\big(1-\frac{y}{N}\big)\log\big(1-\frac{y}{N}\big) + \frac{y}{N}\log\big(\frac{y}{N}\big)\Big) + O(N^{-1})\right\},
\end{aligned}$$

and the fact that the function $\mathsf{y} \mapsto -(1-\mathsf{y})\log(1-\mathsf{y}) - \mathsf{y}\log\mathsf{y}$ is positive and strictly increasing for $\mathsf{y} \in (0, 1/2)$. In the above equivalences we assumed that y/N is bounded away from 0, and k/N is bounded away from 0 and 1. The behavior at the tails can be estimated in a similar way; cf. the proof of Lemma 3.2.

In the sums over $y = 1, \dots, A$, both in the numerator and the denominator in (3.5), there are $O(\sqrt{N})$ terms dominating the other terms. This implies that the right-hand side of (3.5) behaves as $C(k)N^{-\frac{1}{2}}$. It remains to see that the constant $C(k)$ coming from the numerator is summable over k from ϵN to $(1-\epsilon)N$. This constant can be computed using Stirling's approximation:

$$C(k) = \frac{O(1)}{\sqrt{\frac{k}{N}(1-\frac{k}{N})}},$$

where $O(1)$ is independent of k (but depends on β). The sum of these expressions over k is equal to N times the Riemann sum of a convergent integral. Therefore, the sum of the $\mathcal{M}_k$'s is bounded by $\sqrt{N}$, as desired. □

Theorem 1.3 now follows from Lemma 3.4 and Theorem 2.2 with $\lambda = \frac{1}{4}$, so the moment condition is $\mathsf{p} > 12$.

References

[1] T. Alberts, K. Khanin, and J. Quastel, *Intermediate disorder regime for 1+ 1 dimensional directed polymers*, Ann. Probab. **42** (2014), no. 3, 1212–1256, arXiv:1202.4398 [math.PR]. MR-3189070

[2] J. Baik, P. Deift, and T. Suidan, *Combinatorics and random matrix theory*, Graduate Studies in Mathematics, vol. 172, AMS, 2016. MR-3468920

[3] A. Borodin and I. Corwin, *Macdonald processes*, Probab. Theory Relat. Fields **158** (2014), 225–400, arXiv:1111.4408 [math.PR]. MR-3152785

[4] A. Borodin, I. Corwin, and P. Ferrari, *Free energy fluctuations for directed polymers in random media in 1+ 1 dimension*, Comm. Pure Appl. Math. **67** (2014), no. 7, 1129–1214, arXiv:1204.1024 [math.PR]. MR-3207195

[5] I. Corwin, *The Kardar-Parisi-Zhang equation and universality class*, Random Matrices Theory Appl. **1** (2012), arXiv:1106.1596 [math.PR]. MR-2930377

[6] C.-G. Esseen, *On mean central limit theorems*, Kungl. Tekn. Högsk. Handl. Stockholm **121** (1958), 31 pp. MR-0097111

[7] L. Goldstein, *l^1 bounds in normal approximation*, Ann. Probab. **35** (2007), no. 5, 1888–1930, arXiv:0710.3262 [math.PR]. MR-2349578

[8] A. Guionnet and O. Zeitouni, *Concentration of the spectral measure for large matrices*, Electron. Commun. Probab. **5** (2000), 119–136. MR-1781846

[9] D. Huse and C. Henley, *Pinning and roughening of domain wall in ising systems due to random impurities*, Phys. Rev. Lett **54** (1985), no. 25, 2708.

[10] J. Imbrie and T. Spencer, *Diffusion of directed polymers in a random environment*, J. Statist. Phys. **52** (1988), no. 3-4, 609–626. MR-0968950

[11] K. Johansson, *Shape fluctuations and random matrices*, Commun. Math. Phys. **209** (2000), no. 2, 437–476, arXiv:math/9903134 [math.CO]. MR-1737991

[12] M. Ledoux, *On talagrand's deviation inequalities for product measures*, ESAIM: Probability and statistics **1** (1997), 63–87. MR-1399224

[13] M. Meckes, *Some results on random circulant matrices*, High dimensional probability V: the Luminy volume, IMS, 2009, arXiv:0902.2472 [math.PR], pp. 213–223. MR-2797949

[14] N. O'Connell and J. Ortmann, *Tracy-Widom asymptotics for a random polymer model with gamma-distributed weights.*, Electron. J. Probab. **20** (2015), no. 25, 1–18, arXiv:1408.5326 [math.PR]. MR-3325095

[15] T. Seppäläinen, *Lecture notes on the corner growth model*, Unpublished lecture notes (2009).

[16] T. Seppäläinen, *Scaling for a one-dimensional directed polymer with boundary conditions*, Ann. Probab. **40(1)** (2012), 19–73, arXiv:0911.2446 [math.PR]. MR-2917766

[17] T. Seppäläinen, *Variational formulas, busemann functions, and fluctuation exponents for the corner growth model with exponential weights*, arXiv preprint (2017), arXiv:1709.05771 [math.PR].

[18] M. Talagrand, *Concentration of measure and isoperimetric inequalities in product spaces*, Publications Mathematiques de l'IHES **81** (1995), no. 1, 73–205, arXiv:math/9406212 [math.PR]. MR-1361756

[19] C. Tracy and H. Widom, *Level-spacing distributions and the Airy kernel*, Commun. Math. Phys. **159** (1994), no. 1, 151–174, arXiv:hep-th/9211141. MR-1257246

Acknowledgments. We appreciate helpful discussions with Alexander Drewitz, Elizaveta Rebrova, Timo Seppäläinen, and Roman Vershynin. Part of this work was completed when MM and LP participated at a workshop "Analytic Tools in Probability and Applications" at IMA in 2015, and we are grateful to the Institute and the workshop's organizers for hospitality and support. MM is partially supported by grant #315593 from the Simons Foundation. LP is partially supported by the NSF grant DMS-1664617.