The Annals of Probability
2019, Vol. 47, No. 6, 3866–3910
https://doi.org/10.1214/19-AOP1349

FOUR-DIMENSIONAL LOOP-ERASED RANDOM WALK

BY GREGORY LAWLER[1], XIN SUN[2] AND WEI WU[3]

University of Chicago, Columbia University and University of Warwick

The loop-erased random walk (LERW) in $\mathbb{Z}^4$ is the process obtained by erasing loops chronologically for a simple random walk. We prove that the escape probability of the LERW renormalized by $(\log n)^{\frac{1}{3}}$ converges almost surely and in L^p for all $p > 0$. Along the way, we extend previous results by the first author building on slowly recurrent sets. We provide two applications for the escape probability. We construct the two-sided LERW, and we construct a ± 1 spin model coupled with the wired spanning forests on $\mathbb{Z}^4$ with the bi-Laplacian Gaussian field on $\mathbb{R}^4$ as its scaling limit.

1. Introduction. Loop-erased random walk (LERW) is a probability measure on self-avoiding paths introduced by the first author of this paper in [4]. Since then, LERW has become an important model in statistical physics and probability, with close connections to other important subjects such as the uniform spanning tree and the Schramm–Loewner evolution. A key quantity that governs the large scale behavior of LERW is the so-called escape probability, namely, the nonintersection probability of a LERW and an independent simple random walk (SRW) starting at the same point. It is known that $d = 4$ is critical for LERW, in the sense that a LERW and an SRW on $\mathbb{Z}^d$ intersect a.s. if and only if $d \leq 4$. It was shown in [7] that LERW on $\mathbb{Z}^4$ has Brownian motion as its scaling limit after proper normalization. The exact normalization was conjectured but not proved in that paper; in [9], it was determined up to multiplicative constants. The argument uses a weak version of a "mean-field" property for LERW in $\mathbb{Z}^4$. In this paper, we establish the sharp mean-field property for the escape probability of LERW on $\mathbb{Z}^4$ that goes beyond the scaling limit result.

We state our main results for the renormalized escape probability of 4D LERW, Theorems 1.1 and 1.2, in Section 1.1. An outline of the proofs is given in Section 1.2. Then in Sections 1.3 and 1.4 we discuss two applications of the main results, namely a construction of the two-sided LERW in $d = 4$, and a spin field

Received April 2017; revised February 2019.

[1]Supported by NSF Grant DMS-1513036.

[2]Supported by the Simons Foundation as a Junior Fellow at Simons Society of Fellows and by NSF Grants DMS-1811092 and by the Minerva fund at the Department of Mathematics at Columbia University.

[3]Supported in part by NSF Grant DMS-1507019.

MSC2010 subject classifications. 60G50, 60K35.

Key words and phrases. Loop erased random walk, escape probability, uniform spanning forest.

coupled with the wired spanning forests on $\mathbb{Z}^4$ with the bi-Laplacian Gaussian field on $\mathbb{R}^4$ as its scaling limit.

1.1. *Escape probability of LERW.* Given a positive integer d, a process $S = \{S_n\}_{n\in\mathbb{N}}$ on $\mathbb{Z}^d$ is called a *simple random walk* (SRW) on $\mathbb{Z}^d$ if $\{S_{n+1} - S_n\}_{n\in\mathbb{N}}$ are i.i.d. random variables taking uniform distribution on $\{z \in \mathbb{Z}^d : |z| = 1\}$. Here $|\cdot|$ is the Euclidean norm on $\mathbb{R}^d$. Unless otherwise stated, our SRW starts at the origin, namely, $S_0 = 0$. When $S_0 = x$ almost surely, we denote the probability measure of S by $\mathbb{P}^x$.

A *path* on $\mathbb{Z}^d$ is a sequence of vertices such that any two consecutive vertices are neighbors in $\mathbb{Z}^d$. Given a sample S of SRW and $m < n \in \mathbb{N}$, let $S[m, n]$ and $S[n, \infty)$ be the paths $[S_m, S_{m+1} \cdots, S_n]$ and $[S_n, S_{n+1}, \ldots]$, respectively. Given a finite path $\mathcal{P} = [v_0, v_1, \ldots, v_n]$ on $\mathbb{Z}^d$, the *(forward) loop erasure* of $\mathcal{P}$ (denoted by $\mathbf{LE}(\mathcal{P})$) is defined by erasing cycles in $\mathcal{P}$ chronologically. More precisely, we define $\mathbf{LE}(\mathcal{P})$ inductively as follows. The first vertex u_0 of $\mathbf{LE}(\mathcal{P})$ is the vertex v_0 of $\mathcal{P}$. Supposing that u_j has been set, let k be the last index such that $v_k = u_j$. Set $u_{j+1} = v_{k+1}$ if $k < n$; otherwise, let $\mathbf{LE}(\mathcal{P}) := [u_0, \ldots, u_j]$. Suppose S is an SRW on $\mathbb{Z}^d$ ($d \geq 3$). Since S is transient, there is no trouble defining $\mathbf{LE}(S) = \mathbf{LE}(S[0, \infty))$, which we call the *loop-erased random walk* (LERW) on $\mathbb{Z}^d$. LERW on $\mathbb{Z}^2$ can be defined via a limiting procedure but we will not discuss it in this paper.

Let W and S be two independent simple random walks on $\mathbb{Z}^4$ starting at the origin and $\eta = \mathbf{LE}(S)$. Let

$$X_n = (\log n)^{\frac{1}{3}} \mathbb{P}\{W[1, n^2] \cap \eta = \varnothing \mid \eta\}.$$

In [9], building on the work on slowly recurrent sets [8], the first author of this paper proved that $\mathbb{E}[X_n^p] \asymp 1$ for all $p > 0$. In this paper, we show the following.

THEOREM 1.1. *There exists a nontrivial random variable X_∞ such that*

$$\lim_{n\to\infty} X_n = X_\infty \quad \textit{almost surely and in } L^p \textit{ for all } p > 0.$$

We can view X_∞ as the renormalized escape probability of 4D LERW at its starting point. It is the key for our construction of the 4D two-sided LERW in Section 1.3. Our next theorem is similar to Theorem 1.1 with the additional feature of the evaluation of the limiting constant.

THEOREM 1.2. *Let W, W', W'', S be four independent simple random walks on $\mathbb{Z}^4$ starting from the origin and $\eta = \mathbf{LE}(S)$. Then*

$$\lim_{n\to\infty} (\log n)\mathbb{P}\{(W[1, n^2] \cup W'[1, n^2]) \cap \eta = \varnothing, W''[0, n^2] \cap \eta = \{0\}\} = \frac{\pi^2}{24}.$$

Write $\frac{\pi^2}{24}$ in Theorem 1.2 as $\frac{1}{3} \cdot \frac{\pi^2}{8}$. We will see that the constant $\frac{1}{3}$ is universal and is the reciprocal of the number of SRWs other than S. The factor $\pi^2/8$ comes from the bi-harmonic Green function of $\mathbb{Z}^4$ evaluated at (0, 0) and is lattice-dependent. The SRW analog of Theorem 1.2 is proved in [10], Corollary 4.2.5:

$$\lim_{n\to\infty}(\log n)\mathbb{P}\{W[1,n^2]\cap S[0,n^2]=\varnothing, W'[0,n^2]\cap S[1,n^2]=\varnothing\} = \frac{1}{2}\cdot\frac{\pi^2}{8}.$$

Theorems 1.1 and 1.2 are a special case of our Theorem 1.5, whose proof is outlined in Section 1.2. In particular, the asymptotic result is obtained from a refined analysis of slowly recurrent set beyond [8, 9] as well as fine estimates on the harmonic measure of 4D LERW. The explicit constant $\frac{\pi^2}{24}$ is obtained from a "first passage" path decomposition of the intersection of an SRW and a LERW. Here, care is needed because there are several time scales involved. See Section 1.2 for an outline. As a byproduct, at the end of Section 5.2 we obtain an asymptotic result on the long range intersection between SRW and LERW which is of independent interest.

To state the result, we recall the Green function on $\mathbb{Z}^4$ defined by

$$G(x, y) = \sum_{n=0}^{\infty} \mathbb{P}^x[S_n = y].$$

Given a subset $A \subset \mathbb{Z}^4$, the Green function on A is defined by

$$G_A(x, y) = \sum_{n=0}^{\infty} \mathbb{P}^x\big[S_n = y, S[0,n] \subset A\big].$$

It will be technically easier to work on geometric scales. Let $C_n = \{z \in \mathbb{Z}^4 : |z| < e^n\}$ be the discrete disk, $G_n = G_{C_n}$ and

$$G_n^2(w) = \sum_{z\in C_n} G_n(0,z)G_n(z,w).$$

THEOREM 1.3. *Let W, S be independent simple random walks on $\mathbb{Z}^4$ with $W_0 = 0$ and $S_0 = w$. Let $\sigma_n^W = \min\{j : W_j \notin C_n\}$ and $\sigma_n = \min\{j : S_j \notin C_n\}$. If*

$$q_n(w) = \mathbb{P}\{W[0,\sigma_n^W] \cap \mathbf{LE}(S[0,\sigma_n]) \neq \varnothing\},$$

then

$$\lim_{n\to\infty} \max_{n^{-1}\le e^{-n}|w|\le 1-n^{-1}} \left| n q_n(w) - \frac{\pi^2}{24} G_n^2(w)\right| = 0.$$

REMARK 1.4. Theorem 1.3 holds if $W[0,\sigma_n^W] \cap \mathbf{LE}(S[0,\sigma_n])$ is replaced by $W[0,\sigma_n^W] \cap S[0,\sigma_n]$ and $\pi^2/24$ is replaced by $\pi^2/16$. This is the long range estimate for two independent SRWs in [10], Section 4.3. The function $G_n^2(w)$ is the expected number of intersections of $S[0, \sigma_n]$ and $W[0, \sigma_n^W]$. This means that the long-range nonintersection probability of an SRW and an independent LERW is comparable with that of two independent SRWs. This is closely related to the fact that the scaling limit of LERW on $\mathbb{Z}^4$ is Brownian motion, that is, has Gaussian limits.

1.2. *Outline of the proof.* In this subsection, we will first state Theorem 1.5, from which Theorems 1.1 and 1.2 are immediate corollaries. Then we give an outline of its proof, leaving the details to Sections 2–5.

We start by defining some notation. Let $\sigma_n = \min\{j \geq 0 : S_j \notin C_n\}$ and

$$\mathcal{F}_n \text{ be the } \sigma\text{-algebra generated by } \{S_j : j \leq \sigma_n\}. \tag{1.1}$$

We recall that there exist $0 < \beta, c < \infty$ such that for all n, if $z \in C_{n-1}$ and $a \geq 1$,

$$\mathbb{P}^z\{a^{-1}e^{2n} \leq \sigma_n \leq ae^{2n}\} \geq 1 - ce^{-\beta a}. \tag{1.2}$$

For the lower inequality (see, e.g., [12], (12.12)) and the upper inequality follows from the fact that $\mathbb{P}^z\{\sigma_n \leq (k+1)e^{2n} \mid \sigma_n \geq ke^{2n}\}$ is uniformly bounded away from 0.

If $x \in \mathbb{Z}^4$, $V \subset \mathbb{Z}^4$, we write

$$\begin{aligned} H(x, V) &= \mathbb{P}^x\{S[0, \infty) \cap V \neq \varnothing\}, \\ H(V) &= H(0, V), \qquad \mathrm{Es}(V) = 1 - H(V), \\ \overline{H}(x, V) &= \mathbb{P}^x\{S[1, \infty) \cap V \neq \varnothing\}, \qquad \overline{\mathrm{Es}}(V) = 1 - \overline{H}(0, V). \end{aligned}$$

Note that $\overline{\mathrm{Es}}(V) = \mathrm{Es}(V)$, if $0 \notin V$. If $0 \in V$, a standard last-exit decomposition shows that

$$\mathrm{Es}(V^0) = G_{\mathbb{Z}^4 \setminus V^0}(0, 0)\overline{\mathrm{Es}}(V), \tag{1.3}$$

where $V^0 = V \setminus \{0\}$ and $G_{\mathbb{Z}^4 \setminus V^0}$ is the Green's function on $\mathbb{Z}^4 \setminus V^0$. We also write

$$\mathrm{Es}(V; n) = \mathbb{P}\{S[1, \sigma_n] \cap V = \varnothing\},$$

which is clearly decreasing in n.

We have to be a little careful about the definition of the loop-erasures of the random walk and loop-erasures of subpaths of the walk. We will use the following notation:

- η denotes the (forward) loop-erasure of $S[0, \infty)$ and
$$\Gamma = \eta[1, \infty) = \eta[0, \infty) \setminus \{0\}.$$
- ω_n denotes the finite random walk path $S[\sigma_{n-1}, \sigma_n]$.
- $\eta^n = \mathbf{LE}(\omega_n)$ denotes the loop-erasure of $S[\sigma_{n-1}, \sigma_n]$.
- $\Gamma_n = \mathbf{LE}(S[0, \sigma_n]) \setminus \{0\}$, that is, Γ_n is the loop-erasure of $S[0, \sigma_n]$ with the origin removed.

Note that $S[1,\infty)$ is the concatenation of the paths $\omega_1, \omega_2, \ldots$. However, it is not true that Γ is the concatenation of $\eta^1, \eta^2, \ldots$, and that is one of the technical issues that must be addressed in the proof.

Let Y_n, Z_n, G_n be the $\mathcal{F}_n$-measurable random variables

$$Y_n = H(\eta^n), \qquad Z_n = \mathrm{Es}[\Gamma_n], \qquad G_n = G_{\mathbb{Z}^4 \setminus \Gamma_n}(0,0). \tag{1.4}$$

By (1.3), we have $\overline{\mathrm{Es}}(\Gamma_n \cup \{0\}) = G_n^{-1} Z_n$. It is easy to see that $1 \le G_n \le 8$. Furthermore, using the transience of S, we can see that with probability one $G_\infty := \lim_{n\to\infty} G_n$ exists and equals $G_{\mathbb{Z}^4 \setminus \Gamma}(0,0)$.

THEOREM 1.5. *For every* $0 \le r, s < \infty$, *there exists* $0 < c_{r,s} < \infty$, *such that*

$$\lim_{n\to\infty} n^{r/3} \mathbb{E}[Z_n^r G_n^{-s}] = c_{r,s}.$$

Moreover, $c_{3,2} = \pi^2/24$.

Our methods do not compute the constant $c_{r,s}$ except in the case $r = 3$, $s = 2$ (and the trivial case $r = s = 0$).

The proof of Theorem 1.5, which is the technical bulk of this paper, requires several steps which we will outline now. For the remainder of this paper, we fix $r > 0$ and allow constants to depend on r. If $n \in \mathbb{N}$, we let

$$p_n = \mathbb{E}[Z_n^r], \qquad \hat{p}_n = \mathbb{E}[Z_n^3 G_n^{-2}], \tag{1.5}$$

$$h_n = \mathbb{E}[Y_n] = \mathbb{E}[H(\eta^n)], \qquad \phi_n = \prod_{j=1}^{n} e^{-h_j}. \tag{1.6}$$

In Section 2.2, we review and prove some basic estimates on simple random walk that, in particular, give $h_n = O(n^{-1})$, and hence,

$$\phi_n = \phi_{n-1} e^{-h_n} = \phi_{n-1}[1 + O(n^{-1})].$$

In Section 3.1, we revisit the theory of slowly recurrent sets in [8, 9] and obtain quantitative estimates on the escape probability of slowly recurrent sets under a mild assumption (see Definition 3.1). Using these estimates, we prove two propositions in Section 3.2. The first one controls p_n/p_{n+1}.

PROPOSITION 1.6. $p_{n+1} = p_n[1 - O(\log^4 n/n)]$.

The second one gives a good estimate along the subsequence $\{n^4\}$. Let $\tilde{\eta}_n$ denote the (forward) loop-erasure of $S[\sigma_{(n-1)^4+(n-1)}, \sigma_{n^4-n}]$. For $m < n$, we let $A(m,n)$ be the discrete annuli defined by

$$A(m,n) = C_n \setminus C_m = \{z \in C_n : |z| \ge e^m\}.$$

Let $\tilde{\Gamma}_n = \tilde{\eta}_n \cap A((n-1)^4 + 4(n-1), n^4 - 4n)$ and $\tilde{h}_n = \mathbb{E}[H(\tilde{\Gamma}_n)]$.

PROPOSITION 1.7. *There exists $c_0 < \infty$ such that*

$$p_{n^4} = [c_0 + O(n^{-1})] \exp\left\{-r \sum_{j=1}^{n} \tilde{h}_j\right\}.$$

In Section 4, in order to get rid of the subsequence $\{n^4\}$, we prove the following.

PROPOSITION 1.8. *There exists $c < \infty$, $u > 0$ such that*

$$\left|\tilde{h}_n - \sum_{(n-1)^4 < j \le n^4} h_j\right| \le \frac{c}{n^{1+u}}.$$

Proposition 1.8 intuitively says that if the random walk hits $\tilde{\Gamma}_n$ then it does so by hitting exactly one of η^j's. This proposition, which is key for proving our main result, does not follow from the work in [9]. Some of the earlier propositions have been improved here in order to be able to establish this. To rigorously prove this, we need a frequency estimate on cut points of SRW and a large deviation estimate on the harmonic measure of the range of SRW obtained in Sections 2.3 and 4.1, respectively. Propositions 1.6–1.8 readily yield the following.

PROPOSITION 1.9. *For every r, s, there exists constant $c'_{r,s}$, $u > 0$ such that*

$$\mathbb{E}[Z_n^r G_n^{-s}] = c'_{r,s} \phi_n^r [1 + O(n^{-u})].$$

In particular, there exists a constant $c'_{3,2} > 0$ such that

$$\hat{p}_n = c'_{3,2}[1 + O(n^{-u})] \exp\left\{-3 \sum_{j=1}^{n} h_j\right\}. \tag{1.7}$$

In Section 5, we use a path decomposition to study the long-range intersection of an SRW and a LERW and show in Proposition 5.2 that there exists $u > 0$ such that

$$h_n = \frac{8}{\pi^2} \hat{p}_n + O(n^{-1-u}). \tag{1.8}$$

Combined with (1.7), this gives that the limit

$$\lim_{n\to\infty} \left[\log \hat{p}_n + \frac{24}{\pi^2} \sum_{j=1}^{n} \hat{p}_j\right] \tag{1.9}$$

exists and is finite. Note that $\lim_{n\to\infty} \hat{p}_{n+1}/\hat{p}_n = 1$ (see Proposition 1.9). In Section 5.1, we prove an elementary lemma (see Lemma 5.1) on sequences asserting that this combined with (1.9) assures that $\lim_{n\to\infty} n\hat{p}_n = \pi^2/24$. Now (1.7) and (1.8) imply that $\lim_{n\to\infty} 3nh_n = 1$ and

$$\phi_n^3 = \exp\left\{-3\sum_{j=1}^{n} h_j\right\} \sim \frac{c}{n} \quad \text{for some constant } c > 0. \tag{1.10}$$

This combined with Proposition 1.9 concludes the proof of Theorem 1.5. This already implies Theorem 1.2 by changing scales. The proof of Theorem 1.1 will be explained in Section 6.

1.3. *Two sided LERW.* In [5], the first author author proved the existence of *two-sided loop-erased random walk* in $\mathbb{Z}^d$ for $d \geq 5$.

THEOREM 1.10 ([5]). *Given $d \geq 5$, consider the sample of LERW in $\mathbb{Z}^d$, denoted by $\{\eta_i\}_{i\geq 0}$. The $n \to \infty$ limit of $\{\eta_{n+i} - \eta_n\}_{-k\leq i\leq k}$ exists for any $k \in \mathbb{N}$, which defines an ergodic random path $\{\tilde{\eta}_i\}_{i\in\mathbb{Z}}$ in $\mathbb{Z}^d$ called the two-sided LERW.*

The proof of Theorem 1.10 crucially replies on the existence of global cut points for SRW in $\mathbb{Z}^d$ for $d \geq 5$, which is not true for $d \leq 4$. As an application of results in Section 1.1, we extend the existence of the two-sided LERW to $d = 4$ in Section 6. Moreover, X_∞ in Theorem 1.1 is the Radon–Nikodym derivative between the two-sided LERW restricted to nonnegative times and the usual LERW. The existence for $d = 2, 3$ was recently established by the first author author in [11]. A big difference in $d < 4$ compared to $d \geq 4$ case is that the marginal distribution of one side of the path is *not* absolutely continuous with respect to the usual LERW.

Our results addresses the $d = 4$ case of Conjecture 15.12 in [2] by Benjamini–Lyons–Peres–Schramm, which asserts the existence of the two-sided uniform spanning tree in $\mathbb{Z}^d$. This is immediate from Wilson's algorithm [16] that connects LERW and uniform spanning tree (see Section 7.1).

1.4. *A spin field from USF.* As an application of Theorem 1.3, we will construct a sequence of random fields on the integer lattice $\mathbb{Z}^d$ ($d \geq 4$) using uniform spanning tree and show that they converge in distribution to the bi-Laplacian field (Theorem 1.11).

For each positive integer n, let $N = N_n = n(\log n)^{1/4}$. Let $A_N = \{x \in \mathbb{Z}^d : |x| < N\}$. We will construct a ± 1 valued random field on A_N as follows. Recall that a *wired spanning tree* on A_N is a tree on the graph $A_N \cup \{\partial A_N\}$ where we have viewed the boundary ∂A_N as "wired" to a single point. Such a tree produces a *spanning forest* on A_N by removing the edges connected to ∂A_N. We define the uniform spanning forest (USF) on A_N to be the forest obtained by choosing the wired spanning tree of $A_N \cup \{\partial A_N\}$ from the uniform distribution. (Note this is not the same thing as choosing a spanning forest uniformly among all spanning forests of A_N.) We now define the random field on (a rescaling of) $\mathbb{Z}^d$. Let a_n be a sequence of positive numbers (we will be more precise later).

- Choose a USF on A_N. This partitions A_N into (connected) components.
- For each component of the forest, flip a fair coin and assign each vertex in the component value 1 or -1 based on the outcome. This gives a field of spins $\{Y_{x,n} : x \in A_N\}$. If we wish, we can extend this to a field on $x \in \mathbb{Z}^d$ by setting $Y_{x,n} = 0$ for $x \notin A_N$.
- Let $\phi_n(x) = a_n Y_{nx,n}$ which is a field defined on $L_n := n^{-1}\mathbb{Z}^d$.

This random function is constructed in a manner similar to the Edward–Sokal coupling of the FK-Ising model [3]. That coupling says that we can obtain the Ising model on $\mathbb{Z}^d$ by first sample a random configuration $\omega \in \{0, 1\}^{\mathbb{Z}^d}$ according to the so-called random cluster measure, and then flip a fair coin and assign each component of ω value 1 or -1 based on the outcome. The way we construct ϕ_n is similar to the Ising model except that we replace the random cluster measure by the USF measure on $\mathbb{Z}^d$.

It is known that the Ising model has critical dimension $d = 4$, in the sense that mean field critical behaviors are expected for $d \ge 4$ but not for $d \le 3$. In particular, it is believed when $d \ge 4$ the scaling limit of Ising model is a d-dimensional Gaussian Free Field (GFF). For $d \ge 5$, this GFF limit is proved by Aizenman [1], while the critical case $d = 4$ is still open. Our theorem below asserts that the random field ϕ_n we construct has critical dimension $d = 4$, and for $d \ge 4$, we can choose the scaling a_n such that ϕ_n converges to the bi-Laplacian Gaussian field on $\mathbb{R}^d$. Note that when $d = 4$, a bi-Laplacian Gaussian field is log-correlated.

If $h \in C_0^\infty(\mathbb{R}^d)$, we write

$$\langle h, \phi_n \rangle = n^{-d/2} \sum_{x \in L_n} h(x)\phi_n(x).$$

THEOREM 1.11.

- *If $d \ge 5$, there exists $a > 0$ such that if $a_n = an^{(d-4)/2}$, then for every $h_1, \dots, h_m \in C_0^\infty(\mathbb{R}^d)$, the random variables $\langle h_j, \phi_n \rangle$ converge in distribution to a centered joint Gaussian random variable with covariance*

$$\iint h_j(z)h_k(w)|z-w|^{4-d}\, dz\, dw.$$

- *If $d = 4$, if $a_n = \sqrt{3 \log n}$, then for every $h_1, \dots, h_m \in C_0^\infty(\mathbb{R}^d)$ with*

$$\int h_j(z)\, dz = 0, \quad j = 1, \dots, m,$$

the random variables $\langle h_j, \phi_n \rangle$ converge in distribution to a centered Gaussian random variable with variance

$$-\iint h_j(z)h_k(w) \log|z-w|\, dz\, dw.$$

REMARK 1.12.

- Gaussian fields on $\mathbb{R}^d$ with correlations as in Theorem 1.11 is called d-dimensional bi-Laplacian Gaussian field (see [13]).
- For $d = 4$, we could choose the cutoff $N = n(\log n)^\alpha$ for any $\alpha > 0$. We choose $\alpha = \frac{1}{4}$ for concreteness. For $d > 4$, we could do the same construction with no cutoff ($N = \infty$) and get the same result.

By Wilson's algorithm, the two-point correlation function of the field ϕ_n is proportional to the intersection probability of an SRW and a LERW stopped upon hitting ∂A_N. Therefore, Theorem 1.11 essentially follows from Theorem 1.3. In particular, $G_n^2(w)$ there is the discrete biharmonic function that gives the covariance structure of the bi-Laplacian random field in the scaling limit. We will give the full proof of Theorem 1.11 in Section 7, where we only deal with the case $d = 4$. The $d \geq 5$ case can be proved in the same way but is much easier (see [15] for a detailed argument).

2. Preliminaries.

In this section, we recall and prove necessary lemmas about SRW and LERW which will be used frequently in the rest of the paper. Throughout this section, we retain the notation in Section 1.2.

2.1. *Basic notation.*

Given a vertex set $V \subset \mathbb{Z}^d$, ∂V is the set of vertices on $\mathbb{Z}^d \setminus V$ who have a neighbor in V, and $\overline{V} = V \cup \partial V$. A function ϕ on $\overline{V}$ is called *harmonic* on V if for each $v \in V$; we have $\mathbb{E}^v[\phi(S_1)] = \phi(v)$. When we say "$\phi$ is harmonic on V," then it is implicit that ϕ is defined on $\overline{V}$.

We will use c and C to represent constants which may vary line by line. We use the asymptotic notion that two nonnegative functions $f(x)$, $g(x)$ satisfy $f \lesssim g$ if there exists a constant $C > 0$ independent of x such that $f(x) \leq Cg(x)$. We write $f \gtrsim g$ if $g \lesssim f$ and write $f \asymp g$ if $f \lesssim g$ and $g \gtrsim f$. Given a sequence $\{a_n\}$ and a nonnegative sequence $\{b_n\}$, we write $a_n \sim b_n$ if $\lim_{n\to\infty} a_n/b_n = 1$. We write $a_n = O(b_n)$ if $|a_n| \lesssim b_n$. We write $a_n = o(b_n)$ if $\lim_{n\to\infty} |a_n|/b_n = 0$. When $\{b_n\} = \{1\}$, we may write $o(1)$ as $o_n(1)$ to indicate the dependence on n.

We say that a sequence $\{\epsilon_n\}$ of positive numbers is *fast decaying* if it decays faster than every power of n, that is, $n^k \epsilon_n = o_n(1)$ for every $k > 0$. We will write $\{\epsilon_n\}$ for fast decaying sequences. As is the convention for constants, the exact value of $\{\epsilon_n\}$ may change from line to line. We will use implicitly the fact that if $\{\epsilon_n\}$ is fast decaying then so is $\{\epsilon'_n\}$ where $\epsilon'_n = \sum_{m \geq n} \epsilon_m$.

2.2. *Estimates for simple random walk on $\mathbb{Z}^4$.*

In this subsection, we provide facts about simple random walk in $\mathbb{Z}^4$, which will be frequently used in the rest of the paper. We first recall the following facts about intersections of random walks in $\mathbb{Z}^4$ (see [6, 8]).

PROPOSITION 2.1. *There exist* $0 < c_1 < c_2 < \infty$ *such that the following is true. Suppose* S *is a simple random walk on* $\mathbb{Z}^4$ *starting at* 0. *Then*

$$\frac{c_1}{\sqrt{\log n}} \le \mathbb{P}\{S[0,n] \cap S[n+1,\infty] = \varnothing\}$$
$$\le \mathbb{P}\{S[0,n] \cap S[n+1,2n] = \varnothing\} \le \frac{c_2}{\sqrt{\log n}},$$

and if $2 \le \alpha \le n$,

$$c_1 \frac{\log \alpha}{\log n} \le \mathbb{P}\{S[0,n] \cap S[n(1+\alpha^{-1}),\infty) \ne \varnothing\} \le c_2 \frac{\log \alpha}{\log n}. \tag{2.1}$$

Moreover, if S^1 *is an independent simple random walk starting at* $z \in \mathbb{Z}^4$,

$$\mathbb{P}\{S[0,n] \cap S^1[0,\infty) \ne \varnothing\} \le c_2 \frac{\log a}{\log n}, \tag{2.2}$$

where $a = \max\{2, \sqrt{n}/|z|\}$.

An important corollary of Proposition 2.1 is that

$$\sup_n n\mathbb{E}[H(\eta^n)] \le \sup_n n\mathbb{E}[H(\omega_n)] < \infty, \tag{2.3}$$

and hence

$$\exp\{-\mathbb{E}[H(\eta^n)]\} = 1 - \mathbb{E}[H(\eta^n)] + O(n^{-2}).$$

It follows that if ϕ_n is defined as in (1.6) and $m < n$, then

$$\phi_n = \phi_m[1 + O(m^{-1})] \prod_{j=m+1}^{n} [1 - \mathbb{E}[H(\eta^j)]]. \tag{2.4}$$

COROLLARY 2.2. *There exists* $c < \infty$ *such that if* $n \in \mathbb{N}$, $\alpha \ge 2$, $0 < u < 1$, $m = m_{n,\alpha} = (1 + \alpha^{-1})n$ *and* $Y_{n,\alpha} = \max_{j \ge m} H(S_j, S[0,n])$, *then*

$$\mathbb{P}\left\{Y_{n,\alpha} \ge \frac{\log \alpha}{(\log n)^u}\right\} \le \frac{c}{(\log n)^{1-u}}.$$

PROOF. We fix n, α, u and define the stopping time τ by

$$\tau = \min\left\{j \ge m : H(S_j, S[0,n]) \ge \frac{\log \alpha}{(\log n)^u}\right\}.$$

The strong Markov property of S and the definition of τ imply that

$$\begin{aligned}
&\mathbb{P}\{S[0,n] \cap S[m,\infty) \ne \varnothing \mid \tau < \infty\} \\
&\quad \ge \mathbb{P}\{S[0,n] \cap S[\tau,\infty) \ne \varnothing \mid \tau < \infty\} \\
&\quad = \mathbb{E}[\mathbb{P}\{S[0,n] \cap S[\tau,\infty) \ne \varnothing \mid S(\tau), \tau < \infty\}] \\
&\quad = \mathbb{E}[H(S(\tau), S[0,n]) \mid S(\tau), \tau < \infty] \ge \frac{\log \alpha}{(\log n)^u}.
\end{aligned}$$

Therefore, using (2.1),

$$\mathbb{P}\{\tau < \infty\} \le \frac{(\log n)^u}{\log \alpha} \mathbb{P}\{S[0,n] \cap S[m,\infty) \neq \varnothing\} \le \frac{c}{(\log n)^{1-u}}. \qquad \square$$

LEMMA 2.3. *There exists $c > 0$ such that the following holds for all $n \in \mathbb{N}$:*

- *If ϕ is a positive (discrete) harmonic function on C_n and $x \in C_{n-1}$,*

$$\left|\log[\phi(x)/\phi(0)]\right| \le c|x|e^{-n}. \tag{2.5}$$

- *If $m < n$, $V \subset C_{n-m}$, and $\tilde{V} = V \setminus C_{n-m-1}$,*

$$c^{-1}e^{-2m} \le \mathbb{P}\{S[\sigma_n,\infty) \cap C_{n-m} \neq \varnothing \mid \mathcal{F}_n\} \le ce^{-2m}; \tag{2.6}$$

$$c^{-1}e^{-2m}H(\tilde{V}) \le \mathbb{P}\{S[\sigma_n,\infty) \cap V \neq \varnothing \mid \mathcal{F}_n\} \le ce^{-2m}H(V). \tag{2.7}$$

$$\mathrm{Es}(V) \ge \mathbb{P}\{S[0,\sigma_n] \cap V = \varnothing\}\left[1 - ce^{-2m}H(V)\right]. \tag{2.8}$$

PROOF. The inequalities (2.5) and (2.6) are standard estimates; see, for example, [12], Theorem 6.3.8, Proposition 6.4.2. The Harnack principle [12], Theorem 6.3.9, shows that $H(z,V) \asymp H(z',V)$ for $z, z' \in \partial C_{n-m+1}$, and by stopping at time σ_{n-m+1} we see that

$$H(V) \ge \min_{z \in \partial C_{n-m+1}} H(z, \bar{V}).$$

This combined with (2.6) and the strong Markov property gives the upper bound in (2.7). To get the lower bound, one uses the Harnack principle to see that for $z \in \partial C_{n-m+1}$, $H(z, V^+) \asymp H(V^+)$ and $H(z, V \setminus V^+) \asymp H(V \setminus V^+)$, where

$$V^+ = \tilde{V} \cap \{(z_1,\dots,z_4) \in \mathbb{Z}^4 : z_1 \ge 0\}.$$

Finally, (2.8) follows from the upper bound in (2.7) and the strong Markov property. $\square$

LEMMA 2.4. *Let U_n be the event that there exists $k \ge \sigma_n$ with*

$$\mathbf{LE}(S[0,k]) \cap C_{n-\log^2 n} \neq \mathbf{LE}(S[0,\infty)) \cap C_{n-\log^2 n}.$$

Then $\mathbb{P}(U_n)$ is fast decaying.

PROOF. By the loop-erasing process, we can see that the event U_n is contained in the event that either

$$S[\sigma_{n-\frac{1}{2}\log^2 n},\infty) \cap C_{n-\log^2 n} \neq \varnothing \quad \text{or} \quad S[\sigma_n,\infty) \cap C_{n-\frac{1}{2}\log^2 n} \neq \varnothing.$$

The probability that either of these happens is fast decaying by (2.6). $\square$

The next proposition gives a quantitative estimate on the slowly recurrent nature of a simple random path in $\mathbb{Z}^4$.

PROPOSITION 2.5. *If $\Lambda(m,n) = S[0,\infty) \cap A(m,n)$, then the sequences*

$$\mathbb{P}\left\{H[\Lambda(n-1,n)] \geq \frac{\log^2 n}{n}\right\} \quad \textit{and} \quad \mathbb{P}\left\{H(\omega_n) \geq \frac{\log^4 n}{n}\right\}$$

are fast decaying.

PROOF. For any $z \in \mathbb{Z}^4$, let S^z be a simple random walk starting from z independent of S. Let $\Lambda_j^z = \Lambda_j^z(n-1,n) = S^z[0,j] \cap A(n-1,n)$ for $j \in \mathbb{N} \cup \{\infty\}$. By the definition of H and Proposition 2.1, there exists a positive constant c such that for each z with $|z| \geq e^{n-1}$,

$$\mathbb{E}[H(\Lambda_\infty^z)] = \mathbb{P}\{S[0,\infty) \cap S^z[0,\infty) \cap A(n-1,n) \neq \varnothing\} \leq \frac{c}{4n}.$$

From now on, we assume $|z| \geq e^{n-1}$. Then by the Markov inequality,

$$\mathbb{P}\left\{H(\Lambda_\infty^z) \geq \frac{c}{2n}\right\} \leq \frac{1}{2}. \tag{2.9}$$

For each $k \in \mathbb{N}$, let $\tau_k = \inf\{j : H(\Lambda_j^z) \geq ck/n\}$. On the event $\tau_k < \infty$, we have $H(\Lambda_{\tau_k - 1}^z) < ck/n$ and $S_{\tau_k}^z \in A(n-1,n)$. Since

$$H(A \cup B) \leq H(A) + H(B) \quad \text{for any } A, B \subset \mathbb{Z}^4,$$

for n sufficiently large, we have

$$H(\Lambda_{\tau_k}) \leq H(\Lambda_{\tau_k - 1}^z) + H(S_{\tau_k}^z) \leq c\left(k + \frac{1}{2}\right)/n.$$

Moreover, combined with (2.9) we see that for n sufficiently large,

$$\begin{aligned}\mathbb{P}\{\tau_{k+1} < \infty \mid \tau_k < \infty\} &\leq \sum_{w \in A(n-1,n)} \mathbb{P}[S_{\tau_k}^z = w \mid \tau_k < \infty]\mathbb{P}\left[H(\Lambda_\infty^w) \geq \frac{c}{2n}\right] \\ &\leq \frac{1}{2} \sum_{w \in A(n-1,n)} \mathbb{P}[S_{\tau_k}^z = w \mid \tau_k < \infty] = \frac{1}{2}.\end{aligned}$$

Therefore, $\mathbb{P}\{\tau_k < \infty\} \leq 2^{-k}$. Setting $k = \lfloor c^{-1} \log^2 n \rfloor$, we see that the first sequence in Proposition 2.5 is fast decaying.

For the second sequence, note that on the event $\{H(\omega_n) \geq \log^4 n/n\}$, either $\omega_n \not\subset A(n - \log^2 n, n)$ or there exists a $j \in [n - \log^2 n, n]$ such that $H[\Lambda(j-1, j)] \geq \log^2 n/n$. We use (2.6) to see that $\mathbb{P}\{\omega_n \not\subset A(n - \log^2 n, n)\}$ is fast decaying. □

2.3. *Loop-free times.* One of the technical nuisances in the analysis of the loop-erased walk is that if $j < k$, it is not necessarily the case that

$$\mathbf{LE}\big(S[j,k]\big) = \mathbf{LE}\big(S[0,\infty)\big) \cap S[j,k].$$

However, this is the case for special times which we call *loop-free* times. We say that j is a *(global) loop-free time* if

$$S[0,j] \cap S[j+1,\infty) = \varnothing.$$

Proposition 2.1 shows that the probability that j is loop-free is comparable to $(\log j)^{-1/2}$. From the definition of chronological loop erasing, we can see the following. If $j < k$ and j, k are loop-free times, then for all $m \le j < k \le n$,

$$\mathbf{LE}\big(S[m,n]\big) \cap S[j,k] = \mathbf{LE}\big(S[0,\infty)\big) \cap S[j,k] = \mathbf{LE}\big(S[j,k]\big). \tag{2.10}$$

It will be important for us to give upper bounds on the probability that there is *no* loop-free time in a certain interval of time. If $m \le j < k \le n$, let $I(j,k;m,n)$ denote the event that for all $j \le i \le k-1$,

$$S[m,i] \cap S[i+1,n] \neq \varnothing.$$

Proposition 2.1 gives a lower bound on $\mathbb{P}[I(n,2n;0,3n)]$,

$$\mathbb{P}\big[I(n,2n;0,3n)\big] \ge \mathbb{P}\big\{S[0,n] \cap S[2n,3n] \neq \varnothing\big\} \asymp \frac{1}{\log n}.$$

The next lemma shows that

$$\mathbb{P}\big[I(n,2n;0,3n)\big] \asymp 1/\log n \tag{2.11}$$

by giving the matching upper bound.

LEMMA 2.6. *There exists $c < \infty$ such that $\mathbb{P}[I(n,2n;0,\infty)] \le c/\log n$.*

PROOF. Let $E = E_n$ denote the complement of $I(n,2n;0,\infty)$. We need to show that $\mathbb{P}(E) \ge 1 - O(1/\log n)$.

Let $k_n = \lfloor n/(\log n)^{3/4} \rfloor$ and let $A_i = A_{i,n}$ be the event that

$$A_i = \big\{S\big[n+(2i-1)k_n, n+2ik_n\big] \cap S\big[n+2ik_n+1, n+(2i+1)k_n\big] = \varnothing\big\}$$

and consider the events $A_1, A_2, \dots, A_\ell$ where $\ell = \lfloor (\log n)^{3/4}/4 \rfloor$. These are ℓ independent events each with probability greater than $c(\log n)^{-1/2}$ by Proposition 2.1. Therefore,

$$1 - \mathbb{P}(A_1 \cup \cdots \cup A_\ell) = \prod_{i=1}^{\ell} \big[1 - \mathbb{P}(A_i)\big] \le \exp\big\{-O\big((\log n)^{1/4}\big)\big\} = o\left(\frac{1}{\log^3 n}\right).$$

Let $B_i = B_{i,n}$ be the event $\{S[0, n+(2i-1)k_n] \cap S[n+2ik_n, \infty) = \varnothing\}$. By (2.1) in Proposition 2.1, $\mathbb{P}(B_i^c) \le c \log\log n / \log n$. Therefore,

$$\mathbb{P}(B_1 \cap \cdots \cap B_\ell) \ge 1 - \frac{c\ell \log\log n}{\log n} \ge 1 - O\left(\frac{\log\log n}{(\log n)^{1/4}}\right).$$

For $1 \le i \le \ell$, on the event $A_i \cap (B_1 \cap \cdots \cap B_\ell)$ the time $2ik_n$ is loop-free, hence E occurs. Therefore,

$$\mathbb{P}(E) \ge \mathbb{P}\big[(A_1 \cup \cdots \cup A_\ell) \cap (B_1 \cap \cdots \cap B_\ell)\big] \ge 1 - O\left(\frac{\log\log n}{(\log n)^{1/4}}\right). \tag{2.12}$$

This is a good estimate, but we need to improve on it.

Let C_j, $j = 1, \dots, 5$, denote the independent events (depending on n)

$$I\left(n\left[1 + \frac{3(j-1)+1}{15}\right], n\left[1 + \frac{3(j-1)+2}{15}\right]; n + \frac{(j-1)n}{5}, n + \frac{jn}{5}\right).$$

By (2.12), we see that $\mathbb{P}[C_j] \le o(1/(\log n)^{1/5})$, and hence

$$\mathbb{P}(C_1 \cap \cdots \cap C_5) \le o\left(\frac{1}{\log n}\right).$$

Let $D = D_n$ denote the event that at least one of the following ten things happens:

$$S\left[0, n\left(1 + \frac{j-1}{5}\right)\right] \cap S\left[n\left(1 + \frac{3(j-1)+1}{15}\right), \infty\right) \ne \varnothing, \quad j = 1, \dots, 5;$$

$$S\left[0, n\left(1 + \frac{3(j-1)+2}{15}\right)\right] \cap S\left[n\left(1 + \frac{j}{5}\right), \infty\right) \ne \varnothing, \quad j = 1, \dots, 5.$$

Each of these events has probability comparable to $1/\log n$, and hence $\mathbb{P}(D) \asymp 1/\log n$. Also,

$$I(n, 2n; 0, \infty) \subset (C_1 \cap \cdots \cap C_n) \cup D.$$

Therefore, $\mathbb{P}[I(n, 2n; 0, \infty)] \le c/\log n$. $\square$

COROLLARY 2.7.

1. *There exists $c < \infty$ such that if $0 \le j \le j+k \le n$, then*

$$\mathbb{P}\big[I(j, j+k; 0, n)\big] \le \frac{c \log(n/k)}{\log n}. \tag{2.13}$$

2. *Given $0 < \delta < 1$, let $I_{\delta,n} := \bigcup_{j=0}^{n-1} I(j, j+\delta n; 0, n)$. Then there exists a positive constant c such that for all $n \in \mathbb{N}$ and $\delta \in (0, 1)$, we have*

$$\mathbb{P}[I_{\delta,n}] \le \frac{c \log(1/\delta)}{\delta \log n}. \tag{2.14}$$

3. *There exist $c < \infty$ and a positive integer ℓ such that the following holds for all positive integers n. Let $\tilde{I}(m, r)$ denote the event that there is no loop-free point j with $\sigma_m \le j \le \sigma_r$, and let $k = k_n = \lfloor \log n \rfloor$. Then*

$$\mathbb{P}\{\tilde{I}(n - \ell k, n + \ell k) \mid \mathcal{F}_{n-3\ell k}\} \le c/n. \tag{2.15}$$

PROOF.

1. It suffices to prove under the assumption that $k \ge n^{1/2}$. Note that $I(j, j + k; 0, n)$ is contained in the union of the following two events:

$$I(j, j + k; j - k, j + 2k),$$
$$\{S[0, j - k] \cap S[j, n] \ne \varnothing\}, \quad \text{and} \quad \{S[0, j] \cap S[j + k, n] \ne \varnothing\}.$$

Since $k \ge n^{1/2}$, the probability of the first event is $O(1/\log n)$ by Lemma 2.6. By (2.1), the probabilities of the second two events are $O(\log(n/k)/\log n)$. This gives (2.13).

2. By (2.13), $\mathbb{P}\{I(i\delta n/3, (i + 1)\delta n/3; 0, n)\} = O(\log(\delta^{-1})/\log n)$ for all $0 \le i \le \lceil 3/\delta \rceil$. Now (2.14) follows from the fact that $I_{\delta,n}$ can be covered by these $I(i\delta n/3, (i + 1)\delta n/3; 0, n)$'s.

3. We will first consider walks starting at $z \in \partial C_{n-3\ell k}$ (with constants independent of z). Let E_n be the event

$$E_n = \{\sigma_{n-\ell k} \le e^{2n} \le 2e^{2n} \le \sigma_{n+\ell k}, S[\sigma_{n-\ell k}, \infty) \cap C_{n-2\ell k} = \varnothing\}.$$

Using (1.2) and (2.6), we can choose ℓ sufficiently large so that $\mathbb{P}(E_n) \ge 1 - 1/n$. On the event E_n, we have $\tilde{I}(n - \ell k, n + \ell k) \subset I(e^{2n}, 2e^{2n}; 0, \infty)$. Hence, by Lemma 2.6, we have $\mathbb{P}[\tilde{I}(n - \ell k, n + \ell k)] \le O(n^{-1})$.

More generally, if we start the random walk at the origin, stop at time $\sigma_{n-3\ell k}$, and then start again, we can use the result in the previous paragraph. Since $S[\sigma_{n-\ell k}, \infty) \cap C_{n-2\ell k} = \varnothing$ on the event E_n, attachment of the initial part of the walk up to $\sigma_{n-3\ell k}$ will not affect whether a time after $\sigma_{n-\ell k}$ is loop-free. This concludes the proof. $\square$

2.4. *Green function estimates.* Recall the Green function $G(\cdot, \cdot)$ on $\mathbb{Z}^4$ and $G_n(\cdot, \cdot)$ defined in Section 1.1. We write $G(x) = G(x, 0) = G(0, x)$ and $G_n(x) = G_n(x, 0) = G(0, x)$. As a standard estimate (see [12]), we have

$$G(x) = \frac{2}{\pi^2 |x|^2} + O(|x|^{-4}), \quad |x| \to \infty. \tag{2.16}$$

Here and throughout, we use the convention that if we say that a function on $\mathbb{Z}^d$ is $O(|x|^{-r})$ with $r > 0$, we still imply that it is finite at every point. In other words, for lattice functions, $O(|x|^{-r})$ really means $O(1 \wedge |x|^{-r})$. We do not make this assumption for functions on $\mathbb{R}^d$ which could blow up at the origin.

LEMMA 2.8. *For $w \in C_n$, let*

$$\hat{G}_n^2(w) = \sum_{z \in \mathbb{Z}^4} G(0,z) G_n(w,z) = \sum_{z \in C_n} G(0,z) G_n(w,z).$$

Then

$$\hat{G}_n^2(w) = \frac{8}{\pi^2}\big[n - \log|w|\big] + O\big(e^{-n}\big) + O\big(|w|^{-2}\log|w|\big).$$

In particular, if $w \in \partial C_{n-1}$,

$$\hat{G}_n^2(w) = \frac{8}{\pi^2} + O\big(e^{-n}\big). \tag{2.17}$$

PROOF. Let $f(x) = \frac{8}{\pi^2}\log|x|$ and note that

$$\Delta f(x) = \frac{2}{\pi^2|x|^2} + O\big(|x|^{-4}\big) = G(x) + O\big(|x|^{-4}\big),$$

where Δ denotes the discrete Laplacian. Also, we know that

$$f(w) = \mathbb{E}^w\big[f(S_{\sigma_n})\big] - \sum_{z \in C_n} G_n(w,z)\Delta f(z)$$

(this holds for any function f). Since $e^n \le |S_{\sigma_n}| \le e^n + 1$, we have $\mathbb{E}^w[f(S_{\sigma_n})] = \frac{8n}{\pi^2} + O(e^{-n})$. Therefore,

$$\sum_{z \in C_n} G_n(w,z) G(z) = \frac{8}{\pi^2}\big[n - \log|w|\big] + O\big(e^{-n}\big) + \epsilon,$$

where

$$|\epsilon| \le \sum_{z \in C_n} G_n(w,z) O\big(|z|^{-4}\big) \le \sum_{z \in C_n} O\big(|w-z|^{-2}\big) O\big(|z|^{-4}\big).$$

We split the sum on the right-hand side into three pieces:

$$\begin{aligned}\sum_{|z| \le |w|/2} O\big(|w-z|^{-2}\big) O\big(|z|^{-4}\big) &\le c|w|^{-2} \sum_{|z| \le |w|/2} O\big(|z|^{-4}\big) \\ &\le c|w|^{-2}\log|w|,\end{aligned}$$

$$\begin{aligned}\sum_{|z-w| \le |w|/2} O\big(|w-z|^{-2}\big) O\big(|z|^{-4}\big) &\le c|w|^{-4} \sum_{|x| \le |w|/2} O\big(|x|^{-2}\big) \\ &\le c|w|^{-2}.\end{aligned}$$

If we let C_n' the the set of $z \in C_n$ with $|z| > |w|/2$ and $|z-w| > |w|/2$, then

$$\sum_{z \in C_n'} O\big(|w-z|^{-2}\big) O\big(|z|^{-4}\big) \le \sum_{|z| > |w|/2} O\big(|z|^{-6}\big) \le c|w|^{-2}.$$

□

LEMMA 2.9. *If* $1 \leq m < n$ *and* $x \in C_m$,

$$\hat{G}_n^2(x) - G_n^2(x) = \frac{\pi^2}{2} + O(e^{m-n}).$$

PROOF. Set $N = e^n$. Using the martingale $M_t = |S_t|^2 - t$, we see that

$$\sum_{w \in C_n} G_n(x,w) = N^2 - |x|^2 + O(N). \tag{2.18}$$

By the strong Markov property, for all $w \in C_n$,

$$\min_{N \leq |z| \leq N+1} G(z,x) \leq G(w,x) - G_n(w,x) \leq \max_{N \leq |z| \leq N+1} G(z,x).$$

Set $\delta = (1+|x|)/N$. By (2.16), we have

$$G_n(x,w) = G(x,w) - \frac{2}{\pi^2 N^2}\left[1 + O(e^{m-n})\right].$$

Using (2.18), we see that

$$\begin{aligned}
&\sum_{w \in C_n} G_n(x,w) G_n(0,w) \\
&\quad = \sum_{w \in C_n} \left[G(x,w) - \frac{2}{\pi^2 N^2} + O(e^{m-n} N^{-2})\right] G_n(0,w) \\
&\quad = O(\delta) - \frac{2}{\pi^2} + \sum_{w \in C_n} G(x,w) G_n(0,w).
\end{aligned}$$

□

3. Slowly recurrent set and the subsequential limit. Simple random walk paths in $\mathbb{Z}^4$ are "slowly recurrent" sets in the terminology of [8]. In Section 3.1, we will consider a subcollections $\mathcal{X}_n$ of the collection of slowly recurrent sets and give uniform bounds for escape probabilities for such sets. Then in Section 3.2 we use these estimates to prove Propositions 1.6 and 1.7. This section does not rely on notions and results in [8, 9] as we will give a new and self-contained treatment.

3.1. *Sets in* $\mathcal{X}_n$. Given a subset $V \subset \mathbb{Z}^4$ and $m \in \mathbb{N}$ we write

$$V_m = V \cap A(m-1,m), \quad \text{and} \quad h_m = h_{m,V} = H(V_m).$$

Using (2.7), we can see that there exist $0 < c_1 < c_2 < \infty$ such that

$$c_1 h_m \leq H(z, V_m) \leq c_2 h_m \quad \forall z \in C_{m-2} \cup A(m+1, m+2). \tag{3.1}$$

DEFINITION 3.1. Let $\mathcal{X}_n$ denote the collection of subsets V of $\mathbb{Z}^4$ such that for all integers $m \geq \sqrt{n}$,

$$H(V_m) \leq \frac{\log^2 m}{m}.$$

Note that $\mathcal{X}_1 \subset \mathcal{X}_2 \subset \cdots$. Also if $\tilde{V} \subset V$ and $V \in \mathcal{X}_n$, then $\tilde{V} \in \mathcal{X}_n$, The following is an immediate corollary of Proposition 2.5.

PROPOSITION 3.2. *$\mathbb{P}\{S[0,\infty) \notin \mathcal{X}_n\}$ is a fast decaying sequence.*

Let E_m denote the event

$$E_m = E_{m,V} = \{S[1,\sigma_m] \cap V = \varnothing\}.$$

Note that $\mathbb{P}(E_m) = \mathrm{Es}(V;m)$. We will interchangeably use the two notions $\mathbb{P}(E_m)$ and $\mathrm{Es}(V;m)$ throughout this section. We write $\mathrm{hm}_m(z)$ for the harmonic measure of ∂C_m for random walk starting at the origin, that is,

$$\mathrm{hm}_m(z) = \mathbb{P}\{S_{\sigma_m} = z\} \quad \forall z \in \partial C_m.$$

If $V \subset \mathbb{Z}^4$ and $\mathbb{P}(E_m) > 0$, we write

$$\mathrm{hm}_m(z; V) = \mathbb{P}\{S_{\sigma_m} = z \mid E_m\}.$$

By the strong Markov property, we have

$$\mathbb{P}\{S[\sigma_m,\sigma_{m+1}] \cap V \neq \varnothing\} = \sum_{z \in \partial C_m} \mathrm{hm}_m(z) \mathbb{P}^z\{S[0,\sigma_{m+1}] \cap V \neq \varnothing\},$$

and

$$\begin{aligned}\mathbb{P}(E^c_{m+1} \mid E_m) &= \mathbb{P}\{S[\sigma_m,\sigma_{m+1}] \cap V \neq \varnothing \mid E_m\} \\ &= \sum_{z \in \partial C_m} \mathrm{hm}_m(z;V) \mathbb{P}^z\{S[0,\sigma_{m+1}] \cap V \neq \varnothing\}.\end{aligned} \tag{3.2}$$

PROPOSITION 3.3. *There exists $c < \infty$ such that if $V \in \mathcal{X}_n$, $m \geq n/10$, and $\mathbb{P}(E_{m+1} \mid E_m) \geq 1/2$, then $\mathbb{P}(E^c_{m+2} \mid E_{m+1}) \leq c \log^2 n/n$.*

PROOF. As in (3.2), we write

$$\mathbb{P}(E^c_{m+2} \mid E_{m+1}) = \sum_{z \in \partial C_{m+1}} \mathrm{hm}_{m+1}(z;V) \mathbb{P}^z\{S[0,\sigma_{m+2}] \cap V \neq \varnothing\}.$$

Using $\mathbb{P}(E_{m+1} \mid E_m) \geq 1/2$, we claim that there exists $c < \infty$ such that

$$\mathrm{hm}_{m+1}(z;V) \leq c\, \mathrm{hm}_{m+1}(z) \quad \forall z \in \partial C_{m+1}. \tag{3.3}$$

Indeed, we have

$$\begin{aligned}\mathrm{hm}_{m+1}(z;V) &= \frac{\mathbb{P}\{S_{\sigma_{m+1}} = z, E_{m+1}\}}{\mathbb{P}(E_{m+1})} \\ &\leq 2\frac{\mathbb{P}\{S_{\sigma_{m+1}} = z, E_{m+1}\}}{\mathbb{P}(E_m)} \leq 2\mathbb{P}\{S_{\sigma_{m+1}} = z \mid E_m\},\end{aligned}$$

and the Harnack inequality shows that

$$\mathbb{P}\{S_{\sigma_{m+1}} = z \mid E_m\} \le \sup_{w \in \partial C_m} \mathbb{P}^w\{S_{\sigma_{m+1}} = z\} \le c\,\mathrm{hm}_{m+1}(z).$$

Therefore, letting $r_k = \mathbb{P}\{S[\sigma_{m+1}, \sigma_{m+2}] \cap V_k \neq \varnothing\}$ for $k \in \mathbb{N}$, we have

$$\begin{aligned}\mathbb{P}(E^c_{m+2} \mid E_{m+1}) &\le c \sum_{z \in \partial C_{m+1}} \mathrm{hm}_{m+1}(z) \mathbb{P}^z\{S[0, \sigma_{m+2}] \cap V \neq \varnothing\} \\ &= c\mathbb{P}\{S[\sigma_{m+1}, \sigma_{m+2}] \cap V \neq \varnothing\} \le c \sum_{k=1}^{m+2} r_k.\end{aligned}$$

By Definition 3.1, the terms r_k for $k = m, m+1, m+2$ are bounded by

$$\begin{aligned}\mathbb{P}\{S[\sigma_{m+1}, \sigma_{m+2}] \cap (V_m \cup V_{m+1} \cup V_{m+2}) \neq \varnothing\} &\le H(V_m \cup V_{m+1} \cup V_{m+2}) \\ &\le \frac{c \log^2 n}{n}.\end{aligned}$$

Since $r_k \le \mathbb{P}\{S[\sigma_{m+1}, \infty) \cap C_k \neq \varnothing\}$, by (2.6), we have $r_k \le ce^{2(k-m)}$ for $k < m$. Therefore, for λ large enough, $\sum_{k=1}^{m-\lambda \log m} r_k \le cn^{-2}$.

For $m - \lambda \log m \le k \le m-1$, (2.7) and the definition of $\mathcal{X}_n$ imply that

$$r_k \le ce^{-2(m-k)} H(V_k) \le ce^{-2(m-k)} \frac{\log^2 k}{k}.$$

Summing over k gives the result. □

DEFINITION 3.4. Let $\tilde{\mathcal{X}}_n$ denote the set of $V \subset \mathcal{X}_n$ such that $\mathbb{P}(E_n) \ge 2^{-n/4}$.

The particular choice of $2^{-n/4}$ in this Definition 3.4 is rather arbitrary but it is convenient to choose a particular fast decaying sequence. For typical sets in $\mathcal{X}_n$, one expects that $\mathbb{P}(E_n)$ decays as a power in n, so "most" sets in $\mathcal{X}_n$ with $\mathbb{P}(E_n) > 0$ will also be in $\tilde{\mathcal{X}}_n$.

Recall $\mathrm{Es}(V; n) = \mathbb{P}\{S[1, \sigma_n] \cap V = \varnothing\}$, which is decreasing in n. We state the next immediate fact as a proposition so that we can refer to it.

PROPOSITION 3.5. *For any $r > 0$ and any random subset $V \subset \mathbb{Z}^4$,*

$$\mathbb{E}[\mathrm{Es}(V; m)^r; V \notin \tilde{\mathcal{X}}_n] \le \mathbb{P}[V \notin \mathcal{X}_n] + 2^{-rn/4}.$$

In particular, if $\mathbb{P}[V \notin \mathcal{X}_n]$ is fast decaying then so is the left-hand side.

PROPOSITION 3.6. *There exists $c < \infty$ such that if $V \in \tilde{\mathcal{X}}_n$, then*

$$\mathbb{P}(E^c_{j+1} \mid E_j) \le \frac{c \log^2 n}{n}, \quad \frac{3n}{4} \le j \le n.$$

PROOF. If $\mathbb{P}(E_{m+1} \mid E_m) < 1/2$ for all $n/4 \le m \le n/2$, then $\mathbb{P}(E_n) < 2^{-n/4}$ and $V \notin \tilde{\mathcal{X}}_n$. Therefore, we must have $\mathbb{P}(E_{m+1} \mid E_m) \ge 1/2$ for some $n/4 \le m \le n/2$. Now (for n sufficiently large) we can use Proposition 3.3 and induction to conclude that $\mathbb{P}(E_{k+1} \mid E_k) \ge 1/2$ for $m \le k \le n$. The result then follows from Proposition 3.3. $\square$

If we choose n_0 so large that $\frac{c \log^2 n_0}{n_0} \le 1 - 2^{-1/4}$, where c is as in Proposition 3.6, then it follows from Proposition 3.6 that $\tilde{\mathcal{X}}_n \subset \tilde{\mathcal{X}}_{n+1}$ for $n \ge n_0$. We fix the smallest such n_0 and set

$$\tilde{\mathcal{X}} = \bigcup_{j=n_0}^{\infty} \tilde{\mathcal{X}}_j.$$

Combining Propositions 3.3 and 3.6, and the union bound, we have

$$\mathbb{P}\big(E_{n+k}^c \mid E_n\big) \le \frac{ck \log^2 n}{n}, \quad k \in \mathbb{N}, V \in \tilde{\mathcal{X}}_n. \tag{3.4}$$

PROPOSITION 3.7. *There exists $c < \infty$ such that if $V \subset C_n$ and $V \in \tilde{\mathcal{X}}_n$, then*

$$\mathrm{Es}[V; n]\left[1 - \frac{c \log^2 n}{n}\right] \le \mathrm{Es}[V] \le \mathrm{Es}[V; n]. \tag{3.5}$$

PROOF. The upper bound is trivial. For the lower bound, we first use the previous proposition to see that

$$\mathrm{Es}(V; n+1) \ge \mathrm{Es}(V; n)\big[1 - O\big(\log^2 n/n\big)\big].$$

Since $\mathrm{Es}(V) \ge \mathrm{Es}(V; n+1)(1 - \max_{z \in \partial C_{n+1}} H(z, V))$, it suffices to show that there exists c such that for all $z \in \partial C_{n+1}$,

$$H(z, V) \le c \log^2 n/n.$$

This can be done by dividing V into V_j's similarly as in the proof of Proposition 3.3, (see the bound for $\sum_1^{m+2} r_j$ there). $\square$

The next proposition is the key to the analysis of slowly recurrent sets. It says that the distribution of the first visit to ∂C_n given that one has avoided the set V is very close to the unconditioned distribution. We would not expect this to be true for recurrent sets that are not slowly recurrent.

PROPOSITION 3.8. *There exists $c < \infty$ such that if $V \in \tilde{\mathcal{X}}_n$ we have*

$$\mathrm{hm}_n(z; V) \le \mathrm{hm}_n(z)\left[1 + \frac{c \log^3 n}{n}\right] \quad \forall z \in \partial C_n. \tag{3.6}$$

Moreover,

$$\sum_{z\in\partial C_n} \left|\mathrm{hm}_n(z) - \mathrm{hm}_n(z; V)\right| \le \frac{c\log^3 n}{n}. \tag{3.7}$$

PROOF. Let $k = \lfloor \log n \rfloor$. By (3.4), we have

$$\mathbb{P}(E_n^c \mid E_{n-k}) \le \frac{c\log^3 n}{n}. \tag{3.8}$$

Consider a random walk starting on ∂C_{n-k} with the distribution $\mathrm{hm}_{n-k}(\cdot; V)$ and let ν denote the distribution of the first visit to ∂C_n. In other words, ν is the distribution of the first visit to ∂C_n conditioned on the event E_{n-k}. Using (2.5), we see that for $z \in \partial C_n$,

$$\nu(z) = \mathrm{hm}_n(z)\left[1 + O\left(n^{-1}\right)\right].$$

By (3.8), for each $z \in \partial C_n$, we have

$$\begin{aligned}\mathrm{hm}_n(z; V) = \mathbb{P}\{S_{\sigma_n} = z \mid E_n\} &\le \frac{\mathbb{P}(E_{n-k})}{\mathbb{P}(E_n)}\mathbb{P}\{S_{\sigma_n} = z \mid E_{n-k}\}\\ &\le \frac{\nu(z)}{1 - \mathbb{P}(E_n^c \mid E_{n-k})} \le \mathrm{hm}_n(z)\left[1 + O\left(\frac{\log^3 n}{n}\right)\right].\end{aligned}$$

Since $\mathrm{hm}_n(\cdot)$ and $\mathrm{hm}_n(\cdot; V)$ are probability measures on ∂C_n, we have

$$\begin{aligned}\sum_{z\in\partial C_n} \left|\mathrm{hm}_n(z) - \mathrm{hm}_n(z; V)\right| &= 2\sum_{z\in\partial C_n} \left[\mathrm{hm}_n(z; V) - \mathrm{hm}_n(z)\right]_+\\ &\le \frac{c\log^3 n}{n}\sum_{z\in\partial C_n} \mathrm{hm}_n(z) \le \frac{c\log^3 n}{n}.\end{aligned}$$

□

3.2. *Along a subsequence.* In this section, we prove Propositions 1.6 and 1.7 via the estimates proved for sets in $\tilde{\mathcal{X}}$. For $V \in \tilde{\mathcal{X}}_{n^4}$, let

$$V_n^* := V \cap \left\{e^{(n-1)^4 + 4(n-1)} \le |z| < e^{n^4 - 4n}\right\}.$$

PROPOSITION 3.9. *There exists $c < \infty$ such that if $V \in \tilde{\mathcal{X}}_{n^4}$,*

$$\mathrm{Es}\left(V; (n+1)^4\right) = \mathrm{Es}\left(V; n^4\right)\left[1 - H\left(V_{n+1}^*\right) + O\left(\frac{\log^2 n}{n^3}\right)\right].$$

PROOF. Let $\tau_n = \inf\{j : S_j \notin C_{n^4}\}$ and recall E_n and $\mathrm{hm}_n(z; V)$. We observe that $(\mathrm{Es}(V; n) - \mathrm{Es}(V; (n+1)^4))/\mathrm{Es}(V; n^4)$ is bounded by

$$\mathbb{P}\left\{S[\tau_n, \tau_{n+1}] \cap V_{n+1}^* \ne \varnothing \mid E_{n^4}\right\} + \mathbb{P}\left\{S[\tau_n, \tau_{n+1}] \cap \left(V \setminus V_{n+1}^*\right) \ne \varnothing \mid E_{n^4}\right\}.$$

To bound the first term, we have

$$
\begin{aligned}
&\big|\mathbb{P}\{S[\tau_n,\tau_{n+1}]\cap V^*_{n+1}\neq\varnothing \mid E_{n^4}\}-\mathbb{P}\{S[\tau_n,\tau_{n+1}]\cap V^*_{n+1}\neq\varnothing\}\big|\\
&\quad\le \sum_{z\in\partial C_{n^4}}\big|\mathrm{hm}_{n^4}(z;V)-\mathrm{hm}_{n^4}(z)\big|\mathbb{P}^z\{S[0,\tau_{n+1}]\cap V^*_{n+1}\neq\varnothing\}\\
&\quad\le \frac{c\log^3 n}{n^4}\sup_{z\in\partial C_{n^4}}\mathbb{P}^z\{S[0,\tau_{n+1}]\cap V^*_{n+1}\neq\varnothing\}\\
&\quad\le \frac{c\log^3 n}{n^4}\mathbb{P}\{S[\tau_n,\tau_{n+1}]\cap V^*_{n+1}\neq\varnothing\},
\end{aligned}
$$

where the three inequalities are due to the strong Markov property, (3.6) and Harnack inequality respectively.

Using (2.5), we have $\mathbb{P}\{S[\tau_n,\tau_{n+1}]\cap V^*_{n+1}\neq\varnothing\}=H(V^*_{n+1})[1+O(e^{-4n})]$. Hence, it suffices to prove that

$$
\mathbb{P}\{S[\tau_n,\tau_{n+1}]\cap(V\setminus V^*_{n+1})\neq\varnothing \mid E_{n^4}\}=O\left(\frac{\log^2 n}{n^3}\right),
$$

which by the strong Markov property and (3.6), can be further reduced to showing that

$$
\mathbb{P}\{S[\tau_n,\tau_{n+1}]\cap(V\setminus V^*_{n+1})\neq\varnothing\}=O\left(\frac{\log^2 n}{n^3}\right).
$$

Note that $V\setminus V^*_{n+1}$ is contained in the union of C_{n^4-4n} and $O(n)$ sets of the form V_m with $m\ge n^4-4n$. By Definition 3.1 and the union bound,

$$
H\big((V\setminus V^*_{n+1})\cap\{e^{n^4-4n}\le|z|\le e^{(n+1)^4}\}\big)=O\left(\frac{\log^2 n}{n^3}\right).
$$

By Lemma 2.3, $\mathbb{P}\{S[\tau_n,\tau_{n+1}]\cap C_{n^4-4n}\neq\varnothing\}=O(e^{-8n})$. This concludes the proof. □

COROLLARY 3.10. *If $V\in\tilde{\mathcal{X}}_{n^4}$, $m\ge n$, and $m^4\le k\le(m+1)^4$, then*

$$
\mathrm{Es}(V;k)=\mathrm{Es}(V;n^4)\exp\left\{-\sum_{j=n+1}^{m}H(V^*_j)\right\}\left[1+O\left(\frac{\log^4 n}{n}\right)\right].
$$

PROOF. By Proposition 3.9, if $m>n$ we have

$$
\frac{\mathbb{P}(E_{m^4})}{\mathbb{P}(E_{n^4})}=\prod_{j=n+1}^{m}\left[1-H(V^*_j)+O\left(\frac{\log^2 j}{j^3}\right)\right].
$$

By Definition 3.1 and the union bound, we have $H(V_j^*) = O(\log^2 j/j)$. Hence,

$$\frac{\mathbb{P}(E_{m^4})}{\mathbb{P}(E_{n^4})} = \prod_{j=n+1}^{m}\left[e^{-H(V_j^*)} + O\left(\frac{\log^4 j}{j^2}\right)\right]$$

$$= \left[1 + O\left(\frac{\log^4 n}{n}\right)\right]\exp\left\{-\sum_{j=n+1}^{m} H(V_j^*)\right\}.$$

On the other hand, (3.4) implies that $\mathbb{P}(E_k) = \mathbb{P}(E_{m^4})[1 - O(\log^2 m/m)]$ for $m^4 \le k \le (m+1)^4$. This concludes the proof. □

Now we apply our theory to LERW. Recall the setup in Section 1.2.

PROOF OF PROPOSITION 1.6. Let $k = \lceil \log^2 n \rceil$. We will show the stronger result that

$$p_n = \mathbb{E}[\mathrm{Es}(\Gamma_n)^r] = p_{n-k}[1 + O(\log^4 n/n)], \tag{3.9}$$

ad similarly for p_{n+1}.

By Proposition 2.1, $\mathbb{E}[\mathrm{Es}(\Gamma_n)] \ge O(n^{-1/2})$. On the other hand, $\mathrm{Es}(\Gamma_n) \le 1$. Therefore, p_n decays polynomially. By Proposition 3.5,

$$\mathbb{E}[\mathrm{Es}(\Gamma_n)^r] = \mathbb{E}[\mathrm{Es}(\Gamma_n)^r 1_{\Gamma_n \in \tilde{\mathcal{X}}_n}](1 + \epsilon_n), \tag{3.10}$$

where ϵ_n is fast decaying. By (3.5), we have

$$\mathbb{E}[\mathrm{Es}(\Gamma_n)^r] = \mathbb{E}[\mathrm{Es}(\Gamma_n; n)^r][1 + O(\log^2 n/n)]. \tag{3.11}$$

By Lemma 2.4, except for an event of fast decaying probability,

$$\Gamma \cap C_{n-k} \subset \Gamma_n \subset V, \tag{3.12}$$

where $V := (\Gamma \cap C_{n-k}) \cup [S[0,\infty) \setminus C_{n-k}]$. If (3.12) occurs, then we have

$$\mathrm{Es}[\Gamma; n-k] = \mathrm{Es}[V; n-k] \ge \mathrm{Es}[\Gamma_n; n] \ge \mathrm{Es}[\Gamma_n; n+k] \ge \mathrm{Es}[V; n+k].$$

By (3.4), we have

$$\mathbb{E}[\mathrm{Es}[V; n+k]^r; V \in \tilde{\mathcal{X}}_{n-k}]$$

$$= \mathbb{E}[\mathrm{Es}[V; n-k]^r; V \in \tilde{\mathcal{X}}_{n-k}][1 - O(\log^4 n/n)].$$

Since $\mathbb{E}[\mathrm{Es}[V; n+k]^r]$ decays like a power of n, by Proposition 3.5,

$$\mathbb{E}[\mathrm{Es}[V; n+k]^r] = \mathbb{E}[\mathrm{Es}[V; n-k]^r][1 - O(\log^4 n/n)].$$

Now (3.9) follows from (3.11) and

$$\mathbb{E}[\mathrm{Es}(\Gamma_n; n)^r] \ge \mathbb{E}[\mathrm{Es}[V; n+k]^r] = p_{n-k}[1 - O(\log^4 n/n)].$$

A similar argument gives $p_{n+1} = p_{n-k}[1 - O(\log^4 n/n)]$. □

A useful corollary of the proof is

$$\mathbb{E}[\mathrm{Es}(\Gamma; n)^r] = p_n[1 + O(\log^4 n/n)]. \tag{3.13}$$

PROOF OF PROPOSITION 1.7. Let $Q_n = \mathrm{Es}[\Gamma; n^4]$ and

$$\Gamma_n^* = \Gamma \cap A((n-1)^4 + 4(n-1), n^4 - 4n).$$

Then, by Proposition 3.9, if $\Gamma \in \tilde{\mathcal{X}}_{n^4}$, we have

$$Q_{n+1} = Q_n\left[1 - H(\Gamma_{n+1}^*) + O\left(\frac{\log^2 n}{n^3}\right)\right].$$

Applying Proposition 3.5 to $V = \Gamma$, we have

$$\mathbb{E}[Q_{n+1}^r] = \mathbb{E}[Q_n^r[1 - rH(\Gamma_{n+1}^*)]] + \mathbb{E}[Q_n^r]O\left(\frac{\log^2 n}{n^3}\right).$$

Recall $\widetilde{I}(\cdot,\cdot)$ in (2.15). We see that $\mathbb{P}\{\Gamma_{n+1}^* \neq \tilde{\Gamma}_{n+1} \mid \mathcal{F}_{n^4}\}$ is bounded by

$$\begin{aligned}&\mathbb{P}\{\widetilde{I}[n^4 + n, n^4 + 4n] \mid \mathcal{F}_{n^4}\}\\ &\quad + \mathbb{P}\{\widetilde{I}[(n+1)^4 - 4(n+1), (n+1)^4 - (n+1)] \mid \mathcal{F}_{n^4}\},\end{aligned}$$

which by (2.15) is further bounded by $O(n^{-4})$. Therefore,

$$\mathbb{E}[Q_n^r|H(\Gamma_{n+1}^*) - H(\tilde{\Gamma}_{n+1})|] \le O(n^{-4})\mathbb{E}[Q_n^r].$$

Hence,

$$\mathbb{E}[Q_{n+1}^r] = \mathbb{E}[Q_n^r[1 - rH(\tilde{\Gamma}_{n+1})]] + \mathbb{E}[Q_n^r]O\left(\frac{\log^2 n}{n^3}\right). \tag{3.14}$$

Using (2.5) in Lemma 2.3, we can see that

$$\mathbb{E}[H(\tilde{\Gamma}_{n+1}) \mid \mathcal{F}_{n^4}] = \mathbb{E}[H(\tilde{\Gamma}_{n+1})][1 + o(e^{-4n})] = \tilde{h}_{n+1}[1 + o(e^{-4n})]. \tag{3.15}$$

Let $q_n := p_{n^4}$. Combining (3.13), (3.14) and (3.15), we get

$$\begin{aligned} q_{n+1} &= q_n\left[1 - r\tilde{h}_{n+1} + O\left(\frac{\log^2 n}{n^3}\right)\right]\\ &= q_n \exp\{-r\tilde{h}_{n+1}\}\left[1 + O\left(\frac{1}{n^2}\right)\right],\end{aligned}$$

where the last inequality uses $\tilde{h}_{n+1} = O(n^{-1})$. In particular, if $m > n$,

$$q_m = q_n\left[1 + O\left(\frac{1}{n}\right)\right]\exp\left\{-r\sum_{j=n+1}^{m} \tilde{h}_j\right\}$$

from which Proposition 1.7 follows. $\square$

By Proposition 1.6, we see that

$$p_m = p_{n^4}\big[1 + O\big(\log^4 n/n\big)\big] \quad \text{if } n^4 \le m \le (n+1)^4.$$

Combined with Proposition 1.7, we immediately get the following.

COROLLARY 3.11. *There exists $c_0 < \infty$ such that as $m \to \infty$,*

$$p_m = \left[c_0 + O\left(\frac{\log^4 m}{m^{1/4}}\right)\right] \exp\left\{-r \sum_{j=1}^{\lfloor m^{1/4} \rfloor} \tilde{h}_j\right\}.$$

4. From the subsequential limit to the full limit. In this section, we prove Propositions 1.8 and 1.9. The proof of Proposition 1.8 is the technical bulk of this section, which will be given in Section 4.2. Let us first conclude the proof of Proposition 1.9 assuming Proposition 1.8.

PROOF OF PROPOSITION 1.9. Given Proposition 1.8, the $s = 0$ case follows from Corollary 3.11. For the general case, recall that $1 \le G_n \le 8$ and G_n converge to G_∞ almost surely. Moreover, Lemma 2.4 implies that there exists a fast decaying sequence $\{\epsilon_n\}$ such that if $m \ge n$, $\mathbb{P}\{|G_n - G_m| \ge \epsilon_n\} \le \epsilon_n$. Therefore, $\mathbb{E}[\phi_n^{-r} Z_n^r G_n^{-s}] - \mathbb{E}[\phi_n^{-r} Z_n^r G_\infty^{-s}]$ is fast decaying and

$$\big|\mathbb{E}(\phi_n^{-r} Z_n^r G_\infty^{-s}) - \mathbb{E}(\phi_m^{-r} Z_m^r G_\infty^{-s})\big| \le c\big|\mathbb{E}(\phi_n^{-r} Z_n^r) - \mathbb{E}(\phi_m^{-r} Z_m^r)\big|.$$

Take $m \to \infty$, we see that the $s \ne 0$ case follows from the $s = 0$ case. $\square$

4.1. *Harmonic measure of the range of SRW.* We start by proving two estimates for the harmonic measure of the range of random walk.

LEMMA 4.1. *Let*

$$\sigma_n^- = \sigma_n - \lfloor n^{-1/4} e^{2n} \rfloor, \qquad \sigma_n^+ = \sigma_n + \lfloor n^{-1/4} e^{2n} \rfloor,$$
$$n' = \lceil n + n^{4/5} \rceil, \qquad \mathcal{S}_n^- = S[0, \sigma_n^-], \qquad \mathcal{S}_n^+ = S[\sigma_n^+, \sigma_{n'}],$$
$$R_n = \max_{x \in \mathcal{S}_n^-} H(x, \mathcal{S}_n^+) + \max_{y \in \mathcal{S}_n^+} H(y, \mathcal{S}_n^-).$$

Then, for all n sufficiently large,

$$\mathbb{P}\{R_n \ge n^{-1/6}\} \le n^{-1/3}. \tag{4.1}$$

Our proof will actually give a stronger estimate, but (4.1) is all that we need and makes for a somewhat cleaner statement.

PROOF OF LEMMA 4.1. Let $m = n + \lceil \log n \rceil$. Recall from (1.2) that there exists $c_0 < \infty$ such that

$$\mathbb{P}\{\sigma_n \ge c_0 e^{2n} \log n\} \le n^{-1} \quad \text{and} \quad \mathbb{P}\{\sigma_m \ge c_0 e^{2n} n^2 \log n\} \le n^{-1}.$$

Let $V = V_n = S[\sigma_n^+, \sigma_m]$,

$$\tilde{R}_n = \max_{x \in \mathcal{S}_n^-} H(x, V) + \max_{y \in V} H(y, \mathcal{S}_n^-),$$

$$R_n^* = \max_{x \in \mathcal{S}_n^-} H(x, \mathcal{S}_n^+ \setminus V) + \max_{y \in \mathcal{S}_n^+ \setminus V} H(y, \mathcal{S}_n^-),$$

and note that $R_n \le \tilde{R}_n + R_n^*$. We claim that for n sufficiently large,

$$\mathbb{P}\{R_n^* \ge n^{-1/6}/2\} \le O(n^{-1}). \tag{4.2}$$

To see this, let $\ell = \lfloor n + [(\log n)/2] \rfloor$, and let U denote the event $U = \{(S_n^+ \setminus V) \cap C_\ell = \varnothing\}$. By (2.6), $\mathbb{P}(U) \ge 1 - O(n^{-1})$, and on the event U we have $\mathcal{S}_n^- \subset C_n \subset C_\ell \subset (S_n^+ \setminus V)^c$. Therefore, on the event $U \cap \{S[0, \infty) \in \mathcal{X}_n\}$, for $y \in \mathcal{S}_n^-$, $x \in S_n^+ \setminus V$, we have the following two bounds:

$$H(y, S_n^+ \setminus V) \le cH(S[0, \infty) \cap A(n+1, n+n^{4/5})) \le cn^{-1/5} \log^2 n;$$

$$H(x, \mathcal{S}_n^-) \le H(x, C_n) \le c/n.$$

This gives (4.2).

Let $N = N_n = \lceil c_0 e^{2n} \log n \rceil$, $M = M_n = \lceil c_0 e^{2n} n^2 \log n \rceil$ and $k = k_n = \lfloor n^{-1/4} e^{2n}/4 \rfloor$. For each integer $0 \le j \le N/(k+1)$, let $E_j = E_{j,n}$ be event that at least one of the following holds:

$$\max_{0 \le i \le jk} H(S_i, S[(j+1)k, M]) \ge \frac{\log n}{n^{1/4}}, \tag{4.3}$$

$$\max_{(j+1)k \le i \le M} H(S_i, S[0, jk]) \ge \frac{\log n}{n^{1/4}}. \tag{4.4}$$

Then for large enough n, we have $\{\tilde{R}_n \ge n^{-1/3}, jk \le \sigma_n \le (j+1)k\} \subset E_j$.

Recall the notion $Y_{n,\alpha}$ in Corollary 2.2. For fixed j, using the reversibility of simple random walk, the probabilities of both the event in (4.3) and in (4.4) are bounded by $\mathbb{P}[Y_{M,N/k} \ge n^{-1/4} \log n] = O(n^{-3/4})$. Therefore, $\mathbb{P}(E_j) \le O(n^{-3/4})$, and hence

$$\mathbb{P}\{\tilde{R}_n \ge n^{-1/3}\} \le O(n^{-1}) + \frac{N}{k} O(n^{-3/4}) \le O\left(\frac{\log n}{n^{1/2}}\right).$$

Combining this with (4.2) gives the proof. $\square$

The next lemma will use the notion of capacity $\mathrm{cap}(V)$ for a subset $V \subset \mathbb{Z}^4$. We will not review the definition but only recall three key facts:

$$\mathrm{cap}(V \cup V') \le \mathrm{cap}(V) + \mathrm{cap}(V'), \quad V, V' \subset \mathbb{Z}^4; \tag{4.5}$$

$$\mathrm{cap}(\{z + v : v \in V\}) = \mathrm{cap}(V), \quad V \subset \mathbb{Z}^4, z \in \mathbb{Z}^4; \tag{4.6}$$

$$\mathrm{cap}(V) \asymp |z|^2 H(z, V), \quad V \subset C_n, z \notin C_{n+1}. \tag{4.7}$$

See [12], Section 6.5, in particular, Proposition 6.5.1, for definitions and properties. Combining these with the estimates for hitting probabilities, we have

$$\mathbb{E}[\operatorname{cap}(S[0,n^2])] \asymp n^2/\log n.$$

By Markov inequality, there exists $\delta > 0$ such that

$$\mathbb{P}\left\{\operatorname{cap}(S[0,n^2]) \geq \frac{n^2}{\delta \log n}\right\} \leq \delta. \tag{4.8}$$

By iterating (4.8) and using the strong Markov property and subadditivity as in the proof of Proposition 2.5, we see that there exists c, β such that for all n and all $a > 0$,

$$\mathbb{P}\left\{\operatorname{cap}(S[0,n^2]) \geq \frac{an^2}{\log n}\right\} \leq ce^{-\beta a}. \tag{4.9}$$

LEMMA 4.2. *For all $j, m \in \mathbb{N}$, let $L[j,m] = \operatorname{cap}(S[j, j+m])$. For $k, n \in \mathbb{N}$, let $\bar{L}(n;k) = \max_{j \leq n} L[j,k]$. Then for every $u < \infty$*

$$\mathbb{P}\{\bar{L}(n^u e^{2n}; n^{-1/4}e^{2n}) \geq 2n^{-11/10}e^{2n}\} \quad \textit{is fast decaying}. \tag{4.10}$$

PROOF. Let $k = \lceil n^{-1/4}e^{2n}\rceil$. Let U denote the event in (4.10). By the subadditivity of capacity (4.5) and the union bound, we have

$$U \subset \bigcup_{i=1}^{n^{u+1}} \{L[ik,k] \geq n^{-11/10}e^{2n}\}.$$

By (4.6), the events in the union are identically distributed. Therefore,

$$\mathbb{P}(U) \leq n^{u+1}\mathbb{P}\left\{L[0;k] \geq n^{3/20}\frac{e^{2n}}{n^{1/4}n}\right\},$$

which is fast decaying by (4.9). □

4.2. *Proof of Proposition* 1.8. The strategy is to find a $u > 0$, and for each n a random set $U = U(n) \subset \mathbb{Z}^4$ that can be written as a disjoint union

$$U = \bigcup_{j=(n-1)^4+1}^{n^4} U_j \tag{4.11}$$

such that the following four conditions hold where:

$$U \subset \tilde{\Gamma}_n; \tag{4.12}$$

$$U_j \subset \eta^j, \quad j = (n-1)^4 + 1, \dots, n^4; \tag{4.13}$$

$$\mathbb{E}[H(\tilde{\Gamma}_n \setminus U)] + \sum_{(n-1)^4 < j \leq n^4} \mathbb{E}[H(\eta^j \setminus U_j)] \leq O(n^{-(1+u)}); \tag{4.14}$$

$$\max_{(n-1)^4 < j \leq n^4} \max_{x \in U_j} H(x, U \setminus U_j) \leq n^{-u}. \tag{4.15}$$

We will first show that finding such a set gives the result. Taking expectations and using (4.12)–(4.14), we get

$$\mathbb{E}[H(\tilde{\Gamma}_n)] = O(n^{-(1+u)}) + \mathbb{E}[H(U)];$$
$$\sum_{(n-1)^4<j\le n^4} \mathbb{E}[H(U_j)] = O(n^{-(1+u)}) + \sum_{(n-1)^4<j\le n^4} \mathbb{E}[H(\eta^j)].$$

Let $\tilde{S}$ be a simple random walk (independent of U) starting at the origin and let J_n be the number of integers j with $(n-1)^4 < j \le n^4$ and such that $\tilde{S}[0,\infty) \cap U_j \neq \varnothing$. Let $\tilde{\mathbb{P}}$ and $\tilde{\mathbb{E}}$ be the probability and expectation over $\tilde{S}$ with U fixed. Since the U_j are disjoint, (4.15) and the strong Markov property imply for $k \ge 1$,

$$\tilde{\mathbb{P}}\{J_n \ge k+1 \mid J_n \ge k\} \le n^{-u}.$$

Therefore, $\tilde{\mathbb{E}}[J_n] \le \tilde{\mathbb{P}}[J_n \ge 1][1 + O(n^{-u})]$.

Since $\tilde{\mathbb{E}}[J_n] = \sum_{(n-1)^4<j\le n^4} H(U_j)$ and $H(U) = \tilde{\mathbb{P}}[J_n \ge 1]$, we have

$$H(U) \ge [1 - O(n^{-u})] \sum_{(n-1)^4<j\le n^4} H(U_j).$$

Taking expectations over U and using $\mathbb{E}[\tilde{\Gamma}_n] \le O(n^{-1})$, we get

$$\left|\mathbb{E}[H(U)] - \sum_{(n-1)^4<j\le n^4} \mathbb{E}[H(U_j)]\right| \le O(n^{-1-u}).$$

Therefore, it remains to find the sets U and U_j's satisfying (4.12)–(4.15).

Let $\sigma_j^{\pm} = \sigma_j \pm \lfloor j^{-1/4} e^{2j} \rfloor$ as in Lemma 4.1 and $\tilde{\omega}_j = S[\sigma_{j-1}^+, \sigma_j^-]$. We will let U be defined as in (4.11), where $U_j = \eta^j \cap \tilde{\omega}_j$ unless one of the following six events occurs in which case we set $U_j = \varnothing$ (we assume $(n-1)^4 < j \le n^4$):

1. If $j \le (n-1)^4 + 8n$ or $j \ge n^4 - 8n$.
2. If $H(\omega_j) \ge j^{-1} \log^2 j$.
3. If $\omega_j \cap C_{j-8\log n} \neq \varnothing$.
4. If $H(\omega_j \setminus \tilde{\omega}_j) \ge j^{-1-u}$.
5. If it is not true that there exist loop-free points in both $[\sigma_{j-1}, \sigma_{j-1}^+]$ and $[\sigma_j^-, \sigma_j]$.
6. If $\sup_{x\in\tilde{\omega}_j} H(x, S[0, \sigma_{n^4}] \setminus \omega_j) \ge j^{-1/6}$.

We need to show that (4.12)–(4.15) hold for some $u > 0$.

Throughout this proof, we assume n is large enough. The definition of U_j immediately implies (4.13). Combining Conditions 1 and 3, we see that $U_j \subset A((n-1)^4 + 6n, n^4 - 6n)$. Moreover, if there exists loop-free points in $[\sigma_{j-1}, \sigma_{j-1}^+]$ and $[\sigma_j^-, \sigma_j]$, then $\tilde{\eta}_n \cap \tilde{\omega}_j = \eta_j \cap \tilde{\omega}_j$. Therefore, the U_j are disjoint and (4.12) holds. Also, condition 6 immediately yields that (4.15) holds for $u \le 1/6$.

In order to establish (4.14), we first note that

$$\big(\tilde{\Gamma}_n \cup \eta^{(n-1)^4+1} \cup \cdots \cup \eta^{n^4}\big) \setminus U \subset \bigcup_{(n-1)^4 < j \le n^4} V_j,$$

where

$$V_j = \begin{cases} \omega_j & \text{if } U_j = \varnothing, \\ \omega_j \setminus \tilde{\omega}_j & \text{if } U_j = \eta^j \cap \tilde{\omega}_j. \end{cases}$$

Hence, it suffices to find $0 < u \le 1/3$ such that

$$\sum_{(n-1)^4 < j \le n^4} \big(\mathbb{E}\big[H(\omega_j); U_j = \varnothing\big] + \mathbb{E}\big[H(\omega_j \setminus \tilde{\omega}_j)\big]\big) \le cn^{-1-u}.$$

To estimate $\mathbb{E}[H(\omega_j \setminus \tilde{\omega}_j)]$, we use (4.7), and Lemma 4.2 to see that except for an event of fast decaying probability

$$H(\omega_j \setminus \tilde{\omega}_j) \le O\big(j^{-11/10}\big), \tag{4.16}$$

and hence $\mathbb{E}[H(\omega_j \setminus \tilde{\omega}_j)] \le O(j^{-11/10})$ and

$$\sum_{(n-1)^4 < j \le n^4} \mathbb{E}\big[H(\omega_j \setminus \tilde{\omega}_j)\big] \le c \sum_{(n-1)^4 < j \le n^4} j^{-11/10} \le cn^{-\frac{7}{5}}.$$

For $i = 1, 2, 3, 4, 5, 6$, let E_j^i be the event that the ith condition in the definition of U_j holds but none of the previous ones hold. Since $\{U_j = \varnothing\} = E_j^1 \cup \cdots \cup E_j^6$, to estimate $\mathbb{E}[H(\omega_j); U_j = \varnothing]$, we just need to estimate $\sum_{(n-1)^4 < j \le n^4} \mathbb{E}[H(\omega_j); E_j^i]$ case by case:

1. Since $\mathbb{E}[H(\omega_j)] \asymp j^{-1}$ for each j, we have

$$\sum_{(n-1)^4 < j \le n^4} \mathbb{E}\big[H(\omega_j); E_j^1\big] = O\big(n^{-3}\big).$$

2. By Proposition 2.5, $\mathbb{P}\{H(\omega_j) \ge j^{-1} \log^2 j\}$ is fast decaying in j. This takes care of $\sum_{(n-1)^4 < j \le n^4} \mathbb{E}[H(\omega_j); E_j^2]$.

In the event $E_j^3 \cup \cdots \cup E_j^6$, we have $H(\omega_j) < j^{-1} \log^2 j$. Hence,

$$\mathbb{E}\big[H(\omega_j); E_j^3 \cup \cdots \cup E_j^6\big] \le \frac{\log^2 j}{j} \mathbb{P}\big(E_j^3 \cup \cdots \cup E_j^6\big).$$

In particular, it suffices to prove that there exists $u > 0$ such that

$$\mathbb{P}\big(E_j^i\big) \le j^{-u} \quad \text{for } i = 3, 4, 5, 6 \text{ and } (n-1)^4 + 8n < j < n^4 - 8n.$$

3. (2.6) in Lemma 2.3 gives $\mathbb{P}(E_j^3) \le \mathbb{P}\{\omega_j \cap C_{j-\log j} \ne \varnothing\} \le O(j^{-2})$.
4. The bound on $\mathbb{P}(E_j^4)$ is already done in (4.16).

5. Let $I = I_{\delta,n}$ be as in (2.14) substituting in $n = e^{2j} j^{1/16}$ and $\delta = j^{-7/16}$ so that $\delta n = e^{2j} j^{-3/8}$. Using (2.14), we have $\mathbb{P}(I) = o(j^{-1/4})$. Note that the event that there is no loop-free time in either $[\sigma_{j-1}, \sigma_{j-1}^{+}]$ or $[\sigma_j^{-}, \sigma_j]$ is contained in the union of I and the two events

$$\{\sigma_{j+1} \geq e^{2j} j^{1/16}\} \quad \text{and} \quad \{S[e^{2j} j^{1/16}, \infty) \cap S[0, \sigma_{j+1}] \neq \varnothing\}.$$

The probability of the first event is fast decaying by (1.2) and the probability of the second is $o(1/j)$ by (2.1). Hence $\mathbb{P}(E_j^5) = o(j^{-1/4})$.

6. By Lemma 4.1, for large enough n, we have $\mathbb{P}(E_j^6) \leq j^{-1/3}$.

5. Exact relation. In this section, we first prove the elementary lemma promised at the end of Section 1.2. Then we give the asymptotics of the long-range intersection probability of SRW and LERW in terms of $\hat{G}_n^2$ defined in Section 2.4, which concludes our proof of Theorem 1.5.

5.1. *A lemma about a sequence.*

LEMMA 5.1. *Suppose $\beta > 0$, $p_1, p_2, \dots$ is a sequence of positive numbers with $p_{n+1}/p_n \to 1$, and*

$$\lim_{n\to\infty} \left[\log p_n + \beta \sum_{j=1}^{n} p_j \right]$$

exists and is finite. Then

$$\lim_{n\to\infty} n p_n = 1/\beta.$$

PROOF. It suffices to prove the result for $\beta = 1$, for otherwise we can consider $\tilde{p}_n = \beta p_n$. Let

$$a_n = \log p_n + \sum_{j=1}^{n} p_j.$$

The hypothesis implies that $\{a_n\}$ is a Cauchy sequence.

We first claim that for every $\delta > 0$, there exists $N_\delta > 0$ such that if $n \geq N_\delta$ and $p_n = (1 + 2\epsilon)/n$ with $\epsilon \geq \delta$, then there does not exist $r > n$ with

$$p_k \geq \frac{1}{k}, \quad k = n, \dots, r-1,$$
$$p_r \geq \frac{1 + 3\epsilon}{r}.$$

Indeed, suppose these inequalities hold for some n, r. Then

$$\log(p_r/p_n) \geq \log\frac{1+3\epsilon}{1+2\epsilon} - \log(r/n),$$

$$\sum_{j=n+1}^{r} p_j \geq \log(r/n) - O(n^{-1}),$$

and hence for n sufficiently large,

$$a_r - a_n \geq \frac{1}{2}\log\frac{1+3\epsilon}{1+2\epsilon} \geq \frac{1}{2}\log\frac{1+3\delta}{1+2\delta}.$$

Since a_n is a Cauchy sequence, this cannot be true for large n.

We next claim that for every $\delta > 0$, there exists $N_\delta > 0$ such that if $n \geq N_\delta$ and $p_n = (1+2\epsilon)/n$ with $\epsilon \geq \delta$, then there exists r such that

$$\begin{aligned} \frac{1+\epsilon}{k} &\leq p_k < \frac{1+3\epsilon}{k}, \quad k = n, \dots, r-1, \\ p_r &< \frac{1+\epsilon}{r}. \end{aligned} \tag{5.1}$$

To see this, we consider the first r such that $p_r < \frac{1+\epsilon}{r}$. By the previous claim, if such an r exists, then (5.1) holds for n large enough. If no such r exists, then by the argument above for all $r > n$,

$$a_r - a_n \geq \log\frac{1+\epsilon}{1+2\epsilon} + \frac{\epsilon}{2}\log(r/n) - (1+\epsilon)O(n^{-1}).$$

Since the right-hand side goes to infinity as $r \to \infty$, this contradicts the fact that a_n is a Cauchy sequence.

By iterating the last assertion, we can see that for every $\delta > 0$, there exists $N_\delta > 0$ such that if $n \geq N_\delta$ and $p_n = (1+2\epsilon)/n$ with $\epsilon \geq \delta$, then there exists $r > n$ such that

$$p_r < \frac{1+2\delta}{r}, \quad \text{and} \quad p_k \leq \frac{1+3\epsilon}{k}, \quad k = n, \dots, r-1.$$

Let s be the first index greater than r (if it exists) such that either

$$p_k \leq \frac{1}{k} \quad \text{or} \quad p_k \geq \frac{1+2\delta}{k}.$$

Using $p_{n+1}/p_n \to 1$, we can see, perhaps by choosing a larger N_δ if necessary, that

$$\frac{1-\delta}{k} \leq p_s \leq \frac{1+4\delta}{k}.$$

If $p_s \geq (1+2\delta)/k$, then we can iterate this argument with $\epsilon \leq 2\delta$ to see that

$$\limsup_{n\to\infty} np_n \leq 1 + 6\delta.$$

The lim inf can be done similarly. $\square$

5.2. *Long range intersection.* Let S, W be simple random walks with corresponding stopping times σ_n. We will assume that $S_0 = w$, $W_0 = 0$. Let $\eta = \mathbf{LE}(S[0, \sigma_n])$. Note that we are stopping the random walk S at time σ_n but we are allowing the random walk W to run for infinite time.

PROPOSITION 5.2. *There exists $\alpha < \infty$ such that if $n^{-1} \le e^{-n}|w| \le 1 - n^{-1}$, then*

$$\left|\log \mathbb{P}\{W[0,\infty] \cap \eta \neq \varnothing\} - \log[\hat{G}_n^2(w)\hat{p}_n]\right| \le c\frac{\log^\alpha n}{n}.$$

In particular,

$$\mathbb{E}[H(\eta^n)] = \frac{8\hat{p}_n}{\pi^2}\left[1 + O\left(\frac{\log^\alpha n}{n}\right)\right].$$

Throughout this section, let θ_n denote an error term that decays at least as fast as $\log^\alpha n/n$ for some α (with the implicit uniformity of the estimate over all $n^{-1} \le e^{-n}|w| \le 1 - n^{-1}$). θ_n may vary line by line. Then the second assertion in Proposition 5.2 follows immediately from the first and (2.17). We can write the conclusion of the proposition as

$$\mathbb{P}\{W[0,\infty] \cap \eta \neq \varnothing\} = \hat{G}_n^2(w)\hat{p}_n[1 + \theta_n].$$

Note that $\hat{G}_n^2(w) \ge c/n$ if $|w| \le e^n(1 - n^{-1})$, and hence $\hat{G}_n^2(w)\hat{p}_n \le c/n^2$.

We start by giving the sketch of the proof which is fairly straightforward. In the event $\{W[0,\infty] \cap \eta \neq \varnothing\}$, there are typically many points in $W[0,\infty] \cap \eta$. We focus on a particular one. This is analogous to the situation when one is studying the probability that a random walk visits a set. In the latter case, one usually focuses on the first or last visit. In our case with two paths, the notions of "first" and "last" are a little ambiguous so we have to take some care. We will consider the first point on η that is visited by W and then focus on the last visit by W to this first point on η.

To be precise, we write

$$\begin{aligned}
\eta &= [\eta_0, \dots, \eta_m], \\
i &= \min\{t : \eta_t \in W[0,\infty)\}, \\
\rho &= \max\{t \ge 0 : S_t = \eta_i\}, \\
\lambda &= \max\{t : W_t = \eta_i\}.
\end{aligned}$$

Then the event $\{\rho = j; \lambda = k; S_\rho = W_\lambda = z\}$ is the event that:

I: $j < \sigma_n$, $S_j = z$, $W_k = z$,
II: $\mathbf{LE}(S[0, j]) \cap (S[j+1, \sigma_n] \cup W[0, k] \cup W[k+1, \infty)) = \{z\}$,
III: $z \notin S[j+1, \sigma_n] \cup W[k+1, \infty)$.

Viewing the picture at z, we see that $\mathbb{P}\{\mathbf{II} \text{ and } \mathbf{III}\} \sim \hat{p}_n$. Using the slowly recurrent nature of the random walk paths, we expect that as long as z is not too close to 0, w, or ∂C_n, then **I** is almost independent of (**II** and **III**). This then gives

$$\mathbb{P}\{\rho = j; \lambda = k; S_\rho = W_\lambda = z\} \sim \mathbb{P}\{S_j = W_k = z\}\hat{p}_n,$$

and summing over j, k, z gives

$$\mathbb{P}\{W[0,\infty] \cap \eta \neq \varnothing\} \sim \hat{p}_n \hat{G}_n^2(w).$$

The following proof makes this reasoning precise.

PROOF OF PROPOSITION 5.2. Let V be the event

$$V = \{w \notin S[1, \sigma_n], 0 \notin W[1, \infty)\}.$$

Using $\mathbb{P}[0 \notin W[1,\infty)] = \mathbb{P}^w[w \notin S[1,\infty)] = G(0,0)^{-1}$, we can see that $|\mathbb{P}(V) - G(0,0)^{-2}|$ is fast decaying if $n^{-1} \le e^{-n}|w| \le 1 - n^{-1}$. Let $\tau = \max\{j : W_j = 0\}$. Then $\mathbb{P}\{\tau > \sigma_{\log^2 n}\}$ and $\mathbb{P}\{S[0,\infty) \cap C_{\log^2 n} \neq \varnothing\}$ are fast decaying, and hence so is

$$\big|\mathbb{P}\{W[0,\infty] \cap \eta \neq \varnothing \mid V\} - \mathbb{P}\{W[0,\infty] \cap \eta \neq \varnothing\}\big|.$$

Therefore, it suffices to show that

$$\mathbb{P}\big[V \cap \{W[0,\infty] \cap \eta \neq \varnothing\}\big] = \frac{\hat{G}_n^2(w)}{G(0,0)^2}\hat{p}_n[1+\theta_n]. \tag{5.2}$$

Let $E(j,k,z)$, E_z be the events

$$E(j,k,z) = V \cap \{\rho = j; \lambda = k; S_\rho = W_\lambda = z\}, \qquad E_z = \bigcup_{j,k=0}^{\infty} E(j,k,z).$$

Then $\mathbb{P}[V \cap \{W[0,\infty] \cap \eta \neq \varnothing\}] = \sum_{z \in C_n} \mathbb{P}(E_z)$. Let

$$C_n' = C_{n,w}' = \{z \in C_n : |z| \ge n^{-4}e^n, |z - w| \ge n^{-4}e^n, |z| \le (1 - n^{-4})e^n\}.$$

We can use the easy estimate $\mathbb{P}(E_z) \le G_n(w,z)G(0,z)$ to see that

$$\sum_{z \in C_n \setminus C_n'} \mathbb{P}(E_n) \le O(n^{-6}),$$

so it suffices to estimate $\mathbb{P}(E_z)$ for $z \in C_n'$.

We will translate so that z is the origin and will reverse the paths $S[0,\rho]$ and $W[0,\lambda]$. Let $\omega^1, \dots, \omega^4$ be four independent simple random walks starting at the origin and let $x = w - z$, $y = -z$. Let l^i denote the smallest index l such that $|\omega_l^i - y| \ge e^n$. Using the fact that reverse loop-erasing has the same distribution as forward loop-erasing, we see that $\mathbb{P}[E(j,k,z)]$ can be given as the probability of the following event:

$$\begin{gathered}(\omega^3[1,l^3]\cup\omega^4[1,\infty))\cap\mathbf{LE}(\omega^1[0,j])=\varnothing,\\ \omega^2[0,k]\cap\mathbf{LE}(\omega^1[0,j])=\{0\},\\ j<l^1,\quad \omega^1(j)=x,\quad x\notin\omega^1[0,j-1],\\ \omega^2(k)=y,\quad y\notin\omega^2[0,k-1],\end{gathered}$$

where we translate the time reversal of $S[0,j]$, the time reversal of $W[0,k]$, the path $W[k,\infty)$ and $S[j,\infty)$ into ω_1, ω_2, ω_3, ω_4, respectively. Note that $z\in C_n'$ implies

$$n^{-1}e^n\le|y|,|x-y|,|x|\le e^n[1-n^{-1}].$$

We now rewrite this. We fix x, y and let $C_n^y=y+C_n$. Let W^1, W^2, W^3, W^4 be independent random walks starting at the origin and let

$$\begin{aligned}&T^3=\infty,\quad \text{and}\quad T^i=T_n^i=\min\{j:W_j^i\notin C_n^y\}\quad \text{for } i=1 \text{ and } 4;\\ &\tau^1=\min\{m:W_m^1=x\},\quad \text{and}\quad \tau^2=\min\{m:W_m^2=y\};\\ &\hat\Gamma=\hat\Gamma_n=\mathbf{LE}(W^1[0,\tau^1]).\end{aligned}$$

We also override the notation S to denote an simple random walk on $\mathbb{Z}^4$ starting from 0. Let E be the event

$$\hat\Gamma\cap(W^2[1,\tau^2]\cap W^3[1,T^3])=\varnothing\quad\text{and}\quad \hat\Gamma\cap W^4[0,T^4]=\{0\}.$$

Then $\mathbb{P}\{E,\tau^1<T^1,\tau^2<\infty\}$ equals $\mathbb{P}\{V\cap W[0,\infty)\cap\eta\}$ in (5.2). Note that

$$\begin{aligned}\mathbb{P}\{\tau^1<T^1\}&=\frac{G_n(w,z)}{G_n(w,w)}=\frac{G_n(w,z)}{G(0,0)}+o(e^{-n}),\\ \mathbb{P}\{\tau^2<\infty\}&=\frac{G(0,y)}{G(y,y)}=\frac{G(0,z)}{G(0,0)}.\end{aligned}$$

Therefore, in order to prove (5.2), it suffices to prove that

$$p_n'(x,y):=\mathbb{P}\{E\mid\tau^1<T^1,\tau^2<\infty\}=\hat p_n[1+\theta_n]. \tag{5.3}$$

We write Q for the distribution of W_1, W_2, W_3, W_4 under the conditioning $\{\tau^1<T^1,\tau^2<\infty\}$ in (5.3). Then consider two events E_1, E_2 as follows. Let $\hat W=W^2[1,\tau^2]\cup W^3[1,T^3]\cup W^4[0,T^4]$ and let E_0, E_1, E_2 be the events

$$\begin{aligned}E_0&=\{0\notin W^2[1,\tau^2]\cup W^3[1,T^3]\},\\ E_1&=E_0\cap\{\hat W\cap\hat\Gamma\cap C_{n-\log^3 n}=\{0\}\},\\ E_2&=E_1\cap\{\hat W\cap\Theta_n=\varnothing\},\end{aligned}$$

where $\Theta_n = W^1[0, \tau^1] \cap A(n - \log^3 n, 2n)$. Since Q-almost surely

$$\hat{\Gamma} \cap C_{n-\log^3 n} \subset \hat{\Gamma} \subset \Theta_n \cup (\hat{\Gamma} \cap C_{n-\log^3 n}).$$

We have

$$Q(E_2) \le p'_n(x, y) \le Q(E_1).$$

Now to prove (5.3). it suffices to show

$$Q(E_1) = \hat{p}_n[1 + \theta_n], \tag{5.4}$$

$$Q(E_1 \setminus E_2) \le n^{-1}\theta_n. \tag{5.5}$$

For each $z \in \mathbb{Z}^4$ let

$$\phi_x(z) = \phi_{x,n}(z) = \mathbb{P}^z\{\tau^1 < T_n^1\} \quad \text{and} \quad \phi_y(z) = \mathbb{P}^z\{\tau^2 < \infty\}.$$

Therefore, for any path $\omega = [0, \omega^1, \dots, \omega^m]$ with $0, \omega^1, \dots, \omega^{m-1} \in C_n^y \setminus \{x\}$,

$$Q\{[W_0^1, \dots, W_m^1] = \omega\} = \mathbb{P}\{[S_0, \dots, S_m] = \omega\} \frac{\phi_x(\omega^m)}{\phi_x(0)}. \tag{5.6}$$

Similarly, if $\omega = [0, \omega^1, \dots, \omega^m]$ is a path with $y \notin \{0, \omega^1, \dots, \omega^{m-1}\}$, then

$$Q\{[W_0^2, \dots, W_m^2] = \omega\} = \mathbb{P}\{[S_0, \dots, S_m] = \omega\} \frac{\phi_y(\omega^m)}{\phi_y(0)}.$$

Let $\zeta \in \{x, y\}$. By (2.5), there exists a fast decaying sequence ϵ_n such that

$$\phi_\zeta(z) = \phi_\zeta(0)[1 + O(\epsilon_n)] \quad \text{if } |z| \le e^n e^{-\log^2 n}. \tag{5.7}$$

This implies that W^i, S $(i = 1, 2)$ can be coupled on the same probability space such that, except on an event of probability $O(\epsilon_n)$,

$$\hat{\Gamma} \cap C_{n-\log^3 n} = \mathbf{LE}(S[0, \infty)) \cap C_{n-\log^3 n}.$$

Therefore, $Q(E_1) - p_{n-\log n^3 n}$ is fast decaying and hence (5.4) follows from Proposition 1.6.

To prove (5.5), consider the following events whose union covers $E_1 \setminus E_2$:

$$F^2 = E_1 \cap \{W^2[1, \tau^2] \cap \Theta_n \ne \varnothing\}, \qquad F^3 = E_1 \cap \{W^3[1, T^3] \cap \Theta_n \ne \varnothing\},$$
$$F^4 = E_1 \cap \{W^4[1, T^4] \cap \Theta_n \ne \varnothing\}.$$

We are now going to prove $Q(F^i) \le n^{-1}\theta_n$ for $i = 2, 3, 4$, thus proving (5.5).

By Proposition 2.5, we can find a fast decaying sequence δ_n such that

$$\mathbb{P}\left\{H(S[0, \infty) \cap A(n - 1, n)) \ge \frac{\log^2 n}{n}\right\} \le \delta_n^{100}.$$

Let $\rho = \rho_n = \min\{j : |W_j^1 - x| \le e^n \delta_n\}$. Using the strong Markov property of W^1 under Q, we can find a constant $\alpha > 0$ such that

$$Q\{W^1[\rho, \tau^1] \notin \{|z - x| \le e^n \sqrt{\delta_n}\}\} = O(\delta_n^\alpha).$$

Since $|x| \ge e^{-n} n^{-1}$, we know that

$$H(\{|z - x| \le e^n \sqrt{\delta_n}\}) \le n^2 \delta_n.$$

By the strong Markov property, for all $w \in C_n^y$ with $|w - x| \ge e^n \sqrt{\delta_n}$,

$$\phi(0)/\phi(w) \gtrsim \mathbb{P}[\rho_n < T^1] \ge \delta_n^{50}.$$

Using (5.6), we have

$$Q\left\{H(W^1[0, \tau^1] \cap A(n-1, n)) \ge \frac{\log^2 n}{n}\right\} \le \frac{O(\delta_n^{100})}{\mathbb{P}[\rho_n < T^1]} + O(\delta_n^\alpha), \tag{5.8}$$

which is fast decaying. Therefore, $Q[W^1[0, \tau^1] \notin \mathcal{X}_n]$ is fast decaying. Let

$$R_n = H(W^1[0, \tau^1] \cap A(n - 2\log^3 n, n+1)). \tag{5.9}$$

Let $\sigma = \inf\{m : W_m^3 \notin C_{n-\log^3 n}\}$ and $E = \{W^3[1, \sigma] \cap \hat{\Gamma} = \varnothing\}$. Let $\bar{Q}$ be the probability measure conditioning on W^1, W^2, W^4. Then

$$\bar{Q}[F^3] \le \bar{Q}(E) \sum_{z \in \partial C_{n-\log^3 n}} \bar{Q}[S_\sigma = z \mid E] Q[W^3[\sigma, T^3] \cap \hat{\Gamma} = \varnothing \mid W_\sigma^3 = z].$$

On the event $\hat{\Gamma} \in \tilde{\mathcal{X}}_{n-\log^3 n}$, by Proposition 3.8, there exists $c < \infty$ such that

$$\begin{aligned}
&\sum_{z \in \partial C_{n-\log^3 n}} \bar{Q}[W_\sigma^3 = z \mid E] \bar{Q}[W^3[\sigma, T^3] \cap \hat{\Gamma} \ne \varnothing \mid W_\sigma^3 = z] \\
&\quad \le c\bar{Q}[W^3[\sigma, T^3] \cap \hat{\Gamma} \ne \varnothing] \\
&\quad \le c(H(\hat{\Gamma} \cap A(n - 2\log^3 n, n+1)) + Q[W^3[\sigma, \infty) \cap C_{n-2\log^3 n} \ne \varnothing]) \\
&\quad \le c(R_n + Q[W^3[\sigma, \infty) \cap C_{n-2\log^3 n} \ne \varnothing]) \le cR_n + O(e^{-\log^3 n}).
\end{aligned}$$

Note that $\mathrm{Es}(\hat{\Gamma}) \le 2^{-(n-\log^3 n)/4}$ when $\hat{\Gamma} \in \tilde{\mathcal{X}}_n$. Moreover, $Q[\hat{\Gamma}_n \notin \mathcal{X}_n]$ is fast decaying. Applying the union bound to (5.8), we have $Q[R_n \ge \log^7 n/n]$ is fast decaying. Averaging over W^1, W^2, W^4, we have

$$Q(F^3) \le O(\log^7 n/n) Q(E_1) + \epsilon_n,$$

where ϵ_n is fast decaying. This gives $Q(F^3) \le n^{-1}\theta_n$ for some θ_n. The same argument shows $Q(F^4) \le n^{-1}\theta_n$

We still need a similar result to conclude $Q(F^2) \le n^{-1}\theta_n$. By (5.7), we can couple an usual simple random walk S with W^2 (under the Q-probability) such that S and W^2 agree until $\inf\{m : S_m \notin C_{n-\log^3 n}\}$ except for an event of fast decaying probability. The same argument as above reduces proving $Q(F^2) \le n^{-1}\theta_n$ to showing that there exists θ_n such that except for an event of fast decaying Q-probability,

$$Q\{W^2[0,\tau^2] \cap (\hat{\Gamma} \cap A(n - 2\log^3 n, n+1)) \neq \varnothing \mid \hat{\Gamma}\} \le \theta_n$$

which follows from a similar argument for the bound for $Q[R_n \ge \log^7 n/n]$ above. □

As explained in Section 1.2, combined with Lemma 5.1, we conclude the proof of Theorem 1.5. Inserting $\hat{p}_n \sim \pi^2/24$ back to Proposition 5.2, we get the following.

COROLLARY 5.3. *If $n^{-1} \le e^{-n}|w| \le 1 - n^{-1}$, then*

$$\mathbb{P}\{W[0,\infty) \cap \eta \neq \varnothing\} \sim \frac{\pi^2 \hat{G}_n^2(w)}{24n}.$$

More precisely,

$$\lim_{n\to\infty} \max_{n^{-1} \le e^{-n}|w| \le 1-n^{-1}} \left|24n\mathbb{P}\{W[0,\infty) \cap \eta \neq \varnothing\} - \pi^2 \hat{G}_n^2(w)\right| = 0.$$

By a very similar proof, one can show the following variant of Proposition 5.2 that implies Theorem 1.3.

PROPOSITION 5.4. *In the setting of Theorem 1.3, there exists $\alpha < \infty$ such that if $n^{-1} \le e^{-n}|w| \le 1 - n^{-1}$, then*

$$\left|\log \mathbb{P}\{W[0,\sigma_n^W] \cap \mathbf{LE}(S[0,\sigma_n]) \neq \varnothing\} - \log[G_n^2(w)\hat{p}_n]\right| \le c\frac{\log^\alpha n}{n}.$$

6. Two-sided loop-erased random walk. We start by completing the proof of Theorem 1.1. Let $\tilde{Z}_n = \mathrm{Es}(\Gamma; n)$. We first note that the limit $\lim_{n\to\infty} n^{1/3}\tilde{Z}_n$ exists almost surely. We only need consider $\Gamma \in \tilde{\mathcal{X}}$ since otherwise the limit is 0 almost surely. In this case, existence is established by Corollary 3.10, Proposition 1.8 and the fact that $3nh_n \sim 1$. Recall (3.13). We have

$$\mathbb{E}[\tilde{Z}_n^r] = \mathbb{E}[Z_n^r][1 + O(\log^4 n/n)].$$

Since for each r, $\mathbb{E}[n^{r/3}\tilde{Z}_n^r]$ is uniformly bounded, we also get the limit in L^p for each $p > 0$. We then get Theorem 1.1 using (1.2).

If η is an infinite (one-sided) path starting at the origin, and W is a simple random walk with stopping times $\sigma_n^W = \inf\{W_j \notin C_n\}$, we define

$$\phi_\eta(x) = \lim_{n\to\infty} n^{1/3}\mathbb{P}^x\{W[0,\sigma_n^W] \cap \eta = \varnothing\},$$

$$\nabla\phi_\eta = \frac{1}{8}\sum_{|y|=1} \phi_\eta(y).$$

We let $\mathcal{K}$ denote the set of infinite self-avoiding paths starting at the origin such that the limit above exists and is finite for all $x \in \mathbb{Z}^4$ and let $\mathcal{K}^+$ be the set of such η with $\nabla\phi_\eta > 0$. We can restate Theorem 1.1 as follows: if $\eta = \mathbf{LE}(S[0,\infty))$, then with probability one, $\eta \in \mathcal{K}$ and with positive probability $\eta \in \mathcal{K}^+$. Moreover, $\nabla\phi_\eta$ is in L^p for all p.

We can now construct the two-sided loop-erased random walk in $\mathbb{Z}^4$. This is a measure on doubly-infinite self-avoiding paths

$$\omega = [\cdots \omega_{-2}, \omega_{-1}, \omega_0, \omega_1, \omega_2 \cdots]$$

with $\omega_0 = 0$. Each ω can be described in terms of the sequence of one-sided paths

$$\eta^j = [\omega_j, \omega_{j-1}, \omega_{j-2}, \dots] - \omega_j, \quad j \in \mathbb{N},$$

where η^{j+1} is obtained from η^j by choosing $|x| = 1$, attaching x to the beginning of η^j, and translating by $-x$. The transition probabilities for the chain are specified by saying that if $\eta^j = \eta$, then x is chosen with probability $\phi_\eta(x)/\nabla\phi_\eta$. We choose $\eta^0 = \eta = \mathbf{LE}(S[0,\infty))$ tilting by $\nabla\phi_\eta/\mathbb{E}[\nabla\phi_\eta]$.

We give another definition. Suppose

$$\eta = [\eta_{-k}, \eta_{-k-1}, \dots, \eta_{j-1}, \eta_j]$$

is a (finite) self-avoiding walk in $\mathbb{Z}^4$. Let S, W be independent random walk starting at $z = \eta_{-k}$, $w = \eta_j$, respectively, with corresponding stopping times σ_n^S, σ_n^W, respectively, and let E_n be the event

$$\begin{gathered} S[1, \sigma_n^S] \cap \eta = \varnothing, \qquad W[1, \sigma_n^W] \cap \eta = \varnothing, \\ \mathbf{LE}(S[1, \sigma_n^S]) \cap W[1, \sigma_n^W] = \varnothing. \end{gathered} \tag{6.1}$$

We define

$$\mathrm{Es}(\eta) = \lim_{n\to\infty} n^{1/3} \mathbb{P}^{z,w}[E_n]. \tag{6.2}$$

It follows from our theorems that the limit exists. Moreover (see [12], Chapter 9), the limit would be the same if condition (6.1) in the event E_n is replaced by

$$S[1, \sigma_n^S] \cap \mathbf{LE}(W[1, \sigma_n^W]) = \varnothing.$$

From this, we see that $\mathrm{Es}(\eta)$ is translation invariant and also invariant under path reversal. The probability that the two-sided loop-erased walk produces η is

$$8^{-(j+k)} F_\eta \frac{\mathrm{Es}(\eta)}{\mathrm{Es}(0)}.$$

Here $\mathrm{Es}(0)$ is the quantity where η is the trivial walk $[0]$ and F_η is a loop term that can be given in several ways, for example,

$$F_\eta = \prod_{i=-k}^{j} G_{A_i}(\eta_i, \eta_i),$$

where $A_i = \mathbb{Z}^4 \setminus \{\eta_{-k}, \dots, \eta_{i-1}\}$. Although not immediately obvious, this quantity depends only on the set $\{\eta_{-k}, \dots, \eta_j\}$ and not on the ordering of the points.

- For $d \geq 5$, it was constructed in [5], where one can define ϕ_η by
$$\mathbb{P}\{W[0,\infty) \cap \eta = \varnothing\}.$$
In this case, the marginal distribution on the past or future of the path is absolutely continuous with respect to the one-sided measure with a bounded Radon–Nikodym derivative.
- For $d = 4$, it is absolutely continuous with an L^p, but not a uniformly bounded, derivative.
- In [11], the two-sided walk is constructed for $d = 2, 3$ using (6.2) replacing $n^{1/3}$ with a sequence a_n. (The proof there also works for $d > 3$ but does not give as strong a result as above.) If $d = 2$, it is known that $a_n = e^{3n/4}$ works; for $d = 3$, it is expected that we can choose $a_n = e^{\beta n}$ for an appropriate β but this has not been proven.

7. Gaussian limits for the spin field. In this section, we start by reviewing some known facts of UST and random walk Green's function, then proving Theorem 1.11 by applying main estimates of LERW.

7.1. *Uniform spanning trees.* Here we review some facts about the uniform spanning forest (i.e., wired spanning trees) on finite subsets of $\mathbb{Z}^d$ on $\mathbb{Z}^d$. Most of the facts extend to general graphs as well. For more details, see [12], Chapter 9.

Given a finite subset $A \subset \mathbb{Z}^d$, the uniform wired spanning tree in A is a subgraph of the graph $A \cup \{\partial A\}$, choosing uniformly random among all spanning trees of $A \cup \{\partial A\}$. (A spanning tree $\mathcal{T}$ is a subgraph such that any two vertices in $\mathcal{T}$ are connected by a unique simple path in $\mathcal{T}$.) We define the *uniform spanning forest (USF)* on A to be the uniform wired spanning tree restricted to the edges in A. One can also consider the uniform spanning forest on all of $\mathbb{Z}^d$ [2, 14], but we will not need this construction.

The uniform wired spanning tree, and hence the USF, on A can be generated by Wilson's algorithm [16] which we recall here:

- Order the elements of $A = \{x_1, \dots, x_k\}$.
- Start an SRW at x_1 and stop it when in reaches ∂A giving a nearest neighbor path ω. Erase the loops chronologically to produce $\eta = \mathbf{LE}(\omega)$. Add all the edges of η to the tree which now gives a tree $\mathcal{T}_1$ on a subset of $A \cup \{\partial A\}$ that includes ∂A.
- Choose the vertex of smallest index that has not been included and run a simple random walk until it reaches a vertex in $\mathcal{T}_1$. Erase the loops and add the new edges to $\mathcal{T}_1$ in order to produce a tree $\mathcal{T}_2$.
- Continue until all vertices are included in the tree.

Wilson's theorem states that the distribution of the tree is independent of the order in which the vertices were chosen and is uniform among all spanning trees. In particular, we get the following:

- If $x, y \in A$, let S^x, S^y be two independent SRWs starting from x, y, respectively. Then the probability that x, y are in the same component of the USF equals to $\mathbb{P}\{\mathbf{LE}(\omega^x) \cap \omega^y = \varnothing\}$.

Using this characterization, we can see the three regimes for the dimension d. Let us first consider the probabilities that neighboring points are in the same component. Let q_N be the probability that a nearest neighbor of the origin is in a different component as the origin when $A = A_N$. Then

$$q_\infty := \lim_{N\to\infty} q_N > 0, \quad d \geq 5,$$

$$q_N \asymp (\log N)^{-1/3}, \quad d = 4.$$

For $d < 4$, q_N decays like a power of N. For far away points, we have:

- If $d > 4$, and $|x| = n$, the probability that 0 and x are in the same component is comparable to $|x|^{4-d}$. This is true even if $N = \infty$.
- If $d = 4$ and $|x| = n$, the probability that that 0 and x are in the same component is comparable to $1/\log n$. However, if we chose $N = \infty$, the probability would equal to one.

The last fact can be used to show that the USF in all of $\mathbb{Z}^4$ is, in fact, a tree. For $d < 4$, the probability that 0 and x are in the same component is asymptotic to 1 and our construction is not interesting. This is why we restrict to $d \geq 4$.

7.2. *Proof of Theorem* 1.11. Here we give the proof of the theorem by applying Theorem 1.5 and Proposition 5.4. We will only consider the $d = 4$ case here. It suffices to prove the result for $m = 1$, $h_1 = h$, as the general result follows by applying the $k = 1$ result to any linear combination of $h_1, \dots, h_m$.

We fix $h \in C_0^\infty$ with $\int h = 0$ and allow implicit constants to depend on h. We will write just Y_x for $Y_{x,n}$. Let K be such that $h(x) = 0$ for $|x| > K$.

Let us write $L_n = n^{-1}\mathbb{Z}^4 \cap \{|x| \leq K\}$ and $a_n = \sqrt{3\log n}$. Let

$$\langle h, \phi_n\rangle = n^{-2} a_n \sum_{x\in L_n} h(x) Y_{nx} = n^{-2} a_n \sum_{nx\in A_{nK}} h(x) Y_{nx}.$$

Let $q_N(x, y)$ be the probability that x, y are in the same component of the USF on A_N. Note that

$$\mathbb{E}[Y_{x,n} Y_{y,n}] = q_N(x, y),$$

$$\mathbb{E}\big[\langle h, \phi_n\rangle^2\big] = n^{-4} \sum_{x\in L_n} h(x)h(y) a_n^2 q_N(nx, ny).$$

To estimate $\mathbb{E}[\langle h, \phi_n\rangle^2]$, we need the follow two lemmas.

Let $G_N(x, y)$ denote the usual random walk Green's function on A_N, and

$$G_N^2(x, y) = \sum_{z\in A_N} G_N(x, z) G_N(z, y).$$

Note that here the meaning of G_N is not the same as G_n in Section 2.4 with $n = N$ as $A_N \neq C_N$.

LEMMA 7.1. *There exists $c_0 \in \mathbb{R}$ such that if $|x|, |y|, |x-y| \leq N/2$, then*

$$G_N^2(x,y) = \frac{8}{\pi^2}\log\left[\frac{N}{|x-y|}\right] + c_0 + O\left(\frac{|x|+|y|+1}{N} + \frac{1}{|x-y|}\right).$$

PROOF. Let $\delta = N^{-1}[1+|x|+|y|]$ and note that $|x-y| \leq \delta N$. Since

$$\begin{aligned}\sum_{|w|<N(1-\delta)} G_N(x-y,w)G_N(0,w) &\leq \sum_{w\in A_N} G_N(x,w)G_N(y,w)\\ &\leq \sum_{|w|\leq N(1+\delta)} G_N(x-y,w)G_N(0,w),\end{aligned}$$

Lemma 7.1 is reduced to the case $y = 0$, which then follows from Lemmas 2.8 and 2.9. □

LEMMA 7.2. *There exists a sequence r_n with $r_n \leq O(\log\log n)$, a sequence $\epsilon_n \downarrow 0$, such that if $x, y \in L_n$ with $|x-y| \geq 1/\sqrt{n}$,*

$$\left|a_n^2 q_N(nx,ny) - r_n + \log|x-y|\right| \leq \epsilon_n.$$

PROOF. In light of Wilson's algorithm, Lemma 2.9 and Proposition 5.4 yield Lemma 7.2 in the $y = 0$ case. The general case can be reduced to the $y = 0$ case as in Lemma 7.1 by recentering. □

An upper bound for $q_N(x,y)$ can be given in terms of the probability that the paths of two independent random walks starting at x, y, stopped when they leave A_N, intersect. This gives

$$q_N(x,y) \leq \frac{c\log[N/|x-y|]}{\log N}.$$

Let $\delta_n = \exp\{-(\log\log n)^2\}$, which is a function that decays faster than any power of $\log n$. Then

$$q_N(x,y) \leq c\frac{(\log\log n)^2}{\log n}, \quad |x-y| \geq n\delta_n. \tag{7.1}$$

It follows from Lemma 7.2, (7.1) and the trivial inequality $q_N \leq 1$, that

$$\begin{aligned}\mathbb{E}\left[\langle h,\phi_n\rangle^2\right] &= o(1) + n^{-4}\sum_{x,y\in L_n}\left[r_n - \log|x-y|\right]h(x)h(y)\\ &= o(1) - n^{-4}\sum_{x,y\in L_n}\log|x-y|h(x)h(y)\\ &= o(1) - \int h(x)h(y)\log|x-y|\,dx\,dy,\end{aligned}$$

which shows that the second moment has the correct limit. The second equality uses $\int h = 0$ to conclude that

$$\frac{r_n}{n^4} \sum_{x,y \in L_n} h(x)h(y) = o(1).$$

We now consider the higher moments. It is immediate from the construction that the odd moments of $\langle h, \phi_n \rangle$ are identically zero, so it suffices to consider the even moments $\mathbb{E}[\langle h, \phi_n \rangle^{2k}]$. We fix $k \geq 1$ and allow implicit constants to depend on k as well. Let $L_* = L_{n,*}^{2k}$ be the set of $\bar{x} = (x_1, \dots, x_{2k}) \in L_n^{2k}$ such that $|x_j| \leq K$ for all j and $|x_i - x_j| \geq \delta_n$ for each $i \neq j$. We write $h(\bar{x}) = h(x_1) \dots h(x_{2k})$.

Note that $\# L_* \asymp n^{8k}$ and $\#(L_n^{2k} \setminus L_*) \asymp k^2 n^{8k} \delta_n$. In particular,

$$n^{-8k} a_n^{2k} \sum_{\bar{x} \in L_{n,*}^{2k}} h(x_1)h(x_2) \cdots h(x_{2k}) = o_{2k}(\sqrt{\delta_n}).$$

Then we see that

$$\begin{aligned} \mathbb{E}\bigl[\langle h, \phi_n \rangle^{2k}\bigr] &= n^{-8k} a_n^{2k} \sum_{\bar{x} \in L_n^{2k}} h(\bar{x}) \mathbb{E}[Y_{nx_1} \cdots Y_{nx_{2k}}] \\ &= O(\sqrt{\delta_n}) + n^{-8k} a_n^{2k} \sum_{\bar{x} \in L_*} h(\bar{x}) \mathbb{E}[Y_{nx_1} \cdots Y_{nx_{2k}}]. \end{aligned}$$

LEMMA 7.3. *For each k, there exists $c < \infty$ such that the following holds. Suppose $\bar{x} \in L_{n,*}^{2k}$ and let $\omega^1, \dots, \omega^{2k}$ be independent simple random walks started at $nx_1, \dots, nx_{2k}$ stopped when they reach ∂A_N. Let N denote the number of integers $j \in \{2, 3, \dots, 2k\}$ such that*

$$\omega^j \cap (\omega^1 \cup \cdots \cup \omega^{j-1}) \neq \varnothing.$$

Then

$$\mathbb{P}\{N \geq k+1\} \leq c \left[\frac{(\log \log n)^3}{\log n} \right]^{k+1}.$$

Conditioned on Lemma 7.3, we now prove Theorem 1.11 by verifying Wick's formula. We write $y_j = nx_j$ and write Y_j for Y_{y_j}. To calculate $\mathbb{E}[Y_1 \cdots Y_{2k}]$, we first sample our USF which gives a random partition $\mathcal{P}$ of $\{y_1, \dots, y_{2k}\}$. Note that $\mathbb{E}[Y_1 \cdots Y_{2k} \mid \mathcal{P}]$ equals 1 if it is an "even" partition in the sense that each set has an even number of elements. Otherwise, $\mathbb{E}[Y_1 \cdots Y_{2k} \mid \mathcal{P}] = 0$. Any even partition, other than a partition into k sets of cardinality 2, will have $N \geq k+1$. Hence

$$\mathbb{E}[Y_1 \cdots Y_{2k}] = O\left(\left[\frac{(\log \log n)^3}{\log n} \right]^{k+1} \right) + \sum \mathbb{P}(\mathcal{P}_{\bar{y}}),$$

where the sum is over the $(2k-1)!!$ perfect matchings of $\{1, 2, \dots, 2k\}$ and $\mathbb{P}(\mathcal{P}_{\bar{y}})$ denotes the probability of getting this matching for the USF for the vertices $y_1, \dots, y_{2k}$.

Let us consider one of these perfect matchings that for convenience we will assume is $y_1 \leftrightarrow y_2, y_3 \leftrightarrow y_4, \dots, y_{2k-1} \leftrightarrow y_{2k}$. We claim that

$$\begin{aligned}
&\mathbb{P}(y_1 \leftrightarrow y_2, y_3 \leftrightarrow y_4, \dots, y_{2k-1} \leftrightarrow y_{2k}) \\
&\quad = O([\frac{(\log\log n)^3}{\log n}]^{k+1}) + \mathbb{P}(y_1 \leftrightarrow y_2)\mathbb{P}(y_3 \leftrightarrow y_4)\cdots\mathbb{P}(y_{2k-1} \leftrightarrow y_{2k}).
\end{aligned}$$

Indeed, this is just inclusion–exclusion using our estimate on $\mathbb{P}\{N \geq k+1\}$.

If we write $\epsilon_n = \epsilon_{n,k} = (\log\log n)^{3(k+1)}/\log n$, we now see from symmetry that

$$\begin{aligned}
&\mathbb{E}[\langle h, \phi_n\rangle^{2k}\rangle] \\
&\quad = O(\epsilon_n) + n^{-8k} a_n (2k-1)!! \sum_{\bar{x} \in L_*} \mathbb{P}\{nx_1 \leftrightarrow nx_2, \dots, nx_{2k-1} \leftrightarrow nx_{2k}\} \\
&\quad = O(\epsilon_n) + (2k-1)!! \big[\mathbb{E}\big(\langle h, \phi_n\rangle^2\big)\big]^k.
\end{aligned}$$

7.3. *Proof of Lemma* 7.3. Here we fix k and let $y_1, \dots, y_{2k}$ be points with $|y_j| \leq Kn$ and $|y_i - y_j| \geq n\delta_n$ where we recall $\log\delta_n = -(\log\log n)^2$. Let $\omega^1, \dots, \omega^{2k}$ be independent simple random walks starting at y_j stopped when they get to ∂A_N. We let $E_{i,j}$ denote the event that $\omega^i \cap \omega^j \neq \varnothing$ and let $R_{i,j} = \mathbb{P}(E_{i,j} \mid \omega^j)$.

LEMMA 7.4. *There exists $c < \infty$ such that for all i, j and all n sufficiently large*,

$$\mathbb{P}\left\{R_{i,j} \geq c\frac{(\log\log n)^3}{\log n}\right\} \leq \frac{1}{(\log n)^{4k}}.$$

PROOF. We know that there exists $c < \infty$ such that if $|y - z| \geq n\delta_n^2$, then the probability that simple random walks starting at y, z stopped when they reach ∂A_N intersect is at most $c(\log\log n)^2/\log n$. Hence there exists c_1 such that

$$\mathbb{P}\left\{R_{i,j} \leq \frac{c_1(\log\log n)^2}{\log n}\right\} \geq \frac{1}{2}. \tag{7.2}$$

Start a random walk at z and run it until one of three things happens:

- It reaches ∂A_N
- It gets within distance $n\delta_n^2$ of y
- The path is such that the probability that a simple random walk starting at y intersects the path before reaching ∂A_N is greater than $c_1(\log\log n)^2/\log n$.

If the third option occurs, then we restart the walk at the current site and do this operation again. Eventually, one of the first two options will occur. Suppose it takes r trials of this process until one of the first two events occur. Then either $R_{i,j} \leq$

$rc_1(\log\log n)^2/\log n$ or the original path starting at z gets within distance δ_n^2 of y. The latter event occurs with probability $O(\delta_n) = o((\log n)^{-4k})$. Also, using (7.2), we can see the probability that it took at least r steps is bounded by $(1/2)^r$. By choosing $r = c_2 \log\log n$, we can make this probability less than $1/(\log n)^{4k}$. $\square$

PROOF OF LEMMA 7.3. Let R be the maximum of $R_{i,j}$ over all $i \neq j$ in $\{1, \dots, 2k\}$. Then, at least for n sufficiently large,

$$\mathbb{P}\left\{R \geq c\frac{(\log\log n)^3}{\log n}\right\} \leq \frac{1}{(\log n)^{3k}}.$$

Let

$$E_{\cdot,j} = \bigcup_{i=1}^{j-1} E_{i,j}.$$

On the event $R < c(\log\log n)^3/\log n$, we have

$$\mathbb{P}\{E_{\cdot,j} \mid \omega^1, \dots, \omega^{j-1}\} \leq \frac{c(j-1)(\log\log n)^3}{\log n}.$$

If N denotes the number of j for which $E_{\cdot,j}$ occurs, we see that

$$\mathbb{P}\{N \geq k+1\} \leq c\left[\frac{(\log\log n)^3}{\log n}\right]^{k+1}. \qquad \square$$

Acknowledgments. We are grateful to Scott Sheffield for motivating this problem and helpful discussions. We thank Yuval Peres, Robin Pemantle and an anonymous referee for useful comments on an earlier draft. We thank the Isaac Newton Institute for Mathematical Sciences, Cambridge, for support and hospitality during the programme *Random Geometry* where part of this project was undertaken.

REFERENCES

[1] AIZENMAN, M. (1982). Geometric analysis of φ^4 fields and Ising models. I, II. *Comm. Math. Phys.* **86** 1–48. MR0678000

[2] BENJAMINI, I., LYONS, R., PERES, Y. and SCHRAMM, O. (2001). Uniform spanning forests. *Ann. Probab.* **29** 1–65. MR1825141

[3] GRIMMETT, G. (2006). *The Random-Cluster Model. Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **333**. Springer, Berlin. MR2243761

[4] LAWLER, G. F. (1980). A self-avoiding random walk. *Duke Math. J.* **47** 655–693. MR0587173

[5] LAWLER, G. F. (1983). A connective constant for loop-erased self-avoiding random walk. *J. Appl. Probab.* **20** 264–276. MR0698530

[6] LAWLER, G. F. (1985). Intersections of random walks in four dimensions. II. *Comm. Math. Phys.* **97** 583–594. MR0787120

[7] LAWLER, G. F. (1986). Gaussian behavior of loop-erased self-avoiding random walk in four dimensions. *Duke Math*. *J*. **53** 249–269. MR0835809

[8] LAWLER, G. F. (1992). Escape probabilities for slowly recurrent sets. *Probab. Theory Related Fields* **94** 91–117. MR1189088

[9] LAWLER, G. F. (1995). The logarithmic correction for loop-erased walk in four dimensions. In *Proceedings of the Conference in Honor of Jean-Pierre Kahane* (*Orsay*, 1993) 347–361. *J. Fourier Anal. Appl.*, Special Issue. MR1364896

[10] LAWLER, G. F. (2013). *Intersections of Random Walks*. *Modern Birkhäuser Classics*. Birkhäuser/Springer, New York. Reprint of the 1996 edition. MR2985195

[11] LAWLER, G. F. (2018). The infinite two-sided loop-erased random walk. Available at arXiv:1802.06667.

[12] LAWLER, G. F. and LIMIC, V. (2010). *Random Walk*: *A Modern Introduction*. *Cambridge Studies in Advanced Mathematics* **123**. Cambridge Univ. Press, Cambridge. MR2677157

[13] LODHIA, A., SHEFFIELD, S., SUN, X. and WATSON, S. S. (2016). Fractional Gaussian fields: A survey. *Probab. Surv.* **13** 1–56. MR3466837

[14] PEMANTLE, R. (1991). Choosing a spanning tree for the integer lattice uniformly. *Ann*. *Probab.* **19** 1559–1574. MR1127715

[15] SUN, X. and WU, W. (2013). Uniform spanning forests and the bi-Laplacian Gaussian field. Available at arXiv:1312.0059.

[16] WILSON, D. B. (1996). Generating random spanning trees more quickly than the cover time. In *Proceedings of the Twenty-Eighth Annual ACM Symposium on the Theory of Computing* (*Philadelphia*, *PA*, 1996) 296–303. ACM, New York. MR1427525

G. LAWLER
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF CHICAGO
5734 S. UNIVERSITY AVENUE
CHICAGO, ILLINOIS 60637
USA
E-MAIL: lawler@math.uchicago.edu

X. SUN
DEPARTMENT OF MATHEMATICS
COLUMBIA UNIVERSITY
2990 BROADWAY
NEW YORK, NEW YORK 10027
USA
E-MAIL: xinsun@math.columbia.edu

W. WU
STATISTICS DEPARTMENT
UNIVERSITY OF WARWICK
COVENTRY CV4 7AL
UNITED KINGDOM
E-MAIL: w.wu.9@warwick.ac.uk