The Annals of Probability
2012, Vol. 40, No. 2, 535–577
DOI: 10.1214/10-AOP624

UNIFORMITY OF THE UNCOVERED SET OF RANDOM WALK AND CUTOFF FOR LAMPLIGHTER CHAINS

BY JASON MILLER AND YUVAL PERES

Stanford University and Microsoft Research

We show that the measure on markings of $\mathbf{Z}_n^d$, $d \geq 3$, with elements of $\{0, 1\}$ given by i.i.d. fair coin flips on the range $\mathcal{R}$ of a random walk X run until time T and 0 otherwise becomes indistinguishable from the uniform measure on such markings at the threshold $T = \frac{1}{2} T_{\mathrm{cov}}(\mathbf{Z}_n^d)$. As a consequence of our methods, we show that the total variation mixing time of the random walk on the lamplighter graph $\mathbf{Z}_2 \wr \mathbf{Z}_n^d$, $d \geq 3$, has a cutoff with threshold $\frac{1}{2} T_{\mathrm{cov}}(\mathbf{Z}_n^d)$. We give a general criterion under which both of these results hold; other examples for which this applies include bounded degree expander families, the intersection of an infinite supercritical percolation cluster with an increasing family of balls, the hypercube and the Caley graph of the symmetric group generated by transpositions. The proof also yields precise asymptotics for the decay of correlation in the uncovered set.

1. Introduction. Suppose $G = (V, E)$ is a finite, connected graph and X is a lazy random walk on G. This means that X is the Markov chain with state space V and transition kernel

$$p(x, y; G) = \mathbf{P}_x[X(1) = y] = \begin{cases} \dfrac{1}{2}, & \text{if } x = y, \\ \dfrac{1}{2\deg(x)}, & \text{if } \{x, y\} \in E. \end{cases}$$

Let

$$\tau_{\mathrm{cov}}(G) = \min\{t \geq 0 : V \text{ is contained in the range of } X|_{[0,t]}\}$$

be the *cover time* and let $T_{\mathrm{cov}}(G) = \mathbf{E}_\pi[\tau_{\mathrm{cov}}(G)]$ be the *expected cover time*. Here and hereafter, a subscript of π indicates that X is started from stationarity. Let $\tau(y) = \min\{t \geq 0 : X(t) = y\}$ be the first time X hits y and

$$T_{\mathrm{hit}}(G) = \max_{x, y \in V} \mathbf{E}_x[\tau(y)]$$

be the *maximal hitting time*. If (G_n) is a sequence of graphs with $T_{\mathrm{hit}}(G_n) = o(T_{\mathrm{cov}}(G_n))$, then a result of Aldous [4], Theorem 2, implies that $\tau_{\mathrm{cov}}(G_n)$ has a threshold around its mean: $\tau_{\mathrm{cov}}(G_n)/T_{\mathrm{cov}}(G_n) = 1 + o(1)$. Many sequences of

Received January 2010; revised October 2010.
MSC2010 subject classifications. 60J10, 60D05, 37A25.
Key words and phrases. Random walk, uncovered set, lamplighter walk, mixing time, cutoff.

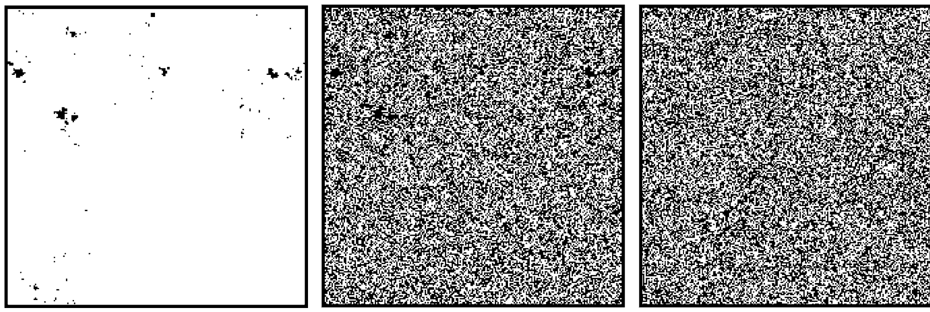

FIG. 1. *The subset* $\mathcal{L}(\frac{1}{2}, \mathbf{Z}_n^2)$ *of* $\mathbf{Z}_n^2$ *consisting of those points unvisited by a random walk* X *run for* $\frac{1}{2}T_{\text{cov}}(\mathbf{Z}_n^2)$, *where* $T_{\text{cov}}(\mathbf{Z}_n^2)$ *is the expected number of steps required for* X *to cover* $\mathbf{Z}_n^2$, *exhibits clustering. Consequently, the marking of* $\mathbf{Z}_n^2$ *by elements of* $\{0,1\}$ *given by the results of i.i.d. coin flips on the range of* X *at time* $\frac{1}{2}T_{\text{cov}}(\mathbf{Z}_n^2)$ *and zero otherwise can be distinguished from a uniform marking.* (a) $\mathcal{L}(\frac{1}{2}, \mathbf{Z}_n^2)$. (b) $\mathcal{L}(\frac{1}{2}, \mathbf{Z}_n^2)$ *marked with i.i.d. coin flips.* (c) $\mathbf{Z}_n^2$ *marked with i.i.d. coin flips.*

graphs satisfy this condition, for example, $\mathbf{Z}_n^d$ for $d \geq 2$, $\mathbf{Z}_2^n$, and the complete graph K_n. When Aldous' condition holds, the set

$$\mathcal{L}(\alpha; G_n) = \{x \in V_n : \tau(x) \geq \alpha T_{\text{cov}}(G_n)\},$$

V_n the vertices of G_n, of α-*late points*, that is, points hit after time $\alpha T_{\text{cov}}(G_n)$, $\alpha \in (0,1)$, often has an interesting structure. The case $G_n = \mathbf{Z}_n^2$ was first studied by Brummelhuis and Hilhorst in [8] where it is shown that $\mathbf{E}|\mathcal{L}(\alpha; \mathbf{Z}_n^2)|$ has growth exponent $2(1-\alpha)$ and that points in $\mathcal{L}_n(\alpha; \mathbf{Z}_n^2)$ are positively correlated. This suggests that $\mathcal{L}(\alpha; G_n)$ has a fractal structure and exhibits clustering. These statements were made precise by Dembo, Peres, Rosen and Zeitouni in [13] where they show that the growth exponent of $|\mathcal{L}(\alpha; \mathbf{Z}_n^2)|$ is $2(1-\alpha)$ *with high probability* in addition to making a rigorous quantification of the clustering phenomenon (see Figure 1 for an illustration of this).

If G_n is either K_n or $\mathbf{Z}_n^d$ for $d \geq 3$, then it is also true that $\log|\mathcal{L}(\alpha; G_n)| \sim (1-\alpha)\log|V_n|$ with high probability. In contrast to $\mathcal{L}(\alpha; \mathbf{Z}_n^2)$, $\mathcal{L}(\alpha; K_n)$ does not exhibit clustering and is "uniformly random" in the sense that conditional on $s_0 = |\mathcal{L}(\alpha; K_n)|$, all subsets of K_n of size s_0 are equally likely. The rapid decay of correlation in $\mathcal{L}(\alpha; \mathbf{Z}_n^d)$ for $d \geq 3$ determined by Brummelhuis and Hilhorst [8] indicates that the clustering phenomenon is also not present in this case and leads one to speculate that $\mathcal{L}(\alpha; \mathbf{Z}_n^d)$ is likewise in some sense "uniformly random."

The purpose of this article is to quantify the degree to which this holds. We use as our measure of uniformity the following statistical test. Let $\mathcal{R}(\alpha; G)$ be the (random) subset of V covered by X at time $\alpha T_{\text{cov}}(G)$ and let $\mu(\cdot; \alpha, G)$ be the probability measure on $\mathcal{X}(G) = \{f : V \to \{0,1\}\}$ given by first sampling $\mathcal{R}(\alpha; G)$

then setting

$$f(x) = \begin{cases} \xi(x), & \text{if } x \in \mathcal{R}(\alpha; G), \\ 0, & \text{otherwise,} \end{cases}$$

where $(\xi(x) : x \in V)$ is a collection of i.i.d. variables such that $\mathbf{P}[\xi(x) = 0] = \mathbf{P}[\xi(x) = 1] = \frac{1}{2}$. The question we are interested in is:

How large does $\alpha \in (0, 1)$ *need to be so that* $\mu(\cdot; \alpha, G)$ *is indistinguishable from the uniform measure* $\nu(\cdot; G)$ *on* $\mathcal{X}(G)$?

It must be that $\alpha \geq 1/2$ in the case of $\mathbf{Z}_n^d$ for $d \geq 2$ since if $\alpha < 1/2$ then

$$\frac{|\mathcal{L}(\alpha; \mathbf{Z}_n^d)| - (1/2)n^d}{n^{d/2}} \to \infty \qquad \text{as } n \to \infty.$$

In particular, the deviations of the number of zeros from $n^d/2$ which arise in a marking from such α far exceed that in the uniform case. By [4], Theorem 2, it is also true that $\alpha \leq 1$ since if $\alpha > 1$ then with high probability $|\mathcal{L}(\alpha; \mathbf{Z}_n^d)| = 0$. The main result of this article is that the threshold for indistinguishability for any sequence of graphs (G_n) with $\lim_{n\to\infty} |V_n| = \infty$ is $\alpha = \frac{1}{2}$ provided random walk on (G_n) is *uniformly locally transient* and satisfies a mild connectivity hypothesis.

We need the following definitions in order to give a precise statement of our results. The ε-*total variation mixing time* of G is

$$T_{\text{mix}}(\varepsilon; G) = \min\Big\{t \geq 0 : \max_{x \in V} \|p^t(x, \cdot; G) - \pi\|_{\text{TV}} \leq \varepsilon\Big\},$$

where $p^t(x, y; G) = \mathbf{P}_x[X(t) = y]$ is the t-step transition kernel of X started at x,

$$\|\mu - \nu\|_{\text{TV}} = \max_{A \subseteq V} |\mu(A) - \nu(A)| = \frac{1}{2} \sum_{x \in V} |\mu(x) - \nu(x)|$$

is the *total variation distance* between the measures μ, ν on V and π is the stationary distribution of X. The ε-*uniform mixing time* of G is

$$T_{\text{mix}}^U(\varepsilon; G) = \min\left\{t \geq 0 : \max_{x, y \in V} \left|\frac{p^t(x, y; G)}{\pi(y)} - 1\right| \leq \varepsilon\right\}.$$

It is a basic fact ([3, 20]; see also Proposition 3.3) that $T_{\text{mix}}^U(\varepsilon; G)$ is within a factor of $\log|V|$ of $T_{\text{mix}}(\varepsilon; G)$, however, for many graphs this factor is constant. Whenever we omit ε and write $T_{\text{mix}}(G), T_{\text{mix}}^U(G)$ it is understood that $\varepsilon = \frac{1}{4}$. *Green's function* of G is

$$g(x, y; G) = \sum_{t=0}^{T_{\text{mix}}^U(G)} p^t(x, y; G),$$

that is, the expected amount of time that X spends at y until time $T_{\mathrm{mix}}^U(G)$ when started at x. For $A \subseteq V$, we set

$$g(x, A; G) = \sum_{y \in A} g(x, y; G).$$

We say that (G_n) is *uniformly locally transient* with *transience function* $\rho : [0, \infty) \times [0, \infty) \to [0, \infty)$ if

$$g(x, A; G_n) \le \rho(d(x, A), \mathrm{diam}(A)) \qquad \text{for all } n \text{ and } x \in V_n, A \subseteq V_n.$$

Here, $d(\cdot, \cdot)$ is the graph distance, $d(x, A) = \min_{y \in A} d(x, y)$, and $\rho(\cdot, s)$ is assumed to be nonincreasing with $\lim_{r \to \infty} \rho(r, s) = 0$ when s is fixed. Let $\rho(r) = \rho(r, 1)$,

$$\overline{\Delta}(G) = \max_{x \in V} \deg(x), \qquad \underline{\Delta}(G) = \min_{x \in V} \deg(x) \quad \text{and} \quad \Delta(G) = \frac{\overline{\Delta}(G)}{\underline{\Delta}(G)}.$$

ASSUMPTION 1.1 (Transience). (G_n) is a sequence of uniformly locally transient graphs with $|V_n| \to \infty$ such that there exists $\Delta_0 > 0$ so that $\Delta(G_n) \le \Delta_0$ for all n and, for each $r > 0$:

(1) $\log|B(x, r)| = o(\log|V_n|)$ as $n \to \infty$, and
(2) $T_{\mathrm{mix}}^U(G_n)\overline{\Delta}^r(G_n) = o(|V_n|)$ as $n \to \infty$.

The reason for the hypothesis $\Delta(G_n) \le \Delta_0$ is that it implies

$$\frac{\pi(x; G_n)}{\pi(y; G_n)} \le \Delta_0 \qquad \text{uniformly in } x, y \in V_n \text{ and } n.$$

In particular, this combined with uniform local transience allows us to conclude that the hitting time of any two points $x, y \in V_n$ is comparable. The purpose of part (1) of Assumption 1.1 is to ensure that for every $r, n > 0$ we can construct an r-net $E_{r,n}$ of V_n whose size at logarithmic scales is comparable to $|V_n|$, that is, $\log|E_{r,n}| = \log|V_n| + o(1)$ as $n \to \infty$. Finally, part (2) of Assumption 1.1 is important since by a union bound it implies that the probability that X hits any fixed ball of finite radius within time $T_{\mathrm{mix}}^U(G_n)$ when initialized from stationarity tends to zero with n.

We will also need to make the following assumption.

ASSUMPTION 1.2 (Connectivity). (G_n) is a sequence of graphs satisfying either:

(1) for every $\gamma > 0$ there exists $R_n^\gamma \to \infty$ as $n \to \infty$ satisfying $R_n^\gamma \le \frac{1}{2} \times \max\{R > 0 : \max_{x \in V_n} |B(x, R)| \le |V_n|^\gamma\}$ such that for every $r > 0$,

$$\frac{T_{\mathrm{mix}}^U(G_n)}{R_n^\gamma} \max_{d(x,A) \ge R_n^\gamma} g(x, A) = o(1) \qquad \text{as } n \to \infty$$

uniformly in $A \subseteq V_n$ with $\mathrm{diam}(A) \le r$, or

(2) a uniform Harnack inequality, that is, for each $\alpha > 1$ there exists $C = C(\alpha) > 0$ such that for every $x, r, R > 0$ with $R/r \geq \alpha$ and positive harmonic function h on $B(x, R)$ we have that

$$\max_{y \in B(x,r)} h(y) \leq C \min_{y \in B(x,r)} h(y).$$

Assumption 1.2 ensures that (G_n) is in some sense well connected. In particular, part (1) is used to show that X is uniformly unlikely to hit a small ball before remixing provided its starting point and the small ball are far enough apart. This hypothesis will be relevant for graphs where $|\partial B(x, r)|$ is comparable to or larger than $|B(x, r)|$, as in the case of $\mathbf{Z}_2^n$ or graphs which are locally tree-like. Part (2) is meant to be applicable for graphs where $|\partial B(x, r)|$ is much smaller than $|B(x, r)|$, as in the case of $\mathbf{Z}_n^d$, and is used to deduce that the empirical average of the probability that successive excursions of X between concentric spheres $\partial B(x, r), \partial B(x, R)$ hit x conditional on their entrance and exit points is well concentrated around its mean provided $R > r$ are large enough.

We now state our main theorem.

THEOREM 1.3. *If (G_n) satisfies Assumptions 1.1 and 1.2, then for every $\varepsilon > 0$,*

$$\lim_{n \to \infty} \left\| \mu\left(\cdot; \frac{1}{2} + \varepsilon, G_n\right) - \nu(\cdot; G_n) \right\|_{\mathrm{TV}} = 0$$

and

$$\lim_{n \to \infty} \left\| \mu\left(\cdot; \frac{1}{2} - \varepsilon, G_n\right) - \nu(\cdot; G_n) \right\|_{\mathrm{TV}} = 1.$$

REMARK 1.4. If (G_n) is a sequence with $|V_n| \to \infty$ and $\sup_n \overline{\Delta}(G_n) < \infty$, then Assumption 1.1 is equivalent to the decay of $g(x, y; G_n)$ in $d(x, y)$ uniformly in n.

Many families satisfy Assumptions 1.1 and 1.2, for example, $\mathbf{Z}_n^d$ for $d \geq 3$, random d-regular graphs whp, also for $d \geq 3$, and the hypercube $\mathbf{Z}_2^n$. We will discuss these and other examples in the next section.

The problem that we consider is closely related to determining the mixing time of the *lamplighter walk*, which we now introduce; recall that $\mathcal{X}(G) = \{f : V \to \{0, 1\}\}$ is the set of markings of V by $\{0, 1\}$. If $G = (V, E)$ is a finite graph, the wreath product $G^\diamond = \mathbf{Z}_2 \wr G$ is the graph $(V^\diamond, E^\diamond)$ whose vertices are pairs (f, x) where $f \in \mathcal{X}(G)$ and $x \in V$. There is an edge between (f, x) and (g, y) if and only if $\{x, y\} \in E$ and $f(z) = g(z)$ for $z \notin \{x, y\}$. $G^\diamond$ is also referred to as the *lamplighter graph* over G since it can be constructed by placing "lamps" at the vertices of G; the first coordinate f of a configuration (f, x) indicates the state of the lamps and the second gives the location of the lamplighter.

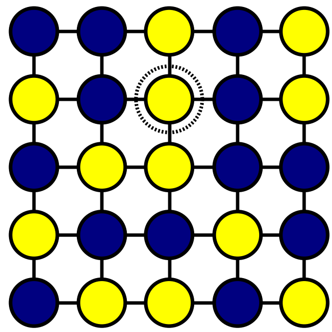

FIG. 2. *A typical configuration of the lamplighter over a* 5×5 *planar grid. The colors indicate the state of the lamps and the dashed circle gives the position of the lamplighter.*

The *lamplighter walk* $X^\diamond$ on G is the random walk on $G^\diamond$. Its transition kernel $p(\cdot,\cdot;G^\diamond)$ can be constructed from $p(\cdot,\cdot;G)$ using the following procedure: given $(f,x) \in V^\diamond$:

(1) sample $y \in V$ adjacent to x using $p(x,\cdot;G)$,
(2) randomize the values of $f(x)$, $f(y)$ using independent fair coin flips,
(3) move the lamplighter from x to y.

See Figure 2 for an example of a typical lamplighter configuration. That both $f(x)$ and $f(y)$ are randomized rather than just $f(y)$ is necessary for reversibility. It is obvious that the stationary distribution of $X^\diamond$ is $\nu(\cdot;G) \times \pi(G)$. For the graphs we consider, the mixing time of $X^\diamond$ is dominated by the mixing time of its first coordinate as it is comparable to $T_{\mathrm{cov}}(G)$ which in turn is much larger than $T_{\mathrm{mix}}(G)$, the mixing time of the second coordinate of $X^\diamond$. This will allow us to deduce $T_{\mathrm{mix}}(G^\diamond) = (\frac{1}{2} + o(1))T_{\mathrm{cov}}(G_n)$ for graphs satisfying Assumptions 1.1, 1.2 from Theorem 1.3.

Random walk on a sequence of graphs (G_n) is said to have a (total variation) cutoff with threshold (a_n) if

$$\lim_{n\to\infty} \frac{T_{\mathrm{mix}}(\varepsilon; G_n)}{a_n} = 1 \qquad \text{for all } \varepsilon \in (0,1).$$

It is believed that many graphs have a cutoff, but establishing this is often quite difficult since it requires a delicate analysis of the behavior of the underlying walk. The term was first coined by Aldous and Diaconis in [2] where they prove cutoff for the top-in-at-random shuffling process. Other early examples include random transpositions on the symmetric group [16], the riffle shuffle and random walk on the hypercube [1]. By making a small modification to the proof of Theorem 1.3, we are able to establish cutoff for the lamplighter walk on base graphs satisfying Assumptions 1.1 and 1.2.

Before we state these results, we will first summarize previous work related to this problem. The mixing time of $G^\diamond$ was first studied by Häggström and Jonasson in [18] in the case $G_n = K_n$ and $G_n = \mathbf{Z}_n$. Their work implies a cutoff with threshold $\frac{1}{2}T_{\text{cov}}(K_n)$ in the former case and that there is no cutoff in the latter. The connection between $T_{\text{mix}}(G^\diamond)$ and $T_{\text{cov}}(G)$ is explored further in [22], in addition to developing the relationship between the relaxation time of $G^\diamond$ and $T_{\text{hit}}(G)$, and $\mathbf{E}[2^{|\mathcal{L}(\alpha;G)|}]$ and $T^U_{\text{mix}}(G^\diamond)$. The results of [22] include a proof of cutoff when $G_n = \mathbf{Z}_n^2$ with threshold $T_{\text{cov}}(\mathbf{Z}_n^2)$ and a general bound that

$$[\tfrac{1}{2} + o(1)]T_{\text{cov}}(G_n) \le T_{\text{mix}}(G_n^\diamond) \le [1 + o(1)]T_{\text{cov}}(G_n), \tag{1.1}$$

whenever (G_n) is a sequence of vertex transitive graphs with $T_{\text{hit}}(G_n) = o(T_{\text{cov}}(G_n))$. It is not possible to improve upon (1.1) without further hypotheses since the lower and upper bounds are achieved by K_n and $\mathbf{Z}_n^2$, respectively.

The bound (1.1) applies to $\mathbf{Z}_n^d$ when $d \ge 3$ since $T_{\text{hit}}(\mathbf{Z}_n^d) \sim c_d n^d$ and $T_{\text{cov}}(\mathbf{Z}_n^d) = c'_d n^d (\log n)$ (see Proposition 10.13, Exercise 11.4 of [20]). This leads [22] to the question of whether there is a threshold for $T_{\text{mix}}((\mathbf{Z}_n^d)^\diamond)$ and, if so, if it is at $\frac{1}{2}T_{\text{cov}}(\mathbf{Z}_n^d)$, $T_{\text{cov}}(\mathbf{Z}_n^d)$ or somewhere in between. By a slight extension of our methods, we are able to show that the threshold is at $\frac{1}{2}T_{\text{cov}}(\mathbf{Z}_n^d)$ when $d \ge 3$, and that the same holds whenever (G_n) satisfies Assumptions 1.1 and 1.2.

THEOREM 1.5. *If (G_n) satisfies Assumptions 1.1 and 1.2, then $T_{\text{mix}}(\varepsilon; G_n^\diamond)$ has a cutoff with threshold $\frac{1}{2}T_{\text{cov}}(G_n)$.*

In order to prove Theorems 1.3 and 1.5, we need to develop a delicate understanding of both the process of coverage and the correlation structure of $\mathcal{L}(\alpha; G_n)$. The proof yields the following theorem, which gives a precise estimate of the decay of correlation in $\mathcal{L}(\alpha; G_n)$ under the additional hypothesis of vertex transitivity.

THEOREM 1.6. *Suppose (G_n) is a sequence of vertex transitive graphs satisfying Assumption 1.1. If (x_n^i) for $1 \le i \le \ell$ is a family of sequences with $x_n^i \in V_n$ and $|x_n^i - x_n^j| \ge r$ for every n and $i \ne j$, then*

$$\begin{aligned}(1 - \delta_{r,\ell})|V_n|^{-\ell\alpha - \delta_{r,\ell}} &\le \mathbf{P}[x_n^i \in \mathcal{L}(\alpha; G_n) \text{ for all } i] \\ &\le (1 + \delta_{r,\ell})|V_n|^{-\ell\alpha + \delta_{r,\ell}},\end{aligned} \tag{1.2}$$

where $\delta_{r,\ell} \to 0$ as $r \to \infty$ while ℓ is fixed. If $\overline{\Delta}(G_n) \to \infty$, we take $r = 1$ and $\delta_{1,\ell} = o(1)$ as $n \to \infty$.

Outline. The remainder of the article is structured as follows. We show in Section 2 that the hypotheses of Theorems 1.3 and 1.5 hold for a number of natural examples. In Section 3, we collect several general estimates that will be used throughout the rest of the article; Proposition 3.2 is in particular of critical importance.

Next, in Section 4 we will develop precise asymptotic estimates for the cover and hitting times of graphs (G_n) satisfying Assumption 1.1. The key idea is that the process by which X hits a point x can be understood by studying the excursions of X from $\partial B(x,r)$ through $\partial B(x,R)$, $r<R$ and then subsequently run for time $\beta T_{\mathrm{mix}}^{U}(G)$, some $\beta>0$, in order to remix. Due to the remixing, these excursions exhibit behavior which is close to that of i.i.d. random walk excursions initialized from stationarity. This has three important consequences. First, our transience assumptions imply that the number $N_H(x)$ of excursions up until the time $\tau(x)$ that x is hit is stochastically dominated from below by a geometric random variable with small parameter p provided $R>r$ are both large. Thus, $N_H(x)$ is typically very large. Second and consequently, the empirical average of the amount of time separating the beginning of successive excursions up to time $\tau(x)$ is very concentrated around its mean $T_{r,R}(x)$. Third, with $p_j(x)$ the probability that the jth excursion E_j hits x by time $\alpha T_{\mathrm{mix}}(G)$ after exiting $B(x,R)$, $\alpha\le\beta$, conditional on both the entrance point of E_j and E_{j+1} to $B(x,r)$, we have that $\frac{1}{k}\sum_{j=1}^{k}p_j(x)$ is also well concentrated around its mean $\overline{p}_{r,R}(x)$. Combining everything, this allows us to deduce the following asymptotic formula for the hitting time of x:

$$\mathbf{E}[\tau(x)]=\big(1+o(1)\big)T_{r,R}(x)\mathbf{E}[N_H(x)]=\frac{(1+o(1))T_{r,R}(x)}{\overline{p}_{r,R}(x)}.$$

For simplicity, we will now restrict our attention to graph families which are vertex transitive. This implies that $T_{r,R}=T_{r,R}(x)$ and $\overline{p}_{r,R}=\overline{p}_{r,R}(x)$ do not depend on x. Consequently, by the Matthews method upper and lower bounds ([21]; see also Theorem 11.2 and Proposition 11.4 of [20]) we infer that

$$T_{\mathrm{cov}}(G_n)=\big(1+o(1)\big)\frac{T_{r,R}}{\overline{p}_{r,R}}\log|V_n|. \tag{1.3}$$

We will now explain how we use these estimates to prove Theorems 1.3 and 1.5 in Section 6. By Proposition 3.2, to give an upper bound on the total variation distance of the i.i.d. marking of the range of random walk run for time $\frac{1}{2}T_{\mathrm{cov}}(G_n)$ from the uniform marking on V_n, it suffices to control the exponential moment of the set of points in V_n which are not visited by two independent random walks, each run for time $\frac{1}{2}T_{\mathrm{cov}}(G_n)$. Equation (1.3) implies that the number $N_{\mathrm{cov}}(x;\alpha)$ of excursions that have occurred by time $\alpha T_{\mathrm{cov}}(G_n)$ satisfies

$$N_{\mathrm{cov}}(x;\alpha)=\big(\alpha+o(1)\big)\log|V_n|/\overline{p}_{r,R}.$$

This in turn implies the tail decay

$$\mathbf{P}[\tau(x)\ge\alpha T_{\mathrm{cov}}(G_n)]=|V_n|^{-\alpha+o(1)}.$$

For points x,y which are far apart, it is unlikely that a single random walk excursion passes through both $B(x,R)$ and $B(y,R)$. That is, the process of hitting well-separated points exhibits mean-field behavior, which in turn allows us to give an efficient estimate of the relevant exponential moment. There are many technical challenges involved in getting all of these estimates to fit together correctly.

Decomposing the process of hitting into excursions between concentric spheres is not new, and is used to great effect, for example, in [10–13]. Our implementation of this idea is new since explicit representations of hitting probabilities and Green's functions in addition to the approximate rotational invariance available in the special case of $\mathbf{Z}_n^d$ are not available in the generality we consider.

We prove Theorem 1.6 in Section 5. This result, which may be of independent interest, is important in Section 6 since it allows us to deduce that points in $\mathcal{L}(\frac{1}{2}; G_n)$ are typically "spread apart." The article ends with a list of related open questions.

2. Examples.

$\mathbf{Z}_n^d$, $d \geq 3$. Although the simplest, this is the motivating example for this work. It is well known (see Section 1.5 of [19]) that there exists a constant $c_d > 0$ so that $g(x, y; \mathbf{Z}_n^d) \leq c_d |x - y|^{2-d}$, which implies uniform local transience. Assumption 1.2(2) is also satisfied since it is also a basic result that random walk on $\mathbf{Z}_n^d$ satisfies a Harnack inequality (see [19], Section 1.4).

Super-critical percolation cluster. Suppose that η_e is a collection of i.i.d. random variables indexed by the edges $e = (x, y)$ of $\mathbf{Z}^d$, $d \geq 3$, taking values in $\{0, 1\}$ such that $\mathbf{P}[\eta_e = 1] = p \in [0, 1]$. An edge e is called open if $\eta_e = 1$. Let $\mathcal{C}(x)$ denote the subset of $\mathbf{Z}^d$ consisting of those elements y that can be connected to x by a path consisting only of open edges. Let C_∞ denote the event that there exists an infinite open cluster and let $p_c = \inf\{p > 0 : \mathbf{P}[C_\infty] > 0\}$. Suppose $p > p_c$. Then it is known that there exists a unique infinite open cluster $\mathcal{C}_\infty$ almost surely. Fix $x \in \mathcal{C}_\infty$ and consider the graph $G_n = B(x, n) \cap \mathcal{C}_\infty$. It follows from the works of Delmotte [9], Deuschel and Pisztora [15], Pisztora [23] and Benjamini and Mossel [6] that the heat kernel for continuous time random walk (CTRW) on G_n has Gaussian tails whp when n is large enough; see the discussion after the statement of Theorem A of [5]. Consequently, Green's function of the CTRW on (G_n) has the same quantitative behavior as for $(\mathbf{Z}_n^d)$. This implies the same is true for the lazy random walk, which in turn yields uniform local transience for (G_n) whp when n is sufficiently large. Therefore there exists $n_0 = n_0(\omega)$ such that $(G_n : n \geq n_0(\omega))$ almost surely satisfies Assumption 1.1. Furthermore, it is a result of Barlow [5] that there exists $n_1 = n_1(\omega)$ such that random walk on $(G_n : n \geq n_1(\omega))$ almost surely satisfies a Harnack inequality and hence Assumption 1.2.

Bounded degree expanders. Suppose that (G_n) is an expander family with uniformly bounded maximal degree such that $|V_n| \to \infty$. Then there exists $T_0 < \infty$ such that $T_{\mathrm{rel}}(G_n) \leq T_0$ for every n where $T_{\mathrm{rel}}(G_n)$ is the relaxation time of lazy random walk on G_n. Equation (12.11) of [20] implies that

$$p^t(x, y; G_n) \leq C\left(\frac{1}{|V_n|} + e^{-t/T_0}\right)$$

and Theorem 12.3 of [20] gives $T_{\mathrm{mix}}^U(G_n) = O(\log|V_n|)$. By Remark 1.4, to check Assumption 1.1, we need only show the uniform decay $g(x, y; G_n)$ in $d(x, y)$. If $t < d(x, y)$, then it is obviously true that $p^t(x, y; G_n) = 0$. Hence,

$$\begin{aligned} g(x, y; G_n) &\le C\left(\frac{O(\log|V_n|)}{|V_n|} + \sum_{t=d(x,y)}^{T_{\mathrm{mix}}^U(G_n)} e^{-t/T_0}\right) \\ &\le C_1 e^{-d(x,y)/T_0} + o(1) \end{aligned} \tag{2.1}$$

as $n \to \infty$. We will now argue that (G_n) satisfies part (1) of Assumption 1.2. Suppose that $\overline{\Delta} \ge \max_{x \in V_n} \deg(x)$ for every n. We can obviously take $R_n^\gamma = \gamma \log|V_n|/(2 \log \overline{\Delta})$, hence we have $T_{\mathrm{mix}}^U(G_n)/R_n^\gamma = O(1)$ as $n \to \infty$. Combining this with (2.1) implies that (G_n) satisfies Assumption 1.2.

Random regular graphs. Suppose that $d \ge 3$ and let $\mathcal{G}_{n,d}$ denote the set of d-regular graphs on n vertices. It is well known [7] that, whp as $n \to \infty$, an element chosen uniformly from $\mathcal{G}_{n,d}$ is an expander. Consequently, whp, a sequence (G_n) where each G_n is chosen independently and uniformly from $\mathcal{G}_{n,d}$, $d \ge 3$, almost surely satisfies the hypotheses of our theorems.

Hypercube. As in the case of super-critical percolation, for $\mathbf{Z}_2^n$ it is easiest to prove bounds for the CTRW which, as we remarked before, easily translate over to the corresponding lazy walk. The transition kernel of the CTRW is

$$p^t(x, y; \mathbf{Z}_2^n) = \frac{1}{2^n}(1 + e^{-2t/n})^{n-|x-y|}(1 - e^{-2t/n})^{|x-y|},$$

where $|x - y|$ is the number of coordinates in which x and y differ. The spectral gap is $1/n$ (see Example 12.15 of [20]) which implies $\Omega(n) = T_{\mathrm{mix}}^U(\mathbf{Z}_2^n) = O(n^2)$ (see Theorem 12.3 of [20]). Suppose that $A \subseteq \mathbf{Z}_2^n$ has diameter s and $d(x, A) = r$. If $y \in A$, we have

$$p^t(x, y; \mathbf{Z}_2^n) \le \frac{1}{2^n}(1 + e^{-2t/n})^{n-r}(1 - e^{-2t/n})^r.$$

It is easy to see that

$$p^t(x, y; \mathbf{Z}_2^n) \le \begin{cases} \left(C_\varepsilon \dfrac{t}{n}\right)^r \exp\left(-\dfrac{t}{C_\varepsilon n}(n - r)\right), & \text{if } t \le \varepsilon n, \\ e^{-\rho_\varepsilon n}, & \text{if } t > \varepsilon n, \end{cases}$$

provided $\varepsilon > 0$ is sufficiently small. Consequently,

$$g(x, A; \mathbf{Z}_2^n) \le C n^{s-r}$$

and therefore $\mathbf{Z}_2^n$ is uniformly locally transient. The other hypotheses of Assumption 1.1 are obviously satisfied. As for Assumption 1.2, we note that in this case, we can take $R_n^\gamma = \gamma n/(2\log_2 n)$. Thus, if $r > 0$ it is easy to see that if $\operatorname{diam}(A) \le s$ and $d(x, A) \ge R_n^\gamma$ we have that

$$\sum_{y \in A} p^t(x, y; \mathbf{Z}_2^n) \le n^s e^{-\rho_\varepsilon n},$$

if $t > \varepsilon n$. On the other hand, if $t \le \varepsilon n$, then we have

$$\sum_{y \in A} p^t(x, y; \mathbf{Z}_2^n) \le n^s \left(\frac{C_\varepsilon t}{n}\right)^{\gamma n/(2\log_2 n)} e{-}t/(2C_\varepsilon).$$

Hence, it is not hard to see that $\mathbf{Z}_2^n$ satisfies Assumption 1.2.

Caley graph of S_n generated by transpositions. Let G_n be the Caley graph of S_n generated by transpositions. By work of Diaconis and Shahshahani [16], $T_{\mathrm{mix}}(G_n) = \Theta(n(\log n))$, which by Theorem 12.3 of [20] implies $T_{\mathrm{mix}}^U(G_n) = O(n^2(\log n)^2)$. We are now going to give a crude estimate of $p^t(\sigma, \tau; S_n)$. By applying an automorphism, we may assume without loss of generality that $\sigma = \mathrm{id}$. Suppose that $d(\mathrm{id}, \tau) = r$ and that $\tau_1, \dots, \tau_r$ are transpositions such that $\tau_r \cdots \tau_1 = \tau$. Then $\tau_1, \dots, \tau_r$ move at most $2r$ of the n elements of $\{1, \dots, n\}$, say, $k_1, \dots, k_{2r}$. Suppose $k_1', \dots, k_{2r}'$ are distinct from $k_1, \dots, k_{2r}$ and $\alpha \in S_n$ is such that $\alpha(k_i) = k_i'$ for $1 \le i \le r$. Then the automorphism of G_n induced by conjugation by α satisfies $\alpha\tau\alpha^{-1} \ne \tau$. Therefore, the size of the set of elements τ' in S_n such that there exists a graph automorphism φ of G_n satisfying $\varphi(\tau) = \tau'$ and $\varphi(\mathrm{id}) = \mathrm{id}$ is at least $\binom{n-2r}{2r} \ge 2^{-2r} n^{2r}((2r)!)^{-1}$ assuming $n \ge 8r$. Therefore,

$$p^t(e, \tau; G_n) \le \frac{2^{2r}(2r)!}{n^{2r}} \quad \text{and} \quad g(e, \tau; G_n) \le C(2^{2r}(2r)!)(\log n)^2 n^{2-2r}.$$

If $\operatorname{diam}(A) = s$, then trivially $|A| \le n^{2s}$ from which it is clear that (G_n) is uniformly locally transient. The other parts of Assumption 1.1 are obviously satisfied by G_n. As for Assumption 1.2, a simple calculation shows that we can take $R_n^\gamma \le \gamma n/4 + O(1)$. Hence setting $R_n^\gamma = \sqrt{n}$, a calculation analogous to the one above, gives that Assumption 1.2 is satisfied.

3. Preliminary estimates. The purpose of this section is to collect several general estimates that will be useful for us throughout the rest of the article.

LEMMA 3.1. *If μ, ν are measures with ν absolutely continuous with respect to μ and*

$$\int \frac{d\nu}{d\mu} d\nu = 1 + \varepsilon,$$

then

$$\|\nu - \mu\|_{\mathrm{TV}} \le \frac{\sqrt{\varepsilon}}{2}.$$

PROOF. This is a consequence of the Cauchy–Schwarz inequality:

$$\begin{aligned}\|\mu-\nu\|_{\mathrm{TV}}^2 &= \left(\frac{1}{2}\int\left|\frac{d\nu}{d\mu}-1\right|d\mu\right)^2 \le \frac{1}{4}\int\left|\frac{d\nu}{d\mu}-1\right|^2 d\mu\\ &= \frac{1}{4}\left(\int\frac{d\nu}{d\mu}\,d\nu-1\right).\end{aligned}$$

□

Let ν denote the uniform measure on $\mathcal{X}(G)=\{f:V\to\{0,1\}\}$.

PROPOSITION 3.2. *Suppose that μ is a measure on $\mathcal{X}(G)$ given by first sampling $\mathcal{R}\subseteq V$ according to a probability μ_0 on 2^V, then, conditional on $\mathcal{R}$ sampling $f\in\mathcal{X}(G)$ by setting*

$$f(x)=\begin{cases}\xi(x), & \text{if } x\in\mathcal{R},\\ 0, & \text{otherwise},\end{cases}$$

where $(\xi(x):x\in V)$ is a collection of i.i.d. random variables with $\mathbf{P}[\xi(x)=0]=\mathbf{P}[\xi(x)=1]=\frac{1}{2}$. Then

$$\int\frac{d\mu}{d\nu}\,d\mu=\iint 2^{|\mathcal{R}^c\cap\mathcal{S}^c|}\,d\mu_0(\mathcal{R})\,d\mu_0(\mathcal{S}).$$

PROOF. Letting $\mu(\cdot|\mathcal{S})$ be the conditional law of μ given $\mathcal{S}$ and $N=|V|$, we have

$$\begin{aligned}\int\frac{d\mu}{d\nu}\,d\mu &= 2^N\int\mu(\{f\})\,d\mu(f)\\ &= 2^N\iint\left(\int\mu(\{f\}|\mathcal{S})\,d\mu_0(\mathcal{S})\right)d\mu(f|\mathcal{R})\,d\mu_0(\mathcal{R}).\end{aligned}$$

Suppose $f\in\mathcal{X}(G)$ is such that $f|_{\mathcal{R}^c}\equiv 0$ for some $\mathcal{R}\subseteq V$. Note that

$$\mu(\{f\}|\mathcal{S})=2^{-|\mathcal{R}\cap\mathcal{S}|-|\mathcal{S}\setminus\mathcal{R}|}\mathbf{1}_{\{f|_{\mathcal{R}\setminus\mathcal{S}}\equiv 0\}}.$$

Hence, the above is equal to

$$\begin{aligned}&2^N\iint\left(\int 2^{-|\mathcal{R}\cap\mathcal{S}|-|\mathcal{S}\setminus\mathcal{R}|}\mathbf{1}_{\{f|_{\mathcal{R}\setminus\mathcal{S}}\equiv 0\}}\,d\mu_0(\mathcal{S})\right)d\mu(f|\mathcal{R})\,d\mu_0(\mathcal{R})\\ &\quad= 2^N\iint 2^{-|\mathcal{R}\cap\mathcal{S}|-|\mathcal{S}\setminus\mathcal{R}|}\left(\int\mathbf{1}_{\{f|_{\mathcal{R}\setminus\mathcal{S}}\equiv 0\}}\,d\mu(f|\mathcal{R})\right)d\mu_0(\mathcal{R})\,d\mu_0(\mathcal{S})\\ &\quad= 2^N\iint 2^{-|\mathcal{R}\cap\mathcal{S}|-|\mathcal{S}\setminus\mathcal{R}|}2^{-|\mathcal{R}\setminus\mathcal{S}|}\,d\mu_0(\mathcal{S})\,d\mu_0(\mathcal{R}).\end{aligned}$$

Simplifying the expression in the exponent gives the result. □

Roughly speaking, the general strategy of our proof will be to show that if $\mathcal{R},\mathcal{R}'$ denote independent copies of the range of random walk on G_n run up to time $(\frac{1}{2}+\varepsilon)T_{\mathrm{cov}}(G_n)$ and $\mathcal{L}=V\setminus\mathcal{R}$, $\mathcal{L}'=V\setminus\mathcal{R}'$ then

$$\mathbf{E}\exp(\zeta|\mathcal{L}\cap\mathcal{L}'|)=1+o(1)\qquad\text{as } n\to\infty \tag{3.1}$$

for $\zeta > 0$. This method cannot be applied directly, however, since this exponential moment blows up even in the case of $\mathbf{Z}_n^3$. To see this, suppose that X, X' are independent random walks on $\mathbf{Z}_n^3$ initialized at stationarity. We divide the cover time $c_3 n^3(\log n)$ into rounds of length n^2. In the first round, with probability $1/4$ we know that X starts in $\mathbf{L}_1 = \mathbf{Z}_n^2 \times \{n/8, \ldots, 3n/8\}$. In each successive round, X has probability $\rho_0 > 0$ strictly bounded from zero in n of not leaving $\mathbf{L}_2 = \mathbf{Z}_n^2 \times \{1, \ldots, n/2\}$ and ending the round in $\mathbf{L}_1$. Since there are $c_3 n(\log n)$ rounds, this means that X does not leave $\mathbf{L}_1$ with probability at least

$$\tfrac{1}{4}\rho_0^{c_3 n \log n} \geq c \exp(-\rho_1 n \log n).$$

Since X' satisfies the same estimate, we therefore have

$$\mathbf{E}\exp(\zeta|\mathcal{L} \cap \mathcal{L}'|) \geq c \exp\left(\frac{\zeta}{2} n^3 - 2\rho_1 n \log n\right) \to \infty \qquad \text{as } n \to \infty.$$

The idea of the proof is to truncate the exponential moment in (3.1) by conditioning the law of random walk run for time $(\frac{1}{2} + \varepsilon) T_{\mathrm{cov}}(G_n)$ *conditional on typical behavior* so that

$$\|\widetilde{\mu}_0 - \mu_0\|_{\mathrm{TV}} = o(1) \qquad \text{as } n \to \infty.$$

We do this in such a way that the uncovered set exhibits a great deal of spatial independence in order to make the exponential moment easy to estimate. To this end, we will condition on two different events. The first is that points in $\mathcal{L}(\frac{1}{2} + \delta; G_n)$ are well separated: for any $x \in V_n$ the number of points in $\mathcal{L}(\frac{1}{2} + \varepsilon; G_n)$ which are contained in a large ball centered at x is at most some constant M. Given this event, we can partition $\mathcal{L}(\frac{1}{2} + \varepsilon; G_n)$ into disjoint subsets $E_1, \ldots, E_M$ such that $x, y \in E_\ell$ distinct implies $d(x, y)$ is large. Observe

$$\mathbf{E}\exp(\zeta|\mathcal{L} \cap \mathcal{L}' \cap E_\ell|) \leq \mathbf{E} \prod_{x \in E_\ell} \left(1 + e^{\zeta} \prod_{j=1}^{N'(x,T)} (1 - q_j'(x))\right),$$

where $N'(x, T)$ is the number of excursions of X' from $\partial B(x, r)$ to $\partial B(x, R)$ by time T and $q_j'(x)$ is the probability the jth such excursion hits x conditional on its entrance and exit points. When T is large, uniform local transience implies that $N'(x, T)$ and $\prod_{j=1}^{k} q_j'(x)$ can be estimated by their mean and, roughly speaking, this is the second event on which we will condition. Finally, we get control of the entire exponential moment by an application of Hölder's inequality.

We finish the section by recording a standard lemma that bounds the rate of decay of the total variation and uniform distances to stationarity:

PROPOSITION 3.3. *For every $s, t \in \mathbf{N}$,*

$$\max_x \|p^{t+s}(x, \cdot) - \pi\|_{\mathrm{TV}} \leq 4 \max_{x,y} \|p^t(x, \cdot) - \pi\|_{\mathrm{TV}} \|p^s(y, \cdot) - \pi\|_{\mathrm{TV}}, \tag{3.2}$$

$$\max_{x,y} \left|\frac{p^{t+s}(x, y)}{\pi(y)} - 1\right| \leq \max_{x,y} \frac{p^s(x, y)}{\pi(y)} \max_x \|p^t(x, \cdot) - \pi\|_{\mathrm{TV}}. \tag{3.3}$$

PROOF. The first part is a standard result; see, for example, Lemmas 4.11 and 4.12 of [20]. The second part is a consequence of the semigroup property:

$$\begin{aligned}\frac{1}{\pi(z)}p^{t+s}(x,z) &= \frac{1}{\pi(z)}\sum_y p^t(x,y)p^s(y,z)\\ &= \frac{1}{\pi(z)}\sum_y [p^t(x,y)-\pi(y)+\pi(y)]p^s(y,z)\\ &\le \left(\max_{y,z}\frac{p^s(y,z)}{\pi(z)}\right)\|p^t(x,\cdot)-\pi\|_{\mathrm{TV}}+1. \qquad \square\end{aligned}$$

Note that (3.2) and (3.3) give

$$\max_x \|p^t(x,\cdot)-\pi\|_{\mathrm{TV}} \le ce^{-c\alpha} \qquad \text{for } t\ge \alpha T_{\mathrm{mix}}(G), \tag{3.4}$$

$$\max_{x,y}\left|\frac{p^{t+s}(x,y)}{\pi(y)}-1\right| \le ce^{-c\alpha} \qquad \text{for } t\ge T_{\mathrm{mix}}^U(G)+\alpha T_{\mathrm{mix}}(G), \tag{3.5}$$

where $c>0$ is a universal constant. We will often use (3.5) without reference, and, for simplicity use that the same inequality holds when $T_{\mathrm{mix}}^U(G)+\alpha T_{\mathrm{mix}}(G)$ is replaced by $\alpha T_{\mathrm{mix}}^U(G)$, perhaps adjusting $c>0$.

4. Hitting and cover times. Throughout, we assume that we have a sequence of graphs (G_n) satisfying Assumption 1.1 with transience function ρ. We will often suppress the index n and refer to an element of (G_n) as G and similarly write V, E for V_n, E_n, respectively. The primary purpose of this section is to develop asymptotic estimates of the maximal hitting and cover times of (G_n). Roughly, these will be given in terms of:

(1) the return time $T_{r,R}(x)$, $x\in V$, of X to $B(x,r)$ after passing through $B(x,R)$, $R>r$, large then allowed to remix, and

(2) the probability $\overline{p}_{r,R}(x)$ that upon entering $B(x,r)$, X subsequently hits x before exiting $B(x,R)$.

The derivation of these formulas requires many technical steps, so we will provide an overview of how everything fits together before delving into the details.

Let $N(x,t)$ be the number of excursions made by X from $\partial B(x,r)$ to $\partial B(x,R)$, then subsequently allowed to remix by running for some multiple of $T_{\mathrm{mix}}^U(G)$, by time t and let $p_j(x)$ be the probability that the jth excursion E_j hits x conditional on the entrance points of E_j and E_{j+1} to $B(x,r)$. Since the $p_j(x)$ are independent, we can express the probability $P(x,t)$ that x has not been hit by time t by the formula

$$P(x,t)=\mathbf{E}\prod_{j=1}^{N(x,t)}\bigl(1-p_j(x)\bigr).$$

We will argue using uniform local transience that we can make $p_j(x)$ as small as we like by choosing $R > r$ large enough. Consequently, we have

$$P(x,t) = \mathbf{E}\left[\exp\left(-\big(1 + O(\rho(r))\big) \sum_{j=1}^{N(x,t)} p_j(x)\right)\right].$$

Our first goal, accomplished in the next subsection, is to show that the empirical mean $\frac{1}{k}\sum_{j=1}^{k} p_j(x)$ is concentrated around its mean $\overline{p}_{r,R}(x)$. Next, in Section 4.2, we will again use concentration to argue that $N(x,t) \approx t/T_{r,R}(x)$. These two steps allow us to conclude that $P(x,t)$ is approximately given by $\exp(-t\overline{p}_{r,R}(x)/T_{r,R}(x))$. That is, $P(x,t)$ is approximately exponential with parameter $\overline{p}_{r,R}(x)/T_{r,R}(x)$ so that the expected hitting time of x is approximately $T_{r,R}(x)/\overline{p}_{r,R}(x)$. In the vertex transitive case, this immediately leads to an estimate of $(T_{r,R}/\overline{p}_{r,R})\log|V|$ for the cover time via the Matthews method ([21]; see also Theorem 11.2 and Proposition 11.4 of [20]). A similar but more complicated formula also holds for graphs which are not vertex transitive and is derived in the second half of Section 4.3.

4.1. *Probability of success.* Fix $R > r$ and let X be a lazy random walk on G. Suppose $A = \{x_1, \dots, x_\ell\} \subseteq V$ where $d(x_i, x_j) \geq 2R$ for $i \neq j$. Let $A(s) = \{x \in V : d(x, A) \leq s\}$ where $d(x, A) = \min_{y \in A} d(x, y)$. Let $\partial A(s) = \{x \in V : d(x, A) = s\}$. The purpose of this section is to prove that the empirical mean of the conditional probability that successive excursions of X from $\partial A(r)$ through $\partial A(R)$ succeed in hitting $x \in A$ given their entrance points concentrates around its mean. We will need to extend our excursions by multiples of the uniform mixing time $T^U_{\mathrm{mix}}(G)$ so we have enough independence to get good concentration.

To this end, we fix $\beta \geq 0$, set $T^U_\beta = \beta T^U_{\mathrm{mix}}(G)$, and define stopping times

$$\tau_0(A) = \min\{t \geq 0 : X(t) \in \partial A(r)\}, \tag{4.1}$$

$$\sigma_0(A) = \min\{t \geq \tau_0(A) : X(t) \notin A(R)\} \tag{4.2}$$

and inductively set

$$\tau_k^\beta(A) = \min\{t \geq \sigma_{k-1}^\beta(A) + T^U_\beta : X(t) \in \partial A(r)\}, \tag{4.3}$$

$$\sigma_k^\beta(A) = \min\{t \geq \tau_k^\beta(A) : X(t) \notin A(R)\}. \tag{4.4}$$

See Figure 3 for an illustration of the stopping times described in (4.1)–(4.4). Fix $\alpha \in [0, \beta]$. Let $S_j^{\alpha,\beta}(x; A)$ be the event that $X(t)$ hits x in $[\tau_j^\beta(A), \sigma_j^\beta(A) + T^U_\alpha]$,

$$p_j^{\alpha,\beta}(x; A) = \mathbf{P}[S_j^{\alpha,\beta}(x; A) | X(\tau_j^\beta(A)), X(\tau_{j+1}^\beta(A))]$$

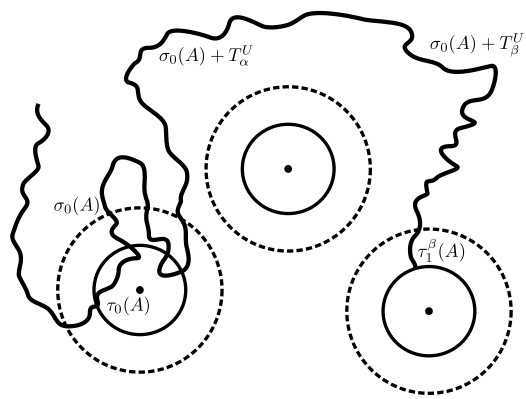

FIG. 3. *The solid and dashed circles represent the boundaries of $A(r)$ and $A(R)$, respectively, and the small points are the elements of A. Note that X may re-enter $A(r)$ during the interval* $[\sigma_k^\beta(A), \tau_{k+1}^\beta(A)]$.

and

$$a_j^{\alpha,\beta}(x;A) = \mathbf{E}\left[\sum_{t=\tau_j^\beta(A)}^{\sigma_j^\beta(A)+T_\alpha^U} \mathbf{1}_{\{X(t)=x\}} \middle| X(\tau_j^\beta(A)), X(\tau_{j+1}^\beta(A))\right].$$

The reason that it is useful to consider $p_{r,R}^{\alpha,\beta}(x;A)$ for $\beta > \alpha$ is that, as we will prove in Lemma 6.3, this allows us to show that the effect of conditioning on the terminal point $X(\tau_{j+1}^\beta(A))$ of the excursion is negligible when $\beta - \alpha$ is large enough. This in turn allows us to use uniform local transience to get that $p_{r,R}^{\alpha,\beta}(x;A)$ can be bounded in terms of the transience function. Finally, we let $\overline{p}_{r,R}^{\alpha,\beta}(x;A) = \mathbf{E}_\pi p_0^{\alpha,\beta}(x;A)$ and $\overline{a}_{r,R}^{\alpha,\beta}(x;A) = \mathbf{E}_\pi a_0^{\alpha,\beta}(x;A)$. For $\beta \geq \alpha \geq 1$ note that

$$\overline{p}_{r,R}^{\alpha,\beta}(x;A) = \overline{p}_{r,R}^{1,1}(x;A) + O\left(\frac{T_\beta^U}{|V|}\right)$$

since a union bound implies that the probability X hits x in the interval $[\sigma_0(A) + T_1^U, \sigma_0(A) + T_\beta^U]$ is $O(T_\beta^U/|V|)$.

By Assumption 1.1, we have that $T_\beta^U \overline{\Delta}^r(G)/|V| = o(1)$ as $n \to \infty$. Note that $\overline{p}_{r,R}^{1,1}(x;A) \geq (2\overline{\Delta}(G))^{-r}$ since the right-hand side bounds from below the probability that X goes directly from $\partial B(x,r)$ to x in r steps. Consequently,

$$\overline{p}_{r,R}^{\alpha,\beta}(x;A) = \big(1+o(1)\big)\overline{p}_{r,R}^{1,1}(x;A). \tag{4.5}$$

From now on, we will write $\overline{p}_{r,R}(x;A)$ for $\overline{p}_{r,R}^{1,1}(x;A)$. By the same argument, it is also true that $\overline{a}_{r,R}^{\alpha,\beta}(x;A) = (1+o(1))\overline{a}_{r,R}^{1,1}(x;A)$ and we will also write $\overline{a}_{r,R}(x;A)$ for $\overline{a}_{r,R}^{1,1}(x;A)$.

LEMMA 4.1. *For each $\delta > 0$ there exists $\gamma_0 > 0$ such that for $\beta - \alpha \ge \gamma_0$ and all n large enough we have*

$$1 - \delta \le \frac{p_j^{\alpha,\beta}(x; A)}{\mathbf{P}[S_j^{\alpha,\beta}(x) | X(\tau_j^\beta(A))]} \le 1 + \delta, \tag{4.6}$$

$$1 - \delta \le \frac{a_j^{\alpha,\beta}(x; A)}{\mathbf{E}[\sum_{t=\tau_j^\beta(A)}^{\sigma_j^\beta(A)+T_\alpha^U} \mathbf{1}_{\{X(t)=x\}} | X(\tau_j^\beta(A))]} \le 1 + \delta. \tag{4.7}$$

In particular, $p_j^{\alpha,\beta}(x; A) \le (1+\delta)\rho(r)$ and $a_j^\beta(x; A) \le (1+\delta)\rho(0)\rho(r)$ where ρ is the transience function.

PROOF. Note that

$$\begin{aligned}
&\mathbf{P}\big[X\big(\sigma_j^\beta(A) + T_\alpha^U\big) = z | X(\tau_j^\beta(A)) = z_j, X(\tau_{j+1}^\beta(A)) = z_{j+1}\big] \\
&\quad = \frac{\mathbf{P}[X(\sigma_j^\beta(A) + T_\alpha^U) = z, X(\tau_j^\beta(A)) = z_j, X(\tau_{j+1}^\beta(A)) = z_{j+1}]}{\mathbf{P}[X(\tau_j^\beta(A)) = z_j, X(\tau_{j+1}^\beta(A)) = z_{j+1}]} \\
&\quad = \left(\frac{\mathbf{P}[X(\tau_{j+1}^\beta(A)) = z_{j+1} | X(\sigma_j^\beta(A) + T_\alpha^U) = z]}{\mathbf{P}[X(\tau_{j+1}^\beta(A)) = z_{j+1} | X(\tau_j^\beta(A)) = z_j]}\right) \\
&\qquad \times \mathbf{P}\big[X\big(\sigma_j^\beta(A) + T_\alpha^U\big) = z | X(\tau_j^\beta(A)) = z_j\big].
\end{aligned}$$

Mixing considerations imply

$$\mathbf{P}[X(\tau_{j+1}^\beta(A)) = z_{j+1} | X(\tau_j^\beta(A)) = z_j] = [1 + O(e^{-c\beta})]\mathbf{P}_\pi[X(\tau_0(A)) = z_{j+1}]$$

and

$$\begin{aligned}
&\mathbf{P}\big[X(\tau_{j+1}^\beta(A)) = z_{j+1} | X\big(\sigma_j^\beta(A) + T_\alpha^U\big) = z\big] \\
&\quad = \big[1 + O\big(e^{-c(\beta-\alpha)}\big)\big]\mathbf{P}_\pi[X(\tau_0(A)) = z_{j+1}].
\end{aligned}$$

Consequently, if μ_j denotes the law of $X(\sigma_j^\beta(A) + T_\alpha^U)$ conditional on $X(\tau_j^\beta(A))$ and $X(\tau_{j+1}^\beta(A))$ and μ is the law of $X(\sigma_j^\beta(A) + T_\alpha^U)$ but conditional only on $X(\tau_j^\beta(A))$, we have $1 - \delta \le d\mu_j/d\mu \le 1 + \delta$ when $\beta - \alpha$ is large enough. Thus,

$$\begin{aligned}
p_j^{\alpha,\beta}(x; A) &= \int \mathbf{P}\big[S_j^{\alpha,\beta}(x) | X(\tau_j^\beta(A)), X\big(\sigma_j^\beta(A) + T_\alpha^U\big) = z, X(\tau_{j+1}^\beta(A))\big]\, d\mu_j(z) \\
&\le (1+\delta) \int \mathbf{P}\big[S_j^{\alpha,\beta}(x) | X(\tau_j^\beta(A)), X\big(\sigma_j^\beta(A) + T_\alpha^U\big) = z\big]\, d\mu(z) \\
&= (1+\delta)\mathbf{P}[S_j^{\alpha,\beta}(x) | X(\tau_j^\beta(A))].
\end{aligned}$$

The lower bound for $p_j^{\alpha,\beta}(x;A)$ and the bounds for $a_j(x;A)$ are proved similarly. $\square$

In the next lemma, we will prove the concentration of $p_j^{\alpha,\beta}(x;A)$ and $a_j^{\alpha,\beta}(x;A)$. The proof consists of three main steps. First, the previous lemma allows us to replace $p_j^{\alpha,\beta}(x;A)$ by $\mathbf{P}[S_j^{\alpha,\beta}(x;A)|X(\tau_j(A))]$ and likewise for $a_j^{\alpha,\beta}(x;A)$. Roughly, the next step is to use a stochastic domination argument to show that we can replace $\mathbf{P}[S_j^{\alpha,\beta}(x;A)|X(\tau_j(A))]$ by i.i.d. variables with law $\mathbf{P}[S_1^{\alpha,\beta}(x,A)|X(\tau_1(A))]$. The result then follows by an application of Cramér's theorem.

LEMMA 4.2. *Fix $r>0$ and $\delta\in(0,1)$. There exists $\gamma_0>0$ depending only on r,δ such that for all $R\geq r$, $\beta-\alpha\geq\gamma_0$ and n large enough we have*

$$\begin{aligned}&\mathbf{P}\Biggl[\sum_{j=1}^k p_j^{\alpha,\beta}(x;A)\notin[1-\delta,1+\delta]\overline{p}_{r,R}(x;A)k\Biggr]\\ &\qquad\leq 4\exp\biggl(-\frac{C\delta^2\overline{p}_{r,R}(x;A)}{\rho(r)}k\biggr)\end{aligned}\tag{4.8}$$

and

$$\begin{aligned}&\mathbf{P}\Biggl[\sum_{j=1}^k a_j^{\alpha,\beta}(x;A)\notin[1-\delta,1+\delta]\overline{a}_{r,R}(x;A)k\Biggr]\\ &\qquad\leq 4\exp\biggl(-\frac{C\delta^2\overline{a}_{r,R}(x;A)}{\rho(r)}k\biggr),\end{aligned}\tag{4.9}$$

where $C>0$ is independent of r,R,δ.

PROOF. Let μ be the measure on $\partial A(r)$ induced by the law of $X(\tau_0(A))$ given that X has a stationary initial distribution. For each $\delta>0$, let $\mathcal{M}(\delta)$ be the set of measures ν on $\partial A(r)$ which are uniformly mutually absolutely continuous with respect to μ in the sense that

$$\max_{z\in\partial A(r)}\left|\frac{\nu(z)}{\mu(z)}-1\right|+\max_{z\in\partial A(r)}\left|\frac{\mu(z)}{\nu(z)}-1\right|\leq\delta.\tag{4.10}$$

Let $\mu_y(z)=\mathbf{P}_y[X(\tau^\gamma(A))=z]$ where $\tau^\gamma(A)=\min\{t\geq T_\gamma^U:X(t)\in\partial A(r)\}$. Mixing considerations imply that $\mu_y\in\mathcal{M}(Ce^{-C\gamma})$ for some $C>0$. Fix $\delta>0$, $\delta'<\delta/2$, and take $\beta-\alpha=\gamma$ so large that $Ce^{-C\gamma}\leq\delta'/2$. Let $\overline{\mu}$, $\underline{\mu}$ be elements of $\mathcal{M}(\delta'/2)$ such that $\mathbf{P}[S_0^{\alpha,\beta}(x)|X(\tau_0(A))=Z]$ where $Z\sim\overline{\mu},\underline{\mu}$ stochastically dominates from above and below, respectively, all other choices in $\mathcal{M}(\delta'/2)$. Assume that γ_0 is chosen sufficiently large so that the previous lemma applies for $\delta'/2$ when n is sufficiently large.

Let $(U_j), (L_j)$ be i.i.d. sequences with laws $\mathbf{P}[S_0^{\alpha,\beta}(x)|X(\tau_0(A)) = Z]$, $Z \sim \overline{\mu}, \underline{\mu}$, respectively. With $\overline{U} = \mathbf{E}U_1$ and $\overline{L} = \mathbf{E}L_1$, obviously

$$(1-\delta')\overline{p}_{r,R}(x;A) \le \overline{L} \le \overline{U} \le (1+\delta')\overline{p}_{r,R}(x;A).$$

By construction, we can find a coupling of $U_j, L_j, p_j^{\alpha,\beta}(x;A)$ so that

$$L_j \le p_j^{\alpha,\beta}(x;A) \le U_j \qquad \text{almost surely for all } j.$$

Corollary 2.4.5 of [14] implies

$$\mathbf{E}e^{\lambda U_1} \le \frac{1}{2\rho(r)}\big(\overline{U}e^{2\lambda\rho(r)} + 2\rho(r) - \overline{U}\big)$$

hence Exercise 2.2.26 of [14] gives that the Fenchel–Legendre transform Λ^* of the law of U_1 satisfies

$$\Lambda^*(u) \ge \widetilde{\Lambda}^*(u) \equiv \frac{u}{2\rho(r)}\log\left(\frac{u}{\overline{U}}\right) + \left(1 - \frac{u}{2\rho(r)}\right)\log\left(\frac{1-u/(2\rho(r))}{1-\overline{U}/(2\rho(r))}\right).$$

As

$$\widetilde{\Lambda}^*(\overline{U}) = (\widetilde{\Lambda}^*)'(\overline{U}) = 0 \quad \text{and} \quad (\widetilde{\Lambda}^*)''(u) \ge \frac{1}{2\rho(r)u}$$

we have

$$\inf_{u \ge (1+\delta')\overline{U}} \Lambda^*(u) \ge \frac{1}{4\rho(r)\overline{U}}(\delta')^2\overline{U}^2 = \frac{(\delta')^2\overline{U}}{4\rho(r)},$$

assuming $\delta' < 1$. Consequently, Cramér's theorem (Theorem 2.2.3, part (c), of [14]) implies that

$$\mathbf{P}\left[\sum_{i=1}^{k} U_i \le (1+\delta')\overline{U}k\right] \ge 1 - 2\exp\left(-\frac{(\delta')^2\overline{U}k}{4\rho(r)}\right). \tag{4.11}$$

An analogous estimate also holds for (L_i) with $\overline{U}$ replaced by $\overline{L}$. The proof of concentration for the $a_j^{\alpha,\beta}(x;A)$ is the same. □

4.2. *Excursion lengths*. We will make use of the same notation in this subsection as in the previous. The main result is Lemma 4.5, which is that the empirical average of successive excursion lengths $\tau_{k+1}^{\beta}(A) - \tau_k^{\beta}(A)$ is exponentially concentrated around its mean. The proof requires two auxiliary inputs. The first, Lemma 4.3, is an estimate of the Radon–Nikodym derivative of the law of random walk conditioned not to hit $A(r)$ with respect to the stationary measure π. The second, Lemma 4.4, gives that the mean length of an excursion does not depend strongly on its starting point. Let $\tau(A) = \tau_0(A)$.

LEMMA 4.3. *For $\alpha, s \geq 0$ we have*

$$\mathbf{P}_y[X(T_\alpha^U) = z|\mathcal{A}] = \bigl[1 + O\bigl(e^{-c\alpha} + |A|\rho(s,r)\bigr) + o(1)\bigr]\pi(z) \qquad \text{as } n \to \infty,$$

where $\mathcal{A} = \{\tau(A) \geq T_\alpha^U, d(X(T_\alpha^U), A) \geq s\}$.

PROOF. For $z \in V$ with $d(z, A) \geq s$, observe

$$\begin{aligned}\mathbf{P}_y[X(T_\alpha^U) = z|\mathcal{A}] &= \frac{\mathbf{P}_y[X(T_\alpha^U) = z, \tau(A) \geq T_\alpha^U]}{\mathbf{P}_y[\mathcal{A}]} \\ &= \frac{\mathbf{P}_y[\tau(A) \geq T_\alpha^U|X(T_\alpha^U) = z]\mathbf{P}_y[X(T_\alpha^U) = z]}{\mathbf{P}_y[\mathcal{A}]} \qquad (4.12) \\ &= \bigl(1 + O(e^{-c\alpha})\bigr)\frac{\mathbf{P}_y[\tau(A) \geq T_\alpha^U|X(T_\alpha^U) = z]\pi(z)}{\mathbf{P}_y[\mathcal{A}]}.\end{aligned}$$

Fix $\alpha' < \alpha$. The idea of the proof is now to argue it is unlikely for $\tau(A)$ to occur in the interval $[T_\alpha^U - T_{\alpha'}^U, T_\alpha^U)$. This allows us to replace T_α^U above by $T_\alpha^U - T_{\alpha'}^U$ in (4.12). This in turn allows us to use mixing considerations to deduce that conditioning on $\{X(T_\alpha^U) = z\}$ has little effect on the probability of $\{\tau(A) \geq T_\alpha^U - T_{\alpha'}^U\}$. We compute

$$\begin{aligned}&\mathbf{P}_y[\tau(A) \geq T_\alpha^U|X(T_\alpha^U) = z] \\ &\quad= \mathbf{P}_y[\tau(A) \geq T_\alpha^U - T_{\alpha'}^U|X(T_\alpha^U) = z] \\ &\qquad - \mathbf{P}_y[T_\alpha^U > \tau(A) \geq T_\alpha^U - T_{\alpha'}^U|X(T_\alpha^U) = z].\end{aligned}$$

We have

$$\begin{aligned}&\mathbf{P}_y[\tau(A) \geq T_\alpha^U - T_{\alpha'}^U|X(T_\alpha^U) = z] \\ &\quad= 1 - \frac{\mathbf{P}_y[\tau(A) < T_\alpha^U - T_{\alpha'}^U, X(T_\alpha^U) = z]}{\mathbf{P}_y[X(T_\alpha^U) = z]} \\ &\quad= 1 - \frac{1 + O(e^{-c\alpha})}{\pi(z)}\mathbf{P}_y[X(T_\alpha^U) = z|\tau(A) < T_\alpha^U - T_{\alpha'}^U] \\ &\qquad\quad \times \mathbf{P}_y[\tau(A) < T_\alpha^U - T_{\alpha'}^U] \\ &\quad= \mathbf{P}_y[\tau(A) \geq T_\alpha^U - T_{\alpha'}^U] + O\bigl(e^{-c(\alpha-\alpha')}\bigr).\end{aligned}$$

Note that

$$\begin{aligned}&\mathbf{P}_y[T_\alpha^U > \tau(A) \geq T_\alpha^U - T_{\alpha'}^U|X(T_\alpha^U) = z] \\ &\quad= \frac{1 + O(e^{-c\alpha})}{\pi(y)\pi(z)}\mathbf{P}_y[T_\alpha^U > \tau(A) \geq T_\alpha^U - T_{\alpha'}^U, X(T_\alpha^U) = z]\pi(y).\end{aligned}$$

By reversing time, we see that this is equal to

$$\begin{aligned}
&\frac{1+O(e^{-c\alpha})}{\pi(y)}\mathbf{P}_z[\tau(A)\le T^U_{\alpha'}, d(X(t),A)>r \text{ for all } T^U_{\alpha'}<t\le T^U_\alpha, X(T^U_\alpha)=y]\\
&\quad\le\frac{1+O(e^{-c\alpha})}{\pi(y)}\mathbf{P}_z[X(T^U_\alpha)=y|\tau(A)\le T^U_{\alpha'}]\mathbf{P}_z[\tau(A)\le T^U_{\alpha'}]\\
&\quad=\bigl(1+O\bigl(e^{-c(\alpha-\alpha')}\bigr)\bigr)\mathbf{P}_z[\tau(A)\le T^U_{\alpha'}].
\end{aligned}$$

A union bound along with uniform local transience implies this is of order $O(|A|\rho(s,r)+o(1))$. With $\mathcal{A}_1=\{d(X(T^U_\alpha),A)\ge s\}$,

$$\begin{aligned}
\mathbf{P}_y[\mathcal{A}]&=\mathbf{P}_y[\tau(A)\ge T^U_\alpha,\mathcal{A}_1]\\
&=\bigl(\mathbf{P}_y[\tau(A)\ge T^U_\alpha-T^U_{\alpha'}|\mathcal{A}_1]-\mathbf{P}_y[T^U_\alpha>\tau(A)\ge T^U_\alpha-T^U_{\alpha'}|\mathcal{A}_1]\bigr)\mathbf{P}_y[\mathcal{A}_1]\\
&=\mathbf{P}_y[\tau(A)\ge T^U_\alpha-T^U_{\alpha'}]+O\bigl(e^{-c(\alpha-\alpha')}+|A|\rho(s,r)+o(1)\bigr),
\end{aligned}$$

the last line coming from a similar analysis as before. Consequently,

$$\frac{\mathbf{P}_y[\tau(A)\ge T^U_\alpha|X(T^U_\alpha)=z]}{\mathbf{P}_y[\mathcal{A}]}=1+O\bigl(e^{-c(\alpha-\alpha')}+|A|\rho(s,r)+o(1)\bigr).$$

Taking $\alpha'=\alpha/2$ gives the lemma. $\square$

Let $\tau_k(A)=\tau_k^0(A)$, $\sigma_k(A)=\sigma_k^0(A)$, and $T_{r,R}(A)=\mathbf{E}_\pi[\tau_1(A)-\tau_0(A)]$. We will now show that mean excursion length does not depend too strongly on the starting point of X. The idea is to argue that X will typically run for some multiple of the mixing time before getting close to A provided it is initialized sufficiently far away from A, then invoke the previous lemma to replace the induced law on V by π.

LEMMA 4.4 (Mean excursion length). *For every $r,\delta>0$ there exists $R_0>r$ such that $R\ge R_0$ implies*

$$(1-\delta)T_{r,R}(A)\le\min_{y\notin A(R)}\mathbf{E}_y\tau_0(A)\le\max_{y\notin A(R)}\mathbf{E}_y\tau_0(A)\le(1+\delta)T_{r,R}(A)$$

for all n large enough.

PROOF. We have that

$$\mathbf{E}_\pi[\tau_1(A)-\tau_0(A)]=\mathbf{E}_\pi[\sigma_0(A)-\tau_0(A)]+\mathbf{E}_\pi[\tau_1(A)-\sigma_0(A)].$$

Obviously,

$$\mathbf{E}_\pi[\sigma_0(A)-\tau_0(A)]\le\max_{y\in A(r)}\mathbf{E}_y\sigma_0(A)\le cT^U_{\mathrm{mix}}(G)$$

for some $c > 0$ since in each interval of length $T^U_{\mathrm{mix}}(G)$, random walk started in $A(r)$ has probability uniformly bounded from below of leaving $A(R)$ provided n is large enough. It is also obvious that

$$\min_{y \notin A(R)} \mathbf{E}_y \tau_0(A) \le \mathbf{E}_\pi[\tau_1(A) - \sigma_0(A)] \le \max_{y \notin A(R)} \mathbf{E}_y \tau_0(A).$$

The previous lemma implies

$$(1-\delta)\mathbf{E}_\pi[\tau_0(A)] \le \mathbf{E}_y[\tau_0(A)|\mathcal{A}] \le T^U_\alpha + (1+\delta)\mathbf{E}_\pi[\tau_0(A)]$$

for all $y \notin A(R)$ provided we choose R, α, s, n large enough to accommodate our choice of δ. Hence,

$$(1-\delta)\mathbf{E}_\pi[\tau_0(A)] \le \mathbf{E}_y[\tau_0(A)] \le (1+\delta)\mathbf{E}_\pi[\tau_0(A)]$$

as it is not difficult to see that $T^U_{\mathrm{mix}}(G) = o(T_{r,R}(A))$ as $n \to \infty$. Therefore

$$\max_{y_1, y_2 \notin A(R)} \frac{\mathbf{E}_{y_1}\tau_0(A)}{\mathbf{E}_{y_2}\tau_0(A)} \le 1+\delta,$$

which proves the lemma. $\square$

We end with the main result of the subsection, the concentration of the empirical average of excursion lengths. The proof is an adaptation of [10], Lemma 24, to our setting and is based on Kac's moment formula ([17], Equation 6) for the first hitting time of a strong Markov process along with the approximate i.i.d. structure of excursion lengths.

LEMMA 4.5 (Concentration of excursions). *For each $\beta \ge 0$ and $r, \delta > 0$ there exists $R_0 > r$ such that*

$$\mathbf{P}_y[\tau^\beta_k(A) \le (1-\delta)T_{r,R}(A)k] \le e^{-C\delta^2 k}, \tag{4.13}$$

$$\mathbf{P}_y[\tau^\beta_k(A) \ge (1+\delta)T_{r,R}(A)k] \le e^{-C\delta^2 k} \tag{4.14}$$

for all $R \ge R_0$, $y \in V$ and n large enough.

PROOF. First of all, it follows from Lemma 4.4 that

$$\max_y \mathbf{E}_y[\tau_0(A)] \le C T_{r,R}(A)$$

for some $C > 0$ provided R, n are sufficiently large. Consequently, Kac's moment formula (see [17], Equation 6) for the first hitting time of a strong Markov process implies for any $j \in \mathbf{N}$ we have that

$$\max_y \mathbf{E}_y[(\tau_0(A))^j] \le j! c^j T^j_{r,R}(A) \tag{4.15}$$

for some $c > 0$. This implies that there exists $\lambda_0 > 0$ so that

$$\max_y \mathbf{E}_y \exp[\lambda\tau_0(A)/T_{r,R}(A)] < \infty \qquad \text{for all } \lambda \in (0, \lambda_0).$$

Using $\mathbf{E}[\sigma_0(A) - \tau_0(A)] = o(T_{r,R}(A))$, a similar argument implies that, by possibly decreasing λ_0,

$$\max_y \mathbf{E}_y \exp[\lambda\sigma_0(A)/T_{r,R}(A)] < \infty \qquad \text{for all } \lambda \in (0, \lambda_0).$$

Combining the strong Markov property with $T_\beta^U = o(T_{r,R}(A))$ yields

$$\max_y \mathbf{E}_y \exp[\lambda\tau_k^\beta(A)/T_{r,R}(A)] < \infty \qquad \text{for all } \lambda \in (0, \lambda_0).$$

Let R_0 be large enough so that the previous lemma implies

$$(1-\delta/2)T_{r,R}(A) \le \min_{y \notin A(R)} \mathbf{E}_y \tau_0(A) \le \max_{y \notin A(R)} \mathbf{E}_y \tau_0(A) \le (1+\delta/2)T_{r,R}(A)$$

for $R \ge R_0$ and n large enough. We compute

$$\begin{aligned}
\max_{y \notin A(R)} \mathbf{E}_y e^{-\theta\tau_0(A)} &\le 1 - \theta \min_{y \notin A(R)} \mathbf{E}_y \tau_0(A) + \theta^2 \max_{y \notin A(R)} \mathbf{E}_y \tau_0^2(A) \\
&\le 1 - \theta(1-\delta/2)T_{r,R}(A) + \zeta\theta^2 \\
&\le \exp\bigl(\zeta\theta^2 - \theta(1-\delta/2)T_{r,R}(A)\bigr),
\end{aligned}$$

where $\zeta = cT_{r,R}^2(A)$ for some $c > 0$. Since $\tau_0(A) \ge 0$, Chebychev's inequality leads to (4.13). Indeed,

$$\begin{aligned}
&\mathbf{P}_y[\tau_k^\beta(A) \le (1-\delta)T_{r,R}(A)k] \\
&\quad \le \exp\bigl(\theta(1-\delta)T_{r,R}(A)k\bigr)\mathbf{E}_y e^{-\theta\tau_k^\beta(A)} \\
&\quad \le \exp\bigl(\theta(1-\delta)T_{r,R}(A)k\bigr)\Bigl[\max_{y \notin A(R)} \mathbf{E}_y e^{-\theta\tau_0(A)}\Bigr]^k \\
&\quad \le \exp\bigl(\theta(1-\delta)T_{r,R}(A)k\bigr)\exp\bigl(\zeta\theta^2 k - \theta(1-\delta/2)T_{r,R}(A)k\bigr).
\end{aligned}$$

Taking

$$\theta = \frac{\delta T_{r,R}(A)}{c_1\zeta}$$

we get that

$$\begin{aligned}
&\mathbf{P}_y[\tau_k^\beta(A) \le (1-\delta)T_{r,R}(A)k] \\
&\quad \le \exp\bigl(\zeta\theta^2 k - \theta T_{r,R}(A)k\delta/2\bigr) \\
&\quad \le \exp\bigl(\zeta\delta^2 T_{r,R}^2(A)k/(c_1^2\zeta^2) - \delta^2 T_{r,R}^2(A)k/(2c_1\zeta)\bigr) \\
&\quad \le \exp(-c\delta^2 k)
\end{aligned}$$

provided we take c_1 sufficiently large.

To prove (4.14), we need to bound

$$\mathbf{P}_y[\tau_k^\beta(A) \geq (1+\delta)T_{r,R}(A)k]$$
$$\leq \exp\bigl(-\theta(1+\delta)T_{r,R}(A)k\bigr)\Bigl(e^{\theta T_\beta^U} \max_y \mathbf{E}_y e^{\theta\tau_0(A)} \max_{y\in A(r)} \mathbf{E}_y e^{\theta[\sigma_0(A)-\tau_0(A)]}\Bigr)^k.$$

We again take

$$\theta = \frac{\delta T_{r,R}(A)}{c_1\zeta}$$

with c_1 to be fixed shortly, and note that

$$\max_y \mathbf{E}_y e^{\theta\tau_0(A)} \leq \bigl(1+o(1)\bigr) \max_{y\notin A(R)} \mathbf{E}_y e^{\theta\tau_0(A)}$$
$$\leq \exp\bigl(\theta(1+\delta/2)T_{r,R}(A) + c_2\zeta\theta^2 + o(1)\bigr).$$

Since $\max_{y\in A(r)} \mathbf{E}_y[\sigma_0(A) - \tau_0(A)] = o(T_{r,R}(A))$ as $n\to\infty$, Kac's formula yields

$$\max_{y\in A(r)} \mathbf{E}_y e^{\theta[\sigma_0(A)-\tau_0(A)]} = 1 + o(1) \qquad \text{as } n\to\infty.$$

Since $T_\beta^U = o(T_{r,R}(A))$ as $n\to\infty$ as well, we have

$$\mathbf{P}_y[\tau_k^\beta(A) \geq (1+\delta)T_{r,R}(A)k]$$
$$\leq \exp\bigl(-\theta(1+\delta)T_{r,R}(A)k + \theta(1+\delta/2)T_{r,R}(A)k + c_2\zeta\theta^2 k + o(1)k\bigr)$$
$$\leq \exp\bigl(-\theta\delta T_{r,R}(A)k/2 + c_2\zeta\theta^2 k + o(1)k\bigr).$$

Taking $c_1 > 0$ large enough gives the result. $\square$

4.3. *Hitting and covering*. The purpose of this subsection is to estimate the maximal hitting time (Lemma 4.6) and cover time (Lemma 4.8).

LEMMA 4.6 (Hitting time estimate). *For every $\delta > 0$ there exists r_0 such that for each $r \geq r_0$ there is an $R_0 > r$ so that if $R \geq R_0$ the following holds. If $A_n = \{x_{n1}, \ldots, x_{n\ell}\} \subseteq V_n$ with $d(x_{ni}, x_{nj}) \geq 2R$ for $i \neq j$ and $y_n \in V_n$ is such that $d(x_{ni}, y_n) \geq 2R$ for all n, then*

$$1-\delta \leq \liminf_{n\to\infty} \frac{\mathbf{E}_{y_n}\tau(x_{ni})}{T_{r,R}(A_n)/\overline{p}_{r,R}(x_{ni};A)} \tag{4.16}$$

$$\leq \limsup_{n\to\infty} \frac{\mathbf{E}_{y_n}\tau(x_{ni})}{T_{r,R}(A_n)/\overline{p}_{r,R}(x_{ni};A)} \leq 1+\delta. \tag{4.17}$$

As the proof of the lemma is long, we pause momentarily to highlight the main steps. The primary tools will be the results from the previous subsections. The first ingredient (though we leave this to the end of the proof) is to argue that it is unlikely for X to hit a point $x_{nk} \in A_n$ in the "remixing" intervals $[\sigma_k^\beta(A) + T_\alpha^U, \sigma_k^\beta(A) + T_\beta^U]$. Once we have established this, it suffices to estimate the expectation of the first time $\widetilde{\tau}(x_{nk})$ that X hits x_{nk} in $\bigcup_k [\tau_k^\beta(A), \sigma_k^\beta(A) + T_\alpha^U]$ in place of the expectation of $\tau(x_{nk})$. In particular, this implies that the probability that x_{nk} is first hit by the $(j+1)$st excursion is well approximated by

$$\mathbf{E}_{y_n}\left[p_{j+1}^{\alpha,\beta}(x_{nk}; A_n) \prod_{i=1}^{j} \bigl(1 - p_i^{\alpha,\beta}(x_{nk}; A_n)\bigr)\right].$$

We now apply the concentration of the empirical mean of the $p_j^{\alpha,\beta}(x; A)$ proved in Lemma 4.2 in order to replace the product with $\exp(-(1 + O(\rho(r))) j \overline{p}_{r,R}(x_{nk}; A_n))$, where we recall that ρ is the transience function. We conclude that the mean number of excursions required to hit x_{nk} is approximately $1/\overline{p}_{r,R}(x_{nk}; A)$. The result now follows by invoking Lemma 4.5.

PROOF OF LEMMA 4.6. We will omit the indices n and i and just write x for x_{ni}, y for y_n and A for A_n. Fix r sufficiently large so that $\rho(r) < \delta^2/100$. Recall that $S_k^{\alpha,\beta}(x; A)$ is the event that X hits x in $[\tau_k^\beta(A), \sigma_k^\beta(A) + T_\alpha^U]$ where $\tau_k^\beta(A), \sigma_k^\beta(A)$ are as in (4.1)–(4.4). Let $N(x; A) = \min\{k \ge 1 : S_k^{\alpha,\beta}(x; A) \text{ occurs}\}$ and let

$$\widetilde{\tau}(x) = \min\{t \ge 0 : X(t) = x \text{ and } t \in I\},$$

where

$$I_k = [\tau_k^\beta(A), \sigma_k^\beta(A) + T_\alpha^U] \quad \text{and} \quad I = \bigcup_k I_k.$$

Then

$$\tau_{N(x;A)}^\beta(A) \le \widetilde{\tau}(x) \le \tau_{N(x;A)+1}^\beta(A).$$

Let

$$W(M; \delta) = \bigcap_{j \ge M} B(j; \delta) \equiv \bigcap_{j \ge M} \{(1-\delta) T_{r,R}(A) j \le \tau_j^\beta(A) \le (1+\delta) T_{r,R}(A) j\}.$$

With $\|\widetilde{\tau}(x)\| = \max_z \mathbf{E}_z \widetilde{\tau}(x)$, note that

$$\begin{aligned}
\mathbf{E}_y \widetilde{\tau}(x) \mathbf{1}_{W^c(M;\delta)} &\le \sum_{j \ge M} \mathbf{E}_y \widetilde{\tau}(x) \mathbf{1}_{B^c(j;\delta)} \\
&\le \sum_{j \ge M} \bigl[\mathbf{E}_y \tau_j^\beta(x) \mathbf{1}_{B^c(j;\delta)} + \|\widetilde{\tau}(x)\| \mathbf{P}_y[B^c(j;\delta)]\bigr] \\
&\le 2C_0 \sum_{j \ge M} [j T_{r,R}(A) + \|\widetilde{\tau}(x)\|] e^{-C\delta^2 j} \qquad (4.18) \\
&\le C_1 \|\widetilde{\tau}(x)\| \sum_{j \ge M} (1+j) e^{-C\delta^2 j} \le C_2 \|\widetilde{\tau}(x)\| \frac{e^{-C\delta^2 M}}{\delta^4}.
\end{aligned}$$

To see the second step, we let

$$\widetilde{\tau}_j(x) = \min\{t \geq \tau_j^\beta(x) : X(t) = x\}.$$

Then we have that

$$\begin{aligned}\mathbf{E}_y \widetilde{\tau}(x) \mathbf{1}_{B^c(j;\delta)} &\leq \mathbf{E}_y \widetilde{\tau}_j(x) \mathbf{1}_{B^c(j;\delta)} = \mathbf{E}_y\big[\big(\tau_j^\beta(x) + \big(\widetilde{\tau}_j(x) - \tau_j^\beta(x)\big)\big)\mathbf{1}_{B^c(j;\delta)}\big] \\ &\leq \mathbf{E}_y \tau_j^\beta(x)\mathbf{1}_{B^c(j;\delta)} + \mathbf{E}_y\big[\big(\widetilde{\tau}_j(x) - \tau_j^\beta(x)\big)|B^c(j;\delta)\big]\mathbf{P}_y[B^c(j;\delta)].\end{aligned}$$

By the strong Markov property, $\mathbf{E}_y[\widetilde{\tau}_j(x) - \tau_j^\beta(x)|B^c(j;\delta)] \leq \|\widetilde{\tau}(x)\|$. In the third step, we used that

$$\begin{aligned}\mathbf{E}_y \tau_j^\beta(A)\mathbf{1}_{B^c(j;\delta)} &\leq (\mathbf{E}_y[\tau_j^\beta(A)]^2)^{1/2}(\mathbf{P}_y[B^c(j;\delta)])^{1/2} \\ &\leq \frac{2T_{r,R}(A)}{\lambda} j\big(\mathbf{E}_y \exp\big(\lambda\tau_j^\beta(A)/(jT_{r,R}(A))\big)\big)^{1/2} Ce^{-C\delta^2 j},\end{aligned}$$

where $\lambda \in (0, \lambda_0)$, λ_0 as in the proof of Lemma 4.5. We used in the fourth step that $T_{r,R}(A) = O(\|\widetilde{\tau}(x)\|)$. Indeed, this is true since uniform local transience implies that with uniformly positive probability more than one excursion is required to hit x and, by Lemma 4.4, the mean length of the second excursion is at least $\frac{1}{2}T_{r,R}(A)$. The final step in (4.18) comes from summing the geometric series. Uniform local transience implies

$$\big|\mathbf{E}_y\widetilde{\tau}(x) - \|\widetilde{\tau}(x)\|\big| \leq \delta\mathbf{E}_y\widetilde{\tau}(x), \tag{4.19}$$

when R is large enough. Consequently, there exists $M > 0$ large enough depending only on δ so that

$$\mathbf{E}_y\widetilde{\tau}(x)\mathbf{1}_{W(M;\delta)} \leq \mathbf{E}_y\widetilde{\tau}(x) \leq (1+\delta)\mathbf{E}_y\widetilde{\tau}(x)\mathbf{1}_{W(M;\delta)}.$$

Now,

$$\begin{aligned}\mathbf{E}_y\tau^\beta_{N(x;A)+1}(A)\mathbf{1}_{W(M;\delta)} &= \mathbf{E}_y\left[N(x;A)\left(\frac{\tau^\beta_{N(x;A)+1}(A)}{N(x;A)}\right)\mathbf{1}_{W(M;\delta)}\right] \\ &\leq (1+\delta)T_{r,R}(x)\mathbf{E}_yN(x;A) + \mathbf{E}_y\tau_M^\beta(A) \\ &\leq (1+\delta)T_{r,R}(A)\mathbf{E}_yN(x;A) + CMT_{r,R}(A).\end{aligned}$$

In order to derive the inequality, we used that if $N(x;A) \geq M$ then by the definition of $W(M;\delta)$ we have $\tau^\beta_{N(x;A)+1}(A)/N(x;A) \leq (1+\delta)T_{r,R}(A)$ and, in case $N(x;A) < M$, we clearly have that $\tau^\beta_{N(x;A)+1}(A) \leq \tau^\beta_M(A)$. The final inequality is a consequence of Lemma 4.4. Similarly, we also have

$$\mathbf{E}_y\tau_{N(x;A)}(A)\mathbf{1}_{W(M;\delta)} \geq (1-\delta)T_{r,R}(A)\mathbf{E}_yN(x;A).$$

Therefore,

$$(1-\delta)T_{r,R}(A)\mathbf{E}_yN(x;A) \leq \mathbf{E}_y\widetilde{\tau}(x) \leq (1+2\delta)T_{r,R}(A)\mathbf{E}_yN(x;A) + CMT_{r,R}(A).$$

By Lemma 4.2,

$$\begin{aligned}
&\overline{p}_{r,R}(x;A)\Big[\exp\big(-(1+\delta)\overline{p}_{r,R}(x;A)j\big) - C\exp\Big(-\frac{C\delta^2\overline{p}_{r,R}(x;A)}{\rho(r)}j\Big)\Big]\\
&\quad\le \mathbf{E}_y p_{j+1}^{\alpha,\beta}(x;A)\exp\Bigg(-[1+O(\rho(r))]\sum_{i=1}^{j} p_i^{\alpha,\beta}(x;A)\Bigg)\\
&\quad\le \overline{p}_{r,R}(x;A)\Big[\exp\big(-(1-\delta)\overline{p}_{r,R}(x;A)j\big) + C\exp\Big(-\frac{C\delta^2\overline{p}_{r,R}(x;A)}{\rho(r)}j\Big)\Big].
\end{aligned}$$

Taking r sufficiently large gives

$$\begin{aligned}
&\mathbf{E}_y N(x;A)\\
&\quad= \sum_{j=1}^{\infty} j\mathbf{P}[N(x;A)=j]\\
&\quad\le CM^2\rho(r) + \sum_{j=M+1}^{\infty} j\big(1+o(1)\big)\big(\overline{p}_{r,R}(x;A)\exp\big(-(1-\delta)\overline{p}_{r,R}(x;A)j\big)\big)\\
&\quad\le 2CM^2\rho(r) + \frac{1+\delta}{\overline{p}_{r,R}(x;A)}.
\end{aligned}$$

Similarly,

$$\mathbf{E}_y N(x;A) \ge \frac{1-\delta}{\overline{p}_{r,R}(x;A)}.$$

Increasing r if necessary so that $M^2\rho(r)\le\delta$ yields

$$\frac{1-2\delta}{\overline{p}_{r,R}(x;A)} \le \mathbf{E}_y N(x;A) \le \frac{1+2\delta}{\overline{p}_{r,R}(x;A)}. \tag{4.20}$$

This proves that

$$\mathbf{E}_y\widetilde{\tau}(x) = \big(1+o(1)\big)\frac{T_{r,R}(A)}{\overline{p}_{r,R}(x;A)} \qquad \text{as } n\to\infty.$$

Let F_k be the event that X hits $A(r)$ in $J_k = [\sigma_k^{\beta}(A)+T_{\alpha}^{U}, \sigma_k^{\beta}(A)+T_{\beta}^{U}]$. With $F=\bigcup_{k=1}^{N(x;A)+1} F_k$, we have

$$\mathbf{E}_y\widetilde{\tau}(x)\mathbf{1}_{F^c} \le \mathbf{E}_y\tau(x) \le \mathbf{E}_y\widetilde{\tau}(x),$$

where we recall that $\tau(x)$ is the first time X hits x.

We now claim that

$$\mathbf{E}_y\widetilde{\tau}(x)\mathbf{1}_{F^c} = \Bigg[1+O\Bigg(\frac{T_{\beta}^{U}|A|\overline{\Delta}^{r}(G)}{|V|\overline{p}_{r,R}^{2}(x;A)} + \overline{p}_{r,R}^{2}(x;A)\Bigg)^{1/2}\Bigg]\mathbf{E}_y\widetilde{\tau}(x). \tag{4.21}$$

Note that this will complete the proof of the lemma as $\overline{p}_{r,R}(x;A) \geq C\underline{\Delta}^{-r}(G)$ so that, by Assumption 1.1, the error term can be made as small as we like by making r, R large enough. Using the Kac moment formula ([17], Equation 6) in the second inequality, we trivially have

$$\begin{aligned}\mathbf{E}_y\widetilde{\tau}(x)\mathbf{1}_F &\leq \mathbf{E}_y\tau(x)\mathbf{1}_F + \|\widetilde{\tau}(x)\|\mathbf{P}[F] \\ &\leq C_1\|\tau(x)\|\sqrt{\mathbf{P}[F]} + \|\widetilde{\tau}(x)\|\mathbf{P}[F].\end{aligned} \tag{4.22}$$

In view of (4.19) we have $\|\tau(x)\| \leq \|\widetilde{\tau}(x)\| \leq (1+\delta)\mathbf{E}_y\widetilde{\tau}(x)$. Thus, using $\mathbf{P}[F] \leq \sqrt{\mathbf{P}[F]}$, we see that we can bound (4.22) from above by $C_2\|\widetilde{\tau}(x)\|\sqrt{\mathbf{P}[F]}$. Using exactly the same proof of (4.20), we have that

$$\mathbf{E}_y[N^2(x;A)] \leq \frac{C_3}{\overline{p}_{r,R}^2(x;A)}. \tag{4.23}$$

Applying (4.23) along with Markov's inequality in the second step, we consequently have

$$\begin{aligned}\mathbf{P}_y[F] &\leq \mathbf{P}_y[F, N(x;A)+1 \leq 1/(\overline{p}_{r,R}(x;A))^2] \\ &\quad + \mathbf{P}[N(x;A)+1 \geq 1/(\overline{p}_{r,R}(x;A))^2] \\ &\leq \sum_{k=1}^{1/(\overline{p}_{r,R}(x;A))^2} \mathbf{P}_y[F_k] + O((\overline{p}_{r,R}(x;A))^2).\end{aligned}$$

Since $|A(r)| \leq |A|\overline{\Delta}^r(G)$, a union bound implies $\mathbf{P}_y[F_k] = O(T_\beta^U|A|\overline{\Delta}^r(G)/|V|)$, which proves (4.21). $\square$

If G were vertex transitive so that $\overline{p}_{r,R}(x)$ and $T_{r,R}(x)$ did not depend on x, then by the Matthews method ([21]; see also Theorem 11.2 and Proposition 11.4 of [20]) it is possible to deduce that $T_{\mathrm{cov}}(G)$ is asymptotically well approximated by $T_{r,R}/\overline{p}_{r,R}\log|V|$. Our goal is to prove something similar even if G is not vertex transitive. The idea of the proof will be to group vertices together based on their hitting time $T_{r,R}(x)/\overline{p}_{r,R}(x)$. In particular, we will argue that the amount of time it takes to cover a set $V_F \subseteq V$ of vertices each of whose hitting time is close T_F is approximately $T_F\log|V_F|$. The cover time of G is then well approximated by $\max_F T_F\log|V_F|$ where F ranges over subsets of vertices with approximately constant hitting time.

The first step in implementing this strategy is to show that if we want to estimate $T_{\mathrm{cov}}(G)$ to a multiple of $\varepsilon T_{\mathrm{cov}}(G)$, $\varepsilon > 0$ fixed, we only need to consider a finite number, depending only on ε, of groups of vertices. This will be accomplished by relating $\overline{p}_{r,R}(x)/T_{r,R}(x)$ to $\pi(x)$ and then invoking Assumption 1.1.

We will now specialize to the case $A = \{x\}$; for simplicity of notation we will omit A. Let

$$O_{r,R}(x) = \frac{\overline{a}_{r,R}(x)}{T_{r,R}(x)}.$$

LEMMA 4.7. *For every $\delta > 0$, there exists r_0 such that if $r \geq r_0$ there is $R_0 > r$ such that $R \geq R_0$ implies*

$$(1-\delta)\pi(x) \leq O_{r,R}(x) \leq (1+\delta)\pi(x)$$

for all n large enough.

PROOF. Let $N(x,T) = \min\{k : \tau_k^\beta(x) \geq T\}$, J_k as in the previous lemma, $J = \bigcup_k J_k$ and $\mathcal{G}(x) = \sigma(X(\tau_j^\beta(x)) : j \geq 1)$. Then

$$\sum_{j=1}^{N(x,T)} a_j^{\alpha,\beta}(x) \leq \mathbf{E}\left[\sum_{t=1}^{T} \mathbf{1}_{\{X(t)=x\}} \mathbf{1}_{\{t \notin J\}} \Big| \mathcal{G}(x)\right] \leq \sum_{j=1}^{N(x,T)+1} a_j^{\alpha,\beta}(x).$$

Lemmas 4.2 and 4.5 give that

$$(1-\delta)T_{r,R}(x) \leq \frac{N(x,T)}{T} \leq (1+\delta)T_{r,R}(x)$$

and

$$(1-\delta)\overline{a}_{r,R}(x) \leq \frac{\sum_{j=1}^{k} a_j^{\alpha,\beta}(x)}{k}(1+\delta)\overline{a}_{r,R}(x)$$

with high probability as $T \to \infty$, for all $r, R, k, n, \beta - \alpha$ large enough. Consequently, using that $(a_j^{\alpha,\beta}(x) : j \geq 1)$ is uniformly bounded, it is not hard to see that

$$(1-\delta)\frac{\overline{a}_{r,R}(x)}{T_{r,R}(x)} \leq \frac{1}{T}\sum_{j=1}^{N(x,T)} a_j^{\alpha,\beta}(x) \leq (1+\delta)\frac{\overline{a}_{r,R}(x)}{T_{r,R}(x)}$$

with high probability as $T \to \infty$, for all $r, R, n, \beta - \alpha$ large enough. The middle term converges to $\pi(x)$ as $T \to \infty$ since

$$\lim_{T\to\infty} \frac{1}{T}\mathbf{E}\sum_{t=1}^{T} \mathbf{1}_{\{X(t)\in A(r)\}} \mathbf{1}_{\{t\in J\}} = 0. \qquad \square$$

Uniform local transience implies that there exists constants $c, C > 0$ so that $c\overline{a}_{r,R}(x) \leq \overline{p}_{r,R}(x) \leq C\overline{a}_{r,R}(x)$; combining this with the previous lemma yields

$$\frac{c\deg(x)}{|E|} \leq \frac{\overline{p}_{r,R}(x)}{T_{r,R}(x)} \leq \frac{C\deg(x)}{|E|}.$$

Let $\varepsilon > 0$ and let

$$H_{n,k}^{\varepsilon} = \left\{x \in V_n : \frac{\Delta(G_n)k\varepsilon}{|E_n|} < \frac{\overline{p}_{r,R}(x)}{T_{r,R}(x)} \leq \frac{\Delta(G_n)(k+1)\varepsilon}{|E_n|}\right\}$$

be a partition of V_n into at most $\Delta_0\varepsilon^{-1}$ subsets, where Δ_0 is the constant from Assumption 1.1. By passing to a subsequence, we may assume without loss of generality that

$$d_k^\varepsilon = \lim_{n\to\infty} d_{n,k}^\varepsilon \equiv \lim_{n\to\infty} \frac{\log|H_{n,k}^\varepsilon|}{\log|V_n|}$$

exists for every k. Note that $d_k^\varepsilon \in [0,1]$ for those k so that $|H_{n,k}^\varepsilon| \neq 0$ for all n large enough and, since the partition is finite, necessarily there exists k so that $d_k^\varepsilon = 1$. In particular, there exists k so that $d_k^\varepsilon \neq 0$. Let

$$C_{n,k}^\varepsilon = \frac{|E_n|}{\underline{\Delta}(G_n)k\varepsilon} d_k^\varepsilon \log|V_n| \quad \text{and} \quad C_n^\varepsilon = \max_k C_{n,k}^\varepsilon. \tag{4.24}$$

LEMMA 4.8 (Cover time estimate). *For each $\delta > 0$, there exists r_0, ε_0 so that if $r \geq r_0$ there is $R_0 > r$ such that $R \geq R_0$ and $\varepsilon \in (0, \varepsilon_0)$ implies*

$$1 - \delta \leq \liminf_{n\to\infty} \frac{T_{\mathrm{cov}}(H_{n,k}^\varepsilon)}{C_{n,k}^\varepsilon} \leq \limsup_{n\to\infty} \frac{T_{\mathrm{cov}}(H_{n,k}^\varepsilon)}{C_{n,k}^\varepsilon} \leq 1 + \delta \tag{4.25}$$

for all k with $d_k^\varepsilon > 0$. Furthermore,

$$1 - \delta \leq \liminf_{n\to\infty} \frac{T_{\mathrm{cov}}(G_n)}{C_n^\varepsilon} \leq \limsup_{n\to\infty} \frac{T_{\mathrm{cov}}(G_n)}{C_n^\varepsilon} \leq 1 + \delta. \tag{4.26}$$

PROOF. Suppose k is such that $d_k^\varepsilon > 0$. Then $|H_{n,k}^\varepsilon| \to \infty$ as $n \to \infty$. Let $r, R, n > 0$ be sufficiently large so that Lemma 4.6 applies with our choice of δ. By Assumption 1.1(1) we have that $\log|B(x,r)| = o(\log|V_n|)$. Consequently, for all n large enough there exists an R-net $E_{n,k}^\varepsilon$ of $H_{n,k}^\varepsilon$ such that

$$\log|E_{n,k}^\varepsilon| = \log|H_{n,k}^\varepsilon| + o(1) \qquad \text{as } n \to \infty.$$

The upper and lower bounds from the Matthews method ([21]; see also Theorem 11.2 and Proposition 11.4 of [20]) combined with the definition of $C_{n,k}^\varepsilon$ imply (4.25). Theorem 2 of [4] implies that

$$\lim_{n\to\infty} \frac{\tau_{\mathrm{cov}}(H_{n,k}^\varepsilon)}{\mathbf{E}\tau_{\mathrm{cov}}(H_{n,k}^\varepsilon)} = 1.$$

As $\tau_{\mathrm{cov}}(G_n) = \max_k \tau_{\mathrm{cov}}(H_{n,k}^\varepsilon)$ and the maximum is over a finite set, it follows that $\tau_{\mathrm{cov}}(G_n) = (1 + o(1)) \max_k T_{\mathrm{cov}}(H_{n,k}^\varepsilon)$. Taking expectations of both sides gives (4.26). $\square$

5. Correlation decay. The purpose of this section is to prove Theorem 1.6. Exactly the same proof will also yield Lemma 5.1, a technical result which will be useful in the next section, which is stated after the proof. Note that vertex transitivity implies $\overline{p}_{r,R}(\cdot)$ and $T_{r,R}(\cdot)$ do not depend on their arguments.

PROOF OF THEOREM 1.6. First, assume that we are in the case of bounded maximal degree. Let A be as in the previous section and let $\delta > 0$ be arbitrary. Fix r so that $\rho(r) \leq \delta^3/100C\ell$ where $\ell = |A|$ and $\overline{p}_{r,R}(x; A) \leq \delta^3$ for all $x \in A$. Let $R_0 > r$ and $\beta - \alpha$ be sufficiently large so that Lemmas 4.2 and 4.5 apply with our choice of δ, r. Finally, let $N(x_i; A) = \min\{k : S_k^{\alpha,\beta}(x_i; A) \text{ occurs}\}$ and $\mathcal{G}(A) = \sigma(p_j^{\alpha,\beta}(x; A) : x \in A, j \geq 1)$. Since $d(x_i, x_j) \geq 2R$, the probability that X neither hits x nor x' in the interval $[\tau_j^{\beta}(x; A), \sigma_j^{\beta}(x; A) + T_\alpha^U]$ is

$$1 - [1 + O(\rho(R))][p_j^{\alpha,\beta}(x; A) + p_j^{\alpha,\beta}(x'; A)]. \tag{5.1}$$

Indeed, the reason for this is that the conditional probability X hits $B(x', R)$ in the same excursion that it hits x given that it hits the latter first is $O(\rho(R))$ and the probability that X hits x before $B(x', R)$ is trivially bounded by $p_j^{\alpha,\beta}(x; A)$. This holds more generally for any subset of A, hence

$$\begin{aligned}
&\mathbf{E}\big[\mathbf{P}[N(x_1; A) > k_1, \dots, N(x_\ell; A) > k_\ell | \mathcal{G}(A)]\big] \\
&\quad = \mathbf{E} \prod_{i=1}^{\ell} \exp\left(-[1 + O(\rho(R))] \sum_{j=1}^{k_i} p_j^{\alpha,\beta}(x_i; A) \right) \\
&\quad = \exp\left(-[1 + O(\delta)] \sum_{i=1}^{\ell} \overline{p}_{r,R}(x_i; A) k_i \right) \\
&\qquad + \sum_{i=1}^{\ell} O\big(\exp(-\overline{p}_{r,R}(x_i; A) k_i/\delta)\big),
\end{aligned} \tag{5.2}$$

where the last equality followed from our choice of r and Lemma 4.2. Let $J_k = [\sigma_k^{\beta}(A) + T_\alpha^U, \sigma_k^{\beta}(A) + T_\beta^U]$, as the in the previous section. Combining this with Lemma 4.5 and that the probability X hits $A(r)$ in J_k is at most $O(T_\beta^U |A| \overline{\Delta}^r(G)/|V|) = o(\overline{p}_{r,R}(x; A))$ for any $x \in A$, we have

$$\begin{aligned}
&\mathbf{P}[\tau(x_1) \geq k T_{r,R}(A)/\overline{p}_{r,R}(x_1; A), \dots, \tau(x_\ell) \geq k T_{r,R}(A)/\overline{p}_{r,R}(x_n; A)] \\
&\quad = \big(1 + o(1)\big) \exp\big(-[1 + O(\delta)]\ell k\big) + O\big(\exp\big(-C\delta^2 k/\rho(r)\big)\big) \\
&\quad = \big(1 + o(1)\big) \exp\big(-[1 + O(\delta)]\ell k\big).
\end{aligned}$$

By vertex transitivity,

$$T_{\text{hit}}(G) = \big(1 + o(1)\big) \frac{T_{r,R}(x_i; A)}{\overline{p}_{r,R}(x_i; A)}.$$

By Lemma 4.8, we know that the cover time is asymptotically $T_{\text{hit}}(G)\log|V|$. Inserting this into (5.2) gives the result for bounded degree.

This proof works also for unbounded degree, but is not quite sufficient for the statement of our theorem since we would like to allow for points in A to be adjacent. There are two parts that break down. First, in Section 4 we proved the concentration of $p_j^{\alpha,\beta}(x; A)$ when $x \in A$ and we also assumed that $x, y \in A$ implies $d(x, y) \geq 2R$. To allow for x, y adjacent, we define

$$p_j^{\alpha,\beta}(y; A) = \mathbf{P}[S_j^{\alpha,\beta}(y; A) | X(\tau_j^\beta(A)), X(\tau_{j+1}^\beta(A))]$$

for $y \in A(r/2)$. It is not difficult to see that for such y, $p_j^{\alpha,\beta}(y; A)$ exhibits nearly the same concentration behavior as for $y \in A$. Second, the estimate (5.1) is no longer good enough since $\rho(1)$ does not decay in n. However, it is not difficult to see that the same probability satisfies the estimate

$$1 - [1 + O(\overline{\Delta}^{-1}(G))][p_j^{\alpha,\beta}(x; A) + p_j^{\alpha,\beta}(x'; A)], \tag{5.3}$$

which suffices since $\overline{\Delta}^{-1}(G_n) \to 0$ as $n \to \infty$. The rest of the proof is the same. $\square$

Vertex transitivity was used only to get that $T_{r,R}(x; A)/\overline{p}_{r,R}(x; A) = (1 + o(1))T_{\text{hit}}(G)$. The same proof works more generally, but leads to more complicated formulas. However, it is not difficult to see that the upper bound takes a very similar form. This result will be especially useful in the next section to show that points which have not been visited by X after time $\frac{1}{2}T_{\text{cov}}(G)$ are typically well separated. Precisely, our estimate is:

LEMMA 5.1. *If (x_n^i) for $1 \leq i \leq \ell$ is a family of sequences with $x_n^i \in H_{n,k(i)}^\varepsilon$ and $|x_n^i - x_n^j| \geq r$ for every n and $i \neq j$,*

$$\mathbf{P}[x_n^i \in \mathcal{L}(\alpha; G_n) \text{ for all } i] \leq (1 + \delta_{r,\ell})|V_n|^{-\ell d_k^\varepsilon \alpha + \delta_{r,\ell}}, \tag{5.4}$$

where $\delta_{r,\ell} \to 0$ as $r \to \infty$ while ℓ is fixed. If $\overline{\Delta}(G_n) \to \infty$ then we take $r = 1$ and $\delta_{1,\ell} = o(1)$ as $n \to \infty$.

6. Total variation bounds. We are now in a position to complete the proof of Theorems 1.3 and 1.5. We will prove the lower bound first since it does not require us to specialize depending on whether (G_n) satisfies part (1) or (2) of Assumption 1.2. As we have explained earlier, the upper bound will be proved by estimating the exponential moment of the set of points not visited by two independent random walks X, X', each run for time $\frac{1}{2}T_{\text{cov}}(G)$. We will use Lemma 5.1 in the proof of Lemma 6.4 to argue that those points $\mathcal{L}$ not visited by X are typically far apart. This will be useful very useful because, as we prove in Section 6.2, the hypothesis of Assumption 1.2 allows us to establish concentration for the empirical average of the conditional probability $q_j(x)$ that excursions between $\partial B(x, r)$ to $\partial B(x, R)$ given both the entry and exit points, where $R > r$ are very large.

6.1. *Lower bound.* We will now prove the lower bound for Theorems 1.3 and 1.5. This is actually just a slight extension of Theorem 4.1 of [22], but we include it for the reader's convenience. Recall from the Introduction that $\mu(\cdot;\alpha,G)$ is the probability measure on $\mathcal{X}(G)=\{f:V\to\{0,1\}\}$ given by first sampling $\mathcal{R}(\alpha;G)$ then setting

$$f(x)=\begin{cases}\xi(x), & \text{if } x\in\mathcal{R}(\alpha;G),\\ 0, & \text{otherwise,}\end{cases}$$

where $(\xi(x):x\in V)$ is a collection of i.i.d. variables such that $\mathbf{P}[\xi(x)=0]=\mathbf{P}[\xi(x)=1]=\frac{1}{2}$ and $\nu(\cdot;G)$ is the uniform measure on $\mathcal{X}(G)$.

LEMMA 6.1 (Lower bound). *For every* $\delta>0$,

$$\lim_{n\to\infty}\|\mu(\cdot;\tfrac{1}{2}-\delta,G_n)-\nu(\cdot;G_n)\|_{\mathrm{TV}}=1.$$

PROOF. For $A\subseteq V$ and $m>0$, let $\tau_{\mathrm{cov}}(A;m)$ be the first time all but m of the vertices of A have been visited by X. For each k such that $d_k^\varepsilon>0$, we will show that

$$\lim_{n\to\infty}\mathbf{P}[\tau_{\mathrm{cov}}(H_{n,k}^\varepsilon;|H_{n,k}^\varepsilon|^\alpha)<(1-\alpha-\delta)C_{n,k}^\varepsilon]=0 \tag{6.1}$$

for each $\delta>0$ and $\varepsilon\in(0,\varepsilon_0(\delta))$. If not, then for some such k,δ,α we have

$$\limsup_{n\to\infty}\mathbf{P}[A_{n,k}(\alpha,\delta)]>0,$$

where

$$A_{n,k}(\alpha,\delta)=\{\tau_{\mathrm{cov}}(H_{n,k}^\varepsilon;|H_{n,k}^\varepsilon|^\alpha)<(1-\alpha-\delta)C_{n,k}^\varepsilon\}.$$

It follows from the Matthews method upper bound ([21]; see also Theorem 11.2 of [20]) that

$$\begin{aligned}&\mathbf{E}[\tau_{\mathrm{cov}}(H_{n,k}^\varepsilon)-\tau_{\mathrm{cov}}(H_{n,k}^\varepsilon;|H_{n,k}^\varepsilon|^\alpha)|A_{n,k}(\alpha,\delta)]\\ &\quad\le\alpha\bigl(1+O(\varepsilon)\bigr)C_{n,k}^\varepsilon\le\alpha(1+\delta/4)C_{n,k}^\varepsilon,\end{aligned}$$

where we take ε so small that the $O(\varepsilon)$ term is at most $\delta/4$. Markov's inequality now implies

$$\mathbf{P}[\tau_{\mathrm{cov}}(H_{n,k}^\varepsilon)<(1-\delta/2)C_{n,k}^\varepsilon|A_{n,k}(\alpha,\delta)]>0.$$

This is a contradiction as Theorem 2 of [4] implies $\tau_{\mathrm{cov}}(H_{n,k}^\varepsilon)/C_{n,k}^\varepsilon\to 1$ in probability.

For each n let $k_0(n)$ be an index that achieves the maximum in $\max_k C_{n,k}^\varepsilon$. Now, (6.1) implies that whp at time $\frac{1}{2}(1-3\delta)T_{\mathrm{cov}}(G_n)=\frac{1}{2}(1-3\delta+O(\varepsilon))C_{n,k_0(n)}^\varepsilon$ the size of the subset of $H_{n,k_0(n)}^\varepsilon$ not visited by X is at least $|H_{n,k_0(n)}^\varepsilon|^{(1+2\delta+O(\varepsilon))/2}$

but less than $|H^{\varepsilon}_{n,k_0(n)}|^{(1+4\delta+O(\varepsilon))/2}$. Thus, the number of zeros in a marking of $H^{\varepsilon}_{n,k_0(n)}$ sampled from $\mu(\cdot;\frac{1}{2}(1-3\delta),G_n)$ is whp at least

$$\tfrac{1}{2}\big|H^{\varepsilon}_{n,k_0(n)}\big| + \big(1+o(1)\big)\big|H^{\varepsilon}_{n,k_0(n)}\big|^{(1+2\delta+O(\varepsilon))/2} \qquad \text{as } n\to\infty.$$

This proves the lemma since the probability of having deviations of this magnitude from the mean tends to zero in a uniform marking. □

6.2. *Concentration of q_j.* Let $\sigma_j(x)=\sigma_j^{0,0}(x)$ and define $\tau_j(x)$ likewise where $\sigma_j^{\alpha,\beta}$, $\tau_j^{\alpha,\beta}$ are as in (4.1)–(4.4). Let $S_j(x)$ be the event that X hits x in the interval $[\tau_j(x),\sigma_j(x)]$ and set $q_j(x)=\mathbf{P}[S_j(x)|X(\tau_j(x)),X(\sigma_j(x))]$. The purpose of this subsection is to study the concentration behavior of $q_j(x)$, which will in turn depend on whether we assume part (1) or (2) of Assumption 1.2; note that $q_j(x)$ differs from $p_j^{\alpha,\beta}(x)$ from Section 4. Indeed, the excursions on which we condition are different since we do not allow the random walk to run for a multiple for $T^U_{\mathrm{mix}}(G)$ after exiting $\partial B(x,R)$ and we condition on the entrance and exit points of the current excursion rather than the entrance points of the current and successive excursion. While both of these changes may seem cosmetic, they affect the concentration behavior, since while $p_j^{\alpha,\beta}(x)$ satisfies (4.6), in locally tree-like graphs it can be that $q_j(x)=1$ with positive probability; see Figure 4 for an illustration of this behavior.

We shall first suppose that (G_n) satisfies Assumption 1.2(1). Let $\varepsilon>0$ be arbitrary, R_n^{γ} be as in Assumption 1.2, $\gamma>0$ to be determined later, and let A be

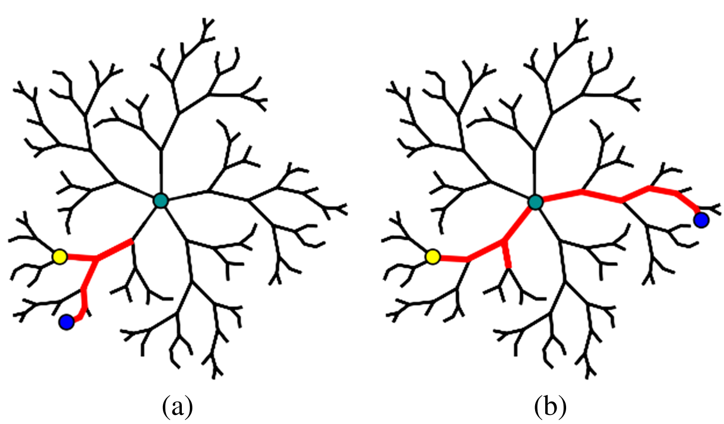

FIG. 4. *The concentration behavior of the $q_j(x)$ is very different from the $p_j^{\alpha,\beta}(x)$ since it is not in general true that $q_j(x)\le C\rho(r)$ while it is true that $p_j^{\alpha,\beta}(x)\le C\rho(r)$. For example, in a graph which is locally tree like as depicted above, it can be that $q_j(x)=1$ for some combinations of entrance and exit points.* (a) *Entrance and exit points of an excursion from $B(x,4)$ to $B(x,6)$, respectively, conditional on which random walk has a low probability of hitting x.* (b) *Entrance and exit points of an excursion from $B(x,4)$ to $B(x,6)$, respectively, conditional on which random walk is forced to hit x.*

a set of points in V_n such that if x, y are distinct in A then $d(x, y) \geq 4R_n^\gamma$. Fix $R > r > 0$ and let $\tau_{k+1}(A) = \min\{t \geq \sigma_k(x) : X(t) \in \partial A(r)\}$. Fix $\beta > 0$ and define indices $i(j, x)$ inductively as follows. Set

$$i(1, x) = \min\{k \geq 1 : \tau_{k+1}(A) - \sigma_k(x) \geq T_\beta^U\}$$

and, for each $j \geq 1$, let

$$i(j+1, x) = \min\{k \geq i(j, x) + 1 : \tau_{k+1}(A) - \sigma_k(x) \geq T_\beta^U\}.$$

When x is clear from the context we will write $i(j)$ for $i(j, x)$.

LEMMA 6.2. *For each $\delta > 0$ and $r > 0$ there exists $R_0 > r$ such that for $R > R_0$ fixed there exists i.i.d. random variables $(I(j, x) : x \in A, j \geq 1)$ which stochastically dominate from above $(i(j, x) : x \in A, j \geq 1)$ and satisfy*

$$\mathbf{P}\big[I\big((1-\delta)j, x\big) \geq j\big] \leq C \exp(-C\delta^2 j)$$

for all n large enough. Let $\mathcal{G}(j, x) = \sigma(\{q_{i(k)}(x) : k \neq j\} \cup \{q_{i(k)}(y) : y \in A \setminus \{x\}\})$. There exists i.i.d. random variables $(Q_j(x) : j \geq 1)$ taking values in $[0, 2\rho(r)]$ such that

$$1 - O(e^{-c\beta}) \leq \frac{\mathbf{E}[q_{i(j)}(x) | \mathcal{G}(j, x)]}{Q_j(x)} \leq 1 + O(e^{-c\beta})$$

and

$$1 - O(e^{-c\beta}) \leq \frac{\overline{p}_{r,R}(x)}{\mathbf{E}Q_j(x)} \leq 1 + O(e^{-c\beta})$$

for all n large enough. Furthermore, the families $\{(Q_j(x) : j \geq 1) : x \in A\}$ are independent.

PROOF. Define stopping times

$$\sigma_{k0}(A) = \min\{t \geq \sigma_k(x) : d(X(t), A) \geq 2R_n^\gamma\},$$
$$\tau_{k1}(A) = \min\{t \geq \sigma_{k0}(x) : d(X(t), A) \leq R_n^\gamma\}.$$

For $j \geq 1$, inductively set

$$\sigma_{kj}(A) = \min\{t \geq \tau_{kj}(A) : d(X(t), A) \geq 2R_n^\gamma\},$$
$$\tau_{k(j+1)}(A) = \min\{t \geq \sigma_{kj}(A) : d(X(t), A) \leq R_n^\gamma\}.$$

Note that $\sigma_{kj}(A) - \tau_{kj}(A) \geq R_n^\gamma$. Thus, for $j_\beta = T_\beta^U / R_n^\gamma$ we have that $\tau_{kj_\beta}(A) \geq \sigma_k(x) + T_\beta^U$. Let $F_k(x) = \{X(t) \in A(r) \text{ for } t \in [\sigma_k(x), \sigma_k(x) + T_\beta^U]\}$. Let x_{kj} be the element in A such that $d(X(\tau_{kj}(A)), x_{kj}) \leq R_n^\gamma$. Observe

$$\mathbf{P}_{X(\tau_{kj}(A))}\big[X(t) \in A(r) \text{ for } t \in [\tau_{kj}(A), \sigma_{kj}(A)] | x_{kj}\big]$$
$$\leq C \max_{d(y, x_{kj}) = R_n^\gamma} g(y, B(x_{kj}, r); G_n).$$

Uniform local transience also yields

$$\mathbf{P}_{X(\sigma_k(x))}\big[X(t) \in A(r) \text{ for } t \in [\sigma_k(x), \tau_{k0}(A)]\big] \le C\rho(R, r) \le \delta/2,$$

provided $R > r$ is large enough. A union bound thus gives

$$\begin{aligned} \mathbf{P}_{X(\tau_k(x))}[F_k(x)] &\le \max_z \max_{d(y,z)=R_n^\gamma} g(y, B(z,r); G_n) \frac{T_\beta^U}{R_n^\gamma} + \delta/2 \\ &\le \delta/2 + o(1) \le \delta \end{aligned} \tag{6.2}$$

as $n \to \infty$ by part (1) of Assumption 1.2. Note that if $x_1, \ldots, x_\ell \in A$ and $j(1), \ldots, j(k)$ are such that $\tau_{j(k)}(x_k) \le \tau_{j(k+1)}(x_{k+1})$ then we have

$$\begin{aligned} &\mathbf{P}\big[F_{j(1)}(x_1), \ldots, F_{j(\ell)}(x_\ell)\big] \\ &\quad = \mathbf{E}\big[\mathbf{P}_{X(\tau_{j(\ell)})(x_\ell)}\big[F_{j(\ell)}(x_\ell)\big]\mathbf{1}_{F_{j(1)}(x_1)} \cdots \mathbf{1}_{F_{j(\ell-1)}(x_{\ell-1})}\big] \\ &\quad \le \delta \mathbf{P}\big[F_{j(1)}(x_1), \ldots, F_{j(\ell-1)}(x_{\ell-1})\big] \le \cdots \le \delta^\ell. \end{aligned}$$

This can of course be repeated with any subset of the above events which implies the stochastic domination claim. It easily now follows from Cramér's theorem that

$$\mathbf{P}\big[I\big((1-\delta)k, x\big) \ge k\big] \le 2\exp(-C\delta^2 k).$$

For the second part of the lemma, we just need to get a bound on $\mu_x(z)/\pi(z)$ where μ_x is the law of random walk started at x conditioned not to get within distance r of A by, say, time $T_{\beta/2}^U$. This can be done in exactly the same way as in the proof of Lemma 4.3. Indeed, the term $|A|\rho(s,r)$ in the statement of that lemma comes from a bound on the probability that random walk at distance s from A hits A in time T_α^U. In the situation of this lemma, the role of s is replaced by R_n^γ and we can use the scheme developed above to estimate the error contributed by this term by $O(\delta)$ provided n is sufficiently large. □

We now turn to the case that (G_n) satisfies part (2) of Assumption 1.2. This case will turn out to be substantially easier, the reason being that the Harnack inequality implies the quenched bound $q_j(x) \le 2C\rho(r)$. We emphasize once more that this is not the case in locally tree-like graphs.

LEMMA 6.3. *If (G_n) satisfies part (2) of Assumption 1.2, then for each $r, \delta > 0$ there exists $R_0 > r$ such that $R \ge R_0$ implies*

$$\begin{aligned} &\mathbf{P}\left[\prod_{j=1}^k \big(1 - q_j(x)\big) \ge \big(1 - (1+\delta)\overline{p}_{r,R}(x)\big)^{k(1+\delta)}\right] \\ &\qquad \le C\big[\exp\big(-C\delta^2 \overline{p}_{r,R}(x)k/\rho(r)\big) + \exp(-C\delta^2 k)\big] \end{aligned} \tag{6.3}$$

for all n large enough.

PROOF. The uniform Harnack inequality implies that $q_j(x) \le 2C\rho(r)$ where $C = C(R/r)$ is the constant from the statement of part (2) of Assumption 1.2. Let $F_j = \{\tau_j(x) - \sigma_{j-1}(x) \le T_\beta^U\}$. Arguing as in the previous lemma and invoking uniform local transience, there exists i.i.d. random variables $\widetilde{F}_j(x)$ with $\mathbf{P}[\widetilde{F}_j(x) = 1] = \delta = 1 - \mathbf{P}[\widetilde{F}_j(x) = 0]$ that stochastically dominate $(\mathbf{1}_{F_j(x)} : j)$ provided R is sufficiently large. We let $\iota(j)$ be the jth smallest index i such that $F_i(x)$ occurs. The lemma now follows from an argument similar to that of Lemma 4.2. Indeed, we can stochastically dominate $q_{\iota(j)}(x)$ from below by i.i.d. random variables L_j with $\mathbf{E}L_j \ge (1-\delta)\overline{p}_{r,R}(x)$ and $L_j \le 10C\rho(r)$. By Cramér's theorem,

$$\mathbf{P}\left[\prod_{j=1}^{k}(1-L_j) \ge \bigl(1-(1+\delta)\overline{p}_{r,R}(x)\bigr)^k\right] \le C\exp\bigl(-C\delta^2\overline{p}_{r,R}(x)k/\rho(r)\bigr).$$

The lemma now follows since, again by Cramér's theorem,

$$\mathbf{P}\bigl[\iota\bigl((1-\delta)k\bigr) \ge k\bigr] \le C\exp(-C\delta^2 k). \qquad \square$$

6.3. *Proof of Theorem* 1.3. We begin by showing that the points not visited by X by time $\frac{1}{2}T_{\mathrm{cov}}(G_n)$ are typically well separated, which in turn will be helpful when we estimate the exponential moment in Proposition 3.2. To this end, we let $\delta > 0$ be arbitrary and assume that $R > r, n_0, \varepsilon$ have been chosen so that for all $n \ge n_0$ we have

$$1-\delta \le \frac{T_{\mathrm{cov}}(G_n)}{C_n^\varepsilon} \le 1+\delta.$$

We may assume without loss of generality that $d_k^\varepsilon > 0$ for all relevant k and, in particular, that $|H_{n,k}^\varepsilon|^{-\delta} \to 0$ for every k. Indeed, Lemmas 4.6 and 4.7 imply that $T_{\mathrm{hit}}(G_n) = \Theta(|V_n|)$, consequently if $\log|H_{n,k}^\varepsilon| \to 0$ as $n \to \infty$ then $T_{\mathrm{cov}}(H_{n,k}^\varepsilon)$ is negligible in comparison to $T_{\mathrm{cov}}(G_n)$. If (G_n) satisfies Assumption 1.2(1) we take R_n^γ as given there. Otherwise, we take $R_n^\gamma = \max\{R > 0 : \max_{x\in V_n}|B(x,R)| \le |V_n|^\gamma\}$.

LEMMA 6.4. *Let $\mathcal{R}(t)$ denote the range of random walk at time t and $\mathcal{L}(t) = V \setminus \mathcal{R}(t)$. Letting*

$$M = \begin{cases} 20\Delta_0 \sup\limits_n \overline{\Delta}^R(G_n)/(\delta\varepsilon d^\varepsilon), & \text{if } \sup_n \overline{\Delta}(G_n) < \infty, \\ 20\Delta_0/(\delta\varepsilon d^\varepsilon), & \text{otherwise}, \end{cases}$$

and

$$\mathcal{T}_0 = \min\Bigl\{T \ge 0 : \max_x |\mathcal{L}(t) \cap B(x, R_n^\gamma)| \le M\Bigr\},$$

we have that $\mathbf{P}[\mathcal{T}_0 > \frac{1+5\delta}{2}T_{\mathrm{cov}}(G_n)] = o(1)$ provided γ is sufficiently small, R is so large that $\delta_{R,m} \le 1$, $d^\varepsilon = \min\{d_k^\varepsilon : d_k^\varepsilon > 0\}$ and $m = 20/d^\varepsilon$. Furthermore, letting

$$\mathcal{T}_1 = \min\{T \ge 0 : |\mathcal{L}(t) \cap H_{n,k}^\varepsilon| \le |H_{n,k}^\varepsilon|^{1/2-\delta} \text{ for all } k\}$$

we have that $\mathbf{P}[\mathcal{T}_1 > \frac{1+5\delta}{2} T_{\mathrm{cov}}(G_n)] = o(1)$.

PROOF. First, suppose that (G_n) has uniformly bounded maximal degree. Fix $R > r$ and let A be an R-net of $H_{n,k}^{\varepsilon}$. Fix $x \in H_{n,k}^{\varepsilon}$ and suppose that $x_1, \dots, x_\ell \in B(x, R_n^{\gamma}) \cap H_{n,k}^{\varepsilon} \cap A$ are distinct. Lemma 5.1 gives us

$$\mathbf{P}\big[x_1, \dots, x_\ell \in \mathcal{L}\big((1+\delta)/2; G_n\big)\big] \le (1+\delta_{R,\ell})|V_n|^{-(1+\delta)\ell d_k^{\varepsilon}/2 + \delta_{R,\ell}}.$$

Consequently, a union bound yields

$$\begin{aligned}\mathbf{P}\big[\big|\mathcal{L}\big((1+\delta)/2; G_n\big) \cap B(x, R_n^{\gamma}) \cap A\big| \ge \ell\big] &\\ \le (1+\delta_{R,\ell})|B(x, R_n^{\gamma})|^{\ell} |V_n|^{-(1+\delta)\ell d_k^{\varepsilon}/2 + \delta_{R,\ell}} &\\ \le (1+\delta_{R,\ell})|V_n|^{(\gamma - (1+\delta)d_k^{\varepsilon}/2)\ell + \delta_{R,\ell}}. &\end{aligned}$$

Hence, choosing $\gamma \le d^{\varepsilon}/4$ the above is $O(|V_n|^{-3})$. Since the number of disjoint R-nets necessary to cover $H_{n,k}^{\varepsilon}$ is at most $\overline{\Delta}^R(G_n)$, the result now follows from a union bound. In the case of unbounded maximal degree, we can skip the step of subdividing the $H_{n,k}^{\varepsilon}$ into R-nets since in this case $\delta_{1,m} \to 0$, otherwise the proof is the same. The second claim is immediate from Markov's inequality and Lemma 5.1. $\square$

We can now complete the proof of Theorem 1.3. We will handle the two cases depending on whether (G_n) satisfies part (1) or (2) of Assumption 1.2. Throughout, we let $N(x, T)$ be the number of such excursions from $\partial B(x, r)$ to $\partial B(x, R)$ that have occurred by time T.

PROOF OF THEOREM 1.3, UNDER ASSUMPTION 1.2(2). Let

$$\mathcal{T}_2 = \min\left\{T \ge 0 \colon \max_{x \in H_{n,k}^{\varepsilon}} \prod_{k=1}^{N(x,T)} \big(1 - q_j(x)\big) \le |H_{n,k}^{\varepsilon}|^{-1/2-\delta} \text{ for all } k\right\}$$

and set

$$\mathcal{T} = \mathcal{T}_0 \vee \mathcal{T}_1 \vee \mathcal{T}_2 \vee \left(\frac{1+5\delta}{2}\right) T_{\mathrm{cov}}(G_n). \tag{6.4}$$

Let $k_0(n)$ be a sequence so that $\liminf_{n \to \infty} d_{k_0(n)}^{\varepsilon} \ge \delta_0 > 0$. For $x \in H_{n,k_0(n)}^{\varepsilon}$, we have

$$\left(\frac{1+3\delta}{2}\right) C_{n,k_0(n)}^{\varepsilon} \ge \left(\frac{1+3\delta+O(\varepsilon)}{2}\right) \frac{\delta_0 T_{r,R}(x) \log|V_n|}{4\rho(r)}$$

for all n large enough. Thus letting $M_{n,k_0(n)}^{\varepsilon}(x) = (1+3\delta)/2 \cdot C_n^{\varepsilon}(x)/T_{r,R}(x)$, we have

$$M_{n,k_0(n)}^{\varepsilon}(x) \ge \left(\frac{1+3\delta+O(\varepsilon)}{2}\right) \frac{\delta_0 \log|V_n|}{4\rho(r)}.$$

Now,

$$\mathbf{P}\big[(1-\delta)T_{r,R}(x)M^{\varepsilon}_{n,k_0(n)}(x) \le \tau_{M^{\varepsilon}_{n,k_0(n)}}(x) \le (1+\delta)T_{r,R}(x)M^{\varepsilon}_{n,k_0(n)}(x)\big]$$
$$\ge 1 - C\exp\biggl(-\frac{C\delta_0\delta^2}{\rho(r)}\log|V_n|\biggr) \ge 1 - O(|V_n|^{-100}),$$

provided we choose r large enough. Choosing $R > r$ sufficiently large, Lemma 6.3 gives us

$$\mathbf{P}\Biggl[\prod_{j=1}^{M^{\varepsilon}_{n,k_0(n)}(x)} (1-q_j(x)) \ge |H^{\varepsilon}_{n,k_0(n)}|^{-1/2-\delta}\Biggr] \le O(|V|^{-100}).$$

Combining everything,

$$\mathbf{P}\biggl[\mathcal{T} \ne \biggl(\frac{1+5\delta}{2}\biggr)T_{\mathrm{cov}}(G_n)\biggr] = o(1) \qquad \text{as } n \to \infty. \tag{6.5}$$

Let μ be the probability on $\mathcal{X}(G_n)$ given by first sampling $\mathcal{R} \subseteq V_n$ according to μ_0, the measure on subsets of V_n given by running X to time $(1+5\delta)/2 \cdot T_{\mathrm{cov}}(G_n)$, then sampling $f|_{\mathcal{R}}$ by marking with i.i.d. fair coins and $f|_{V_n \setminus \mathcal{R}} \equiv 0$. Define $\widetilde{\mu}$ similarly except by sampling $\mathcal{R} \subseteq V_n$ according to $\widetilde{\mu}_0$, the measure given by running X up to time $\mathcal{T}$ rather than $(1+5\delta)/2 \cdot T_{\mathrm{cov}}(G_n)$. As a consequence of (6.5),

$$\|\mu - \widetilde{\mu}\|_{\mathrm{TV}} \le \mathbf{P}\biggl[\mathcal{T} \ne \biggl(\frac{1+5\delta}{2}\biggr)T_{\mathrm{cov}}(G_n)\biggr] = o(1) \qquad \text{as } n \to \infty.$$

Suppose we have two independent random walks X, X' on G_n, each with stationary initial distribution, and let $\mathcal{T}, \mathcal{T}'$ be stopping times for each as in (6.4). Let $\mathcal{R}, \mathcal{R}'$ be their ranges at time $\mathcal{T}, \mathcal{T}'$, respectively, and $\mathcal{L} = V_n \setminus \mathcal{R}$, $\mathcal{L}' = V_n \setminus \mathcal{R}'$. Let $q'_j(x)$ be the quantity analogous to $q_j(x)$ for X' and $\mathcal{G} = \sigma(q'_j(x) : j \ge 1)$. The previous lemma implies that we can divide $\mathcal{L}$ into M disjoint sets $A_1, \dots, A_M$ such that if $x, y \in A_\ell$ with $x \ne y$ then $d(x,y) \ge R_n^{\gamma} > R$. Consequently, letting $\mathcal{G}(A_\ell) = \otimes_{x \in A_\ell}\mathcal{G}(x)$ we have

$$\mathbf{E}[\exp(\zeta|\mathcal{L} \cap \mathcal{L}' \cap A_\ell|)|\mathcal{G}(A_\ell)] \le \prod_{x \in A_\ell}\Biggl(1 + e^{\zeta}\Biggl(\prod_{j=1}^{N(x,\mathcal{T}')}(1-q'_j(x))\Biggr)\Biggr)$$
$$\le \exp\biggl(e^{\zeta}\sum_k |H^{\varepsilon}_{n,k}|^{-\delta}\biggr).$$

Since $A_1, \dots, A_M$ cover $\mathcal{L}$, it follows from Hölder's inequality that

$$\begin{aligned} \mathbf{E}\exp(\zeta|\mathcal{L} \cap \mathcal{L}'|) &\le \biggl[\exp\biggl(e^{\zeta M}\sum_k |H^{\varepsilon}_{n,k}|^{-\delta}\biggr)\biggr]^{1/M} \\ &\le 1 + 2\frac{\exp(\zeta M)}{M}\sum_k |H^{\varepsilon}_{n,k}|^{-\delta}. \end{aligned} \tag{6.6}$$

$\square$

PROOF OF THEOREM 1.3, UNDER ASSUMPTION 1.2(1). Let

$$\mathcal{T}_2 = \min\left\{T \geq 0 : \max_k \max_{x \in H_{n,k}^\varepsilon} \frac{(1+2\delta)\log|H_{n,k}^\varepsilon|}{2N(x,T)\overline{p}_{r,R}(x)} \leq 1\right\}$$

and

$$\mathcal{T} = \mathcal{T}_0 \vee \mathcal{T}_1 \vee \mathcal{T}_2 \vee \left(\frac{1+5\delta}{2}\right) T_{\mathrm{cov}}(G_n). \tag{6.7}$$

It follows from Lemmas 4.5 and 4.8 and the definition of $H_{n,k}^\varepsilon$ that

$$\mathbf{P}\left[\mathcal{T} \neq \left(\frac{1+5\delta}{2}\right) T_{\mathrm{cov}}(G_n)\right] = o(1) \qquad \text{as } n \to \infty. \tag{6.8}$$

Let μ be the probability on $\mathcal{X}(G_n)$ given by first sampling $\mathcal{R} \subseteq V_n$ according to μ_0, the measure on subsets of V_n given by running X to time $(1+5\delta)/2 \cdot T_{\mathrm{cov}}(G_n)$, then sampling $f|_{\mathcal{R}}$ by marking with i.i.d. fair coins and $f|_{V_n \setminus \mathcal{R}} \equiv 0$. Define $\widetilde{\mu}$ similarly except by sampling $\mathcal{R} \subseteq V_n$ according to $\widetilde{\mu}_0$, the measure given by running X up to time $\mathcal{T}$ rather than $(1+5\delta)/2 \cdot T_{\mathrm{cov}}(G_n)$. As a consequence of (6.8),

$$\|\mu - \widetilde{\mu}\|_{\mathrm{TV}} \leq \mathbf{P}\left[\mathcal{T} \neq \left(\frac{1+5\delta}{2}\right) T_{\mathrm{cov}}(G_n)\right] = o(1) \qquad \text{as } n \to \infty.$$

Suppose we have two independent random walks X, X' on G_n, each with stationary initial distribution, and let $\mathcal{T}, \mathcal{T}'$ be stopping times for each as in (6.7). Using the same notation as the previous proof, by the definition of $\mathcal{T}_2'$, we have

$$\begin{aligned}
&\mathbf{E}\big[\mathbf{E}[\exp(\zeta|\mathcal{L} \cap \mathcal{L}' \cap A_\ell|)|\mathcal{G}(A_\ell)]\big] \\
&\quad \leq \mathbf{E}\prod_{x \in A_\ell}\left(1 + e^\zeta\left(\prod_{j=1}^{N(x,\mathcal{T}')}(1 - q_j'(x))\right)\right) \\
&\quad \leq \mathbf{E}\prod_{x \in A_\ell}\left(1 + e^\zeta\left(\prod_{j=1}^{N(x)}(1 - q_j'(x))\right)\right),
\end{aligned} \tag{6.9}$$

where $N(x) = (1+2\delta)\log|H_{n,k}^\varepsilon|/2\overline{p}_{r,R}(x)$ and k is such that $x \in H_{n,k}^\varepsilon$. Let

$$\widetilde{N}(x) = (1-\delta)N(x) \geq \frac{(1+\delta/2)\log|H_{n,k}^\varepsilon|}{2\overline{p}_{r,R}(x)}.$$

Observe that (6.9) is bounded by

$$\mathbf{E}\prod_{x \in A_\ell}\left(1 + e^\zeta\left(\prod_{j=1}^{\widetilde{N}(x)}(1 - q_{i(j)}'(x)) + \mathbf{1}_{\{I(\widetilde{N}(x)) > N(x)\}}\right)\right).$$

As A_ℓ satisfies the hypotheses of Lemma 6.2, this is in turn bounded by

$$\mathbf{E}\prod_{x\in A_\ell}\left(1+e^{\zeta}\left(\prod_{j=1}^{\widetilde{N}(x)}(1-(1-\delta/4)Q_j'(x))\right)+O(|V_n|^{-100})\right)$$
$$\leq \exp\left(e^{\zeta}\sum_k |H_{n,k}^{\varepsilon}|^{-\delta}\right).$$

The theorem now follows from Hölder's inequality, as in the previous proof. □

6.4. *The lamplighter.*

PROOF OF THEOREM 1.5. This is proved by making several small modifications to the proof of Theorem 1.3. Namely, rather than considering the range of X run up to time $\mathcal{T}$ as in either (6.4) or (6.7), one considers the range $\widetilde{\mathcal{R}}(x)$ of X run up to time $\mathcal{T}$, conditioned on the event $\{X(\mathcal{T})=x\}$ for a given point x. Exactly the same argument shows that the total variation distance of the law $\widetilde{\mu}_x$ on markings $\mathcal{X}(G_n)$ induced by i.i.d. coin flips on $\widetilde{\mathcal{R}}(x)$ and 0 on $(\widetilde{\mathcal{R}}(x))^c$ from the uniform measure on $\mathcal{X}(G_n)$ is $o(1)$. This implies that the law μ_x on markings of $\mathcal{X}(G_n)$ given by i.i.d. coin flips on the range $\mathcal{R}(x)$ of X run up to time $T=\frac{1+\varepsilon}{2}T_{\mathrm{cov}}(G_n)$, conditioned on $\{X(T)=x\}$, and the uniform measure is $o(1)$. At time T, the random walk is well mixed, from which the result is clear. □

7. Further questions.

1. Theorem 1.3 yields a wide class of examples where the threshold for indistinguishability is at $\frac{1}{2}T_{\mathrm{cov}}$, and $\mathbf{Z}_n^2$ is an example where the threshold is at T_{cov}. Does there exist a sequence (G_n) of vertex transitive graphs where the threshold is at $\alpha T_{\mathrm{cov}}(G_n)$ for $\alpha\in(1/2,1)$?

2. Our statistical test for uniformity is only valid for $\alpha>1/2$. For $\alpha\leq 1/2$, the natural reference measure is i.i.d. markings conditioned on the number of zeros being on the order of $|V|^{1-\alpha}$. Can analogous results be proved in this setting?

3. Our definition of uniform local transience is given in terms of Green's function summed up to the *uniform mixing time*. Does it suffice to assume only the uniform decay of

$$g(x,y;G)=\sum_{t=1}^{T}p^t(x,y;G),$$

where $T=T_{\mathrm{mix}}(G)$ or even $T=T_{\mathrm{rel}}(G)$?

4. The complete graph K_n does not satisfy the hypotheses of Theorem 1.3 yet the lamplighter walk on K_n has a threshold at $\frac{1}{2}T_{\mathrm{cov}}(K_n)$. Is there a more general theorem allowing for a unified treatment of this case?

Acknowledgment. J. Miller thanks the Theory Group at Microsoft Research for support through a summer internship, during which the research for this article was conducted.

REFERENCES

[1] ALDOUS, D. (1983). Random walks on finite groups and rapidly mixing Markov chains. In *Seminar on Probability, XVII. Lecture Notes in Math.* **986** 243–297. Springer, Berlin. MR0770418

[2] ALDOUS, D. and DIACONIS, P. (1986). Shuffling cards and stopping times. *Amer. Math. Monthly* **93** 333–348. MR0841111

[3] ALDOUS, D. and FILL, J. Reversible Markov chains and random walks on graphs. Available at http://www.stat.berkeley.edu/~aldous/RWG/book.html.

[4] ALDOUS, D. J. (1991). Threshold limits for cover times. *J. Theoret. Probab.* **4** 197–211. MR1088401

[5] BARLOW, M. T. (2004). Random walks on supercritical percolation clusters. *Ann. Probab.* **32** 3024–3084. MR2094438

[6] BENJAMINI, I. and MOSSEL, E. (2003). On the mixing time of a simple random walk on the super critical percolation cluster. *Probab. Theory Related Fields* **125** 408–420. MR1967022

[7] BRODER, A. and SHAMIR, E. (1987). On the second eigenvalue of random regular graphs. In *28th Annual Symposium on Foundations of Computer Science* 286–294. ACM, New York.

[8] BRUMMELHUIS, M. J. A. M. and HILHORST, H. J. (1991). Covering of a finite lattice by a random walk. *Phys. A* **176** 387–408. MR1130067

[9] DELMOTTE, T. (1999). Parabolic Harnack inequality and estimates of Markov chains on graphs. *Rev. Mat. Iberoam.* **15** 181–232. MR1681641

[10] DEMBO, A., PERES, Y. and ROSEN, J. (2003). Brownian motion on compact manifolds: Cover time and late points. *Electron. J. Probab.* **8** 14 pp. (electronic). MR1998762

[11] DEMBO, A., PERES, Y., ROSEN, J. and ZEITOUNI, O. (2001). Thick points for planar Brownian motion and the Erdős–Taylor conjecture on random walk. *Acta Math.* **186** 239–270. MR1846031

[12] DEMBO, A., PERES, Y., ROSEN, J. and ZEITOUNI, O. (2004). Cover times for Brownian motion and random walks in two dimensions. *Ann. of Math.* (2) **160** 433–464. MR2123929

[13] DEMBO, A., PERES, Y., ROSEN, J. and ZEITOUNI, O. (2006). Late points for random walks in two dimensions. *Ann. Probab.* **34** 219–263. MR2206347

[14] DEMBO, A. and ZEITOUNI, O. (1998). *Large Deviations Techniques and Applications*, 2nd ed. *Applications of Mathematics* (*New York*) **38**. Springer, New York. MR1619036

[15] DEUSCHEL, J.-D. and PISZTORA, A. (1996). Surface order large deviations for high-density percolation. *Probab. Theory Related Fields* **104** 467–482. MR1384041

[16] DIACONIS, P. and SHAHSHAHANI, M. (1981). Generating a random permutation with random transpositions. *Z. Wahrsch. Verw. Gebiete* **57** 159–179. MR0626813

[17] FITZSIMMONS, P. J. and PITMAN, J. (1999). Kac's moment formula and the Feynman–Kac formula for additive functionals of a Markov process. *Stochastic Process. Appl.* **79** 117–134. MR1670526

[18] HÄGGSTRÖM, O. and JONASSON, J. (1997). Rates of convergence for lamplighter processes. *Stochastic Process. Appl.* **67** 227–249. MR1449833

[19] LAWLER, G. F. (1991). *Intersections of Random Walks*. Birkhäuser, Boston, MA. MR1117680

[20] LEVIN, D. A., PERES, Y. and WILMER, E. L. (2009). *Markov Chains and Mixing Times*. Amer. Math. Soc., Providence, RI. MR2466937

[21] MATTHEWS, P. (1988). Covering problems for Markov chains. *Ann. Probab.* **16** 1215–1228. MR0942764

[22] PERES, Y. and REVELLE, D. (2004). Mixing times for random walks on finite lamplighter groups. *Electron. J. Probab.* **9** 825–845. MR2110019

[23] PISZTORA, A. (1996). Surface order large deviations for Ising, Potts and percolation models. *Probab. Theory Related Fields* **104** 427–466. MR1384040

DEPARTMENT OF MATHEMATICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA 94305
USA
E-MAIL: jmiller@math.stanford.edu

MICROSOFT RESEARCH
THEORY GROUP
REDMOND, WASHINGTON 98052
USA
E-MAIL: peres@microsoft.com