The Annals of Probability
2013, Vol. 41, No. 2, 636–670
DOI: 10.1214/11-AOP713

PAINTING A GRAPH WITH COMPETING RANDOM WALKS[1]

BY JASON MILLER

Stanford University

Let X_1, X_2 be independent random walks on $\mathbf{Z}_n^d$, $d \geq 3$, each starting from the uniform distribution. Initially, each site of $\mathbf{Z}_n^d$ is unmarked, and, whenever X_i visits such a site, it is set irreversibly to i. The mean of $|\mathcal{A}_i|$, the cardinality of the set $\mathcal{A}_i$ of sites painted by i, once all of $\mathbf{Z}_n^d$ has been visited, is $\frac{1}{2}n^d$ by symmetry. We prove the following conjecture due to Pemantle and Peres: for each $d \geq 3$ there exists a constant α_d such that $\lim_{n\to\infty} \mathrm{Var}(|\mathcal{A}_i|)/h_d(n) = \frac{1}{4}\alpha_d$ where $h_3(n) = n^4$, $h_4(n) = n^4(\log n)$ and $h_d(n) = n^d$ for $d \geq 5$. We will also identify α_d explicitly and show that $\alpha_d \to 1$ as $d \to \infty$. This is a special case of a more general theorem which gives the asymptotics of $\mathrm{Var}(|\mathcal{A}_i|)$ for a large class of transient, vertex transitive graphs; other examples include the hypercube and the Caley graph of the symmetric group generated by transpositions.

1. Introduction. Suppose that X_1, X_2 are independent random walks on a graph $G = (V, E)$ starting from stationarity. Initially, each vertex of G is unmarked, and, whenever X_i visits such a site, it is marked i irreversibly. If both X_1 and X_2 visit a site for the first time simultaneously, then the mark is chosen by the flip of an independent fair coin. Let $\mathcal{A}_i$ be the set of sites marked i once every vertex of G has been visited. By symmetry, it is obvious that $\mathbf{E}|\mathcal{A}_i| = \frac{1}{2}|V|$. The purpose of this manuscript is to derive precise asymptotics for $\mathrm{Var}(|\mathcal{A}_i|)$ for many families of graphs.

The process by which a single random walk covers a graph has been studied extensively. Examples of interesting statistics include the expected amount of time it takes for the random walk to visit every site [4, 14], the growth exponent of the set of sites visited most frequently [3] and the clustering and correlation structure of the last visited points [2, 5, 15]. The motivation for this work is to understand better how multiple random walks cover a graph.

The investigation of the statistical properties of $\mathcal{A}_i$ was first proposed in the work of Gomes Jr. et al. [9]. Their motivation was to study the technical challenges associated with physical problems involving interacting random walks. They estimate the growth of $\mathbf{E}|\mathcal{B}|$ where $\mathcal{B}$ is the interface separating $\mathcal{A}_1$ from $\mathcal{A}_2$ in the special case of the one-cycle $\mathbf{Z}_n^1$. As with $\mathbf{E}|\mathcal{A}_i|$, computing $\mathbf{E}|\mathcal{B}|$ for $\mathbf{Z}_n^d$ becomes

Received March 2010; revised April 2011.
[1]Supported in part by NSF Grants DMS-04-06042 and DMS-08-06211.
MSC2010 subject classifications. 60G50, 60F99.
Key words and phrases. Random walk, competing random walks, variance.

trivial for $d \geq 3$ since it is easy to see that, with probability strictly between 0 and 1, for any pair of adjacent vertices x, y, X_2 will hit y before X_1, conditional on the event that X_1 hits x first. On the other hand, estimating $\mathrm{Var}(|\mathcal{A}_i|)$ in this setting is challenging since its expansion in terms of correlation functions exhibits significant cancellation which, when ignored, leads to bounds that are quite imprecise. We will develop this point further at the end of the Introduction.

The problem we consider here was formulated by Hilhorst, though in a slightly different setting. Rather than considering the sets of sites $\mathcal{A}_1, \mathcal{A}_2$ *first painted* by X_1, X_2, respectively, it is also natural to study the sets $\widetilde{\mathcal{A}}_1, \widetilde{\mathcal{A}}_2$ of sites *most recently painted* by X_1, X_2, respectively. In other words, in the latter formulation the constraint that the marks are irreversible is removed. It turns out that these two classes of problems are equivalent, which is to say $(\mathcal{A}_1, \mathcal{A}_2) \stackrel{d}{=} (\widetilde{\mathcal{A}}_1, \widetilde{\mathcal{A}}_2)$. This helpful observation, which follows from the time-reversibility of random walk, was made and communicated to us by Comets.

We restrict our attention to *lazy* walks X_1, X_2 to avoid issues of periodicity, and in particular to ensure that the random walk has a unique stationary distribution. That is, the one-step transition kernel is given by

$$p(x, y; G) = \begin{cases} \dfrac{1}{2}, & \text{if } x = y, \\ \dfrac{1}{2\deg(x)}, & \text{if } x \sim y, \\ 0, & \text{otherwise,} \end{cases}$$

where $x \sim y$ means that x is adjacent to y in G. The particular choice of holding probability $\frac{1}{2}$ is not important for the proof; indeed, any $\lambda \in (0, 1)$ would suffice. Our proofs also work in the setting of continuous time walks. Let $p^t(\cdot, \cdot; G)$ be the t-step transition kernel of a lazy random walk on G and $\pi(\cdot; G)$ its unique stationary distribution.

Our main result is the precise asymptotics for $\mathrm{Var}(|\mathcal{A}_i|)$ on tori of dimension at least three, thus verifying a conjecture due to Pemantle and Peres [8], page 35.

THEOREM 1.1. *Suppose that $G_n = \mathbf{Z}_n^d$, $d \geq 3$. There exists a finite constant $\alpha_d > 0$ such that*

$$\lim_{n \to \infty} \frac{\mathrm{Var}(|\mathcal{A}_i|)}{h_d(n)} = \frac{1}{4}\alpha_d,$$

where

$$h_d(n) = \begin{cases} n^4, & \text{if } d = 3, \\ n^4(\log n), & \text{if } d = 4, \text{ and} \\ n^d, & \text{if } d \geq 5. \end{cases}$$

Our proof allows us to identify α_d explicitly and is given as follows. Let

$$G(x;\mathbf{Z}^d)=\mathbf{E}_0\sum_{t=0}^{\infty}\mathbf{1}_{\{X(t)=x\}} \tag{1.1}$$

be the Green's function for lazy random walk on $\mathbf{Z}^d$. This is the amount of time a random walk initialized at 0 spends at x before escaping to ∞. For $d \geq 5$,

$$\alpha_d=\frac{1}{G^2(0;\mathbf{Z}^d)}\sum_{y\in\mathbf{Z}^d}G^2(y;\mathbf{Z}^d). \tag{1.2}$$

It is not difficult to see that $\alpha_d \to 1$ as $d \to \infty$, so that $\mathrm{Var}(|\mathcal{A}_1|) \approx \frac{1}{4}n^d$ for d and n large is close to the variance of an i.i.d. marking. For $d = 4$,

$$\alpha_4=\lim_{n\to\infty}\frac{1}{G^2(0;\mathbf{Z}^4)\log n}\sum_{y\in\mathbf{Z}^4:\,|y|\leq n}G^2(y;\mathbf{Z}^4); \tag{1.3}$$

we will explain why this limit exists and is positive and finite in Proposition 2.1. The definition of α_3 is slightly more involved. Let $\mathbf{T}^3$ denote the three-dimensional continuum torus, $p^t(\cdot,\cdot;\mathbf{T}^3)$ the transition kernel for Brownian motion on $\mathbf{T}^3$ and

$$g^T(x,y;\mathbf{T}^3)=\int_0^T p^t(x,y;\mathbf{T}^3)\,dt.$$

Now set

$$\begin{aligned}\alpha_3^T&=\frac{1}{G^2(0;\mathbf{Z}^3)}\int_{\mathbf{T}^3}\int_{\mathbf{T}^3}\left(g^{T/2}(x,y;\mathbf{T}^3)-\frac{1}{2}T\right)^2dx\,dy,\\ \alpha_3&=\lim_{T\to\infty}\alpha_3^T.\end{aligned} \tag{1.4}$$

The reason that the limit exists and is positive and finite is that $p^t(x,y;\mathbf{T}^3)$ converges to the uniform density exponentially fast in t; see Proposition 3.1 for a discrete version of this statement.

Throughout the rest of the article, for functions f, g, we say that $f = O(g)$ if there exists constants c_1, c_2 such that $|f| \leq c_1 + c_2|g|$. We say that $f = \Omega(g)$ if there exists constants c_1, c_2 so that $|f| \geq c_1 + c_2|g|$. We say that $f = \Theta(g)$ if $f = O(g)$ and $f = \Omega(g)$. Finally, we say $f = o(g)$ if $\lim_{n\to\infty} f(n)/g(n) = 0$.

We note that the problem for $d = 1$ is trivial: $\mathrm{Var}(|\mathcal{A}_1|) = \Theta(n^2)$. Indeed, observe that with positive probability, the distance between X_1 and X_2 at time 0 is at least $\frac{1}{4}n$. In cn^2 steps (for c large enough), X_1 has positive probability of covering the entire cycle while X_2 has positive probability of not leaving an interval of length $\frac{1}{4}n$ containing its starting point. On this event, $|\mathcal{A}_1| \geq \frac{3}{4}n$. This proves our claim as the upper bound is trivial. For $d = 2$, the asymptotics of $\mathrm{Var}(|\mathcal{A}_1|)$ remains open.

One interesting remark is that the variance for $d = 3, 4$ is significantly higher than that of an i.i.d. marking. The results of Theorem 1.1 should also be contrasted with the behavior of the variance of the range $\mathcal{R}$ of random walk on $\mathbf{Z}^d$ run up to the *cover time* $T_{\mathrm{cov}}(\mathbf{Z}_n^d)$ of $\mathbf{Z}_n^d$, which is the expected amount of time it takes for a single random walk to visit every site. When $d \geq 3$, $T_{\mathrm{cov}}(\mathbf{Z}_n^d) \sim c_d n^d(\log n)$; see [13]. For $d \geq 5$, it follows from work of Jain and Orey [10] that $\mathrm{Var}(|\mathcal{R}|) = \Theta(n^d(\log n))$. For $d = 3, 4$, it follows from work of Jain and Pruitt [11] that $\mathrm{Var}(|\mathcal{R}|)$ is $\Theta(n^3(\log n)^2)$ and $\Theta(n^4(\log n))$, respectively.

This work opens the doors to many other problems involving two random walks. Natural next steps include CLTs for the fluctuations of $|\mathcal{A}_i|$ and for the number of sites painted by i at time t, as well as the development of an understanding of the geometrical properties of the clusters of $\mathcal{A}_i$. The latter seem to be connected to the theory of *random interlacements*. This is a model developed by Sznitman in [17] to describe the microscopic structure of the points visited by a random walk on $\mathbf{Z}_n^d$, $d \geq 3$, at times un^d for $u > 0$—that is, when a constant order of vertices have been visited. Roughly speaking, the model is a Poisson process on $W^* \times (0, \infty)$, where W^* is the space of doubly-infinite paths on $\mathbf{Z}^d$ modulo time-shifts. For a point (X, U) realized in this process, one should think of X as describing a random walk trajectory (an "interlacement") and U a time parameter. The model was first developed to study the process of disconnection of a discrete cylinder by random walk [6] and has been subsequently applied to understand the fine geometrical structure of random walk in many different settings [18, 19]. Sznitman's theory generalizes to the setting of k random walks by labeling each interlacement with an element of $\{1, \dots, k\}$ i.i.d. at random. Studying the structure of the clusters in the $\mathcal{A}_i$ using this general theory is an interesting research direction.

Theorem 1.1 is a special case of a much more general result, which gives the asymptotics of $\mathrm{Var}(|\mathcal{A}_i|)$ for many other graphs, such as the hypercube and the Caley graph of the symmetric group generated by transpositions. We will now review some additional terminology which is necessary to give a precise statement of the result. Recall that the *uniform mixing time* of random walk on G is

$$T_{\mathrm{mix}}(G) = \min\left\{t \geq 0 : \max_{x,y} \left| \frac{p^t(x, y; G)}{\pi(y; G)} - 1 \right| \leq \frac{1}{4}\right\},$$

and the *Green's function* for G is

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

that is, the expected amount of time X_i spends at y up until $T_{\mathrm{mix}}(G)$ when started from x. Let $\tau_i(x) = \min\{t \geq 0 : X_i(t) = x\}$ be the first time X_i hits x; we will omit i if there is only one random walk under consideration. Throughout the rest of the article, $a \wedge b = \min(a, b)$ for $a, b \in \mathbf{R}$.

ASSUMPTION 1.2. (G_n) is a sequence of vertex transitive graphs with $|V_n| \to \infty$ such that:

(1) $T_{\mathrm{mix}}(G_n) = o(|V_n|/(\log|V_n|)^2)$ and $\lim_{n\to\infty} T_{\mathrm{mix}}(G_n) = \infty$;
(2) $\sum_{y \neq x_0} g^2(x_0, y; G_n) = o(T_{\mathrm{mix}}(G_n)/\log|V_n|)$ for each $x_0 \in V_n$ fixed;
(3) there exists $\rho_0 < 1$ so that $\mathbf{P}_x[\tau(y) \wedge \tau(z) \leq T_{\mathrm{mix}}(G_n)] \leq \rho_0$ uniformly in n and $x, y, z \in V_n$ distinct.

The purpose of (1) is that in many cases we will perform union bounds over time-scales whose length is proportional to $T_{\mathrm{mix}}(G_n)$, and the hypothesis gives us explicit control on how the number of terms in these bounds relates to the size of V_n. Part (2) gives us control on the tail behavior of g and, finally, part (3) says that with uniformly positive probability the walks we consider do not hit adjacent points within the mixing time. Note that vertex transitivity implies g is constant along the diagonal. Part (3) implies that the number of times random walk started at x returns to x before the mixing time is stochastically dominated by a geometric random variable whose parameter depends only on ρ_0. Consequently, we see that there exists $g_0 > 0$ such that $g(x, x; G_n) \leq g_0$ uniformly in x and n.

Assume that (G_n) is a sequence of vertex transitive graphs, and let

$$f_{n,c}(x, y) = \mathbf{P}_x[\tau(y) \leq cT_{\mathrm{mix}}(G_n)], \tag{1.5}$$

$$\overline{f}_{n,c} = \sum_y f_{n,c}(x, y)\pi(y; G_n). \tag{1.6}$$

Note that $\overline{f}_{n,c}$ does not depend on the choice of x since if we replaced x with x', by vertex transitivity we may precompose $f_{n,c}$ with an automorphism of G_n which sends x to x'.

The general theorem is:

THEOREM 1.3. *Suppose that (G_n) satisfies Assumption 1.2. Let*

$$F_{n,c} = \sum_{x,y} \big(f_{n,c}(x, y) - \overline{f}_{n,c}\big)^2.$$

There exists $\gamma > 0$ so that for every $c \geq 2$, we have

$$\mathrm{Var}(|\mathcal{A}_i|) = \big(\tfrac{1}{4} + O(\Delta_n)\big)F_{n,c} + O\big(e^{-\gamma c}(T_{\mathrm{mix}}(G_n))^2\big) \tag{1.7}$$

as $n \to \infty$ where

$$\Delta_n = \frac{T_{\mathrm{mix}}(G_n)\log|V_n|}{|V_n|}.$$

Applying this to the special cases of the hypercube and the Caley graph of S_n generated by transpositions leads to the following corollary.

COROLLARY 1.4. *Suppose that $G_n = (V_n, E_n)$ is either the hypercube $\mathbf{Z}_2^n$ or the Caley graph of S_n generated by transpositions. Then*

$$\mathrm{Var}(|\mathcal{A}_i|) = \tfrac{1}{4}\big(1 + o(1)\big)|V_n|.$$

In particular, the first-order asymptotics of the variance are exactly the same as for an i.i.d. marking.

Throughout the remainder of the article, all graphs under consideration shall satisfy Assumption 1.2. In most examples, it will be that $T_{\mathrm{mix}}^2(G_n) = o(F_{n,c})$ so that the second term in (1.7) is negligible. In this case, taking $c = 2$ in (1.7) provides a means to compute not only the magnitude but also the constant in the first order asymptotics of the variance. In some cases, such as $G_n = \mathbf{Z}_n^3$, the constant can even be computed when $F_{n,c} = \Theta((T_{\mathrm{mix}}(G_n))^2)$.

The challenge in obtaining Theorems 1.1 and 1.3 is that the cancellation in the expansion of the variance is quite significant which, when ignored, yields only an upper bound that can be off by as much as a multiple of $T_{\mathrm{mix}}(G_n)$. We will now illustrate this point in the case of $\mathbf{Z}_n^d$ for $d \geq 3$. It will turn out that the contribution to the variance from the sites visited by both X_1, X_2 simultaneously is negligible, and hence we will ignore this possibility in the present discussion. Observe

$$\begin{aligned}\mathrm{Var}&\left(\sum_x \mathbf{1}_{\{\tau_1(x) < \tau_2(x)\}}\right)\\ &= \sum_{x,y}\big(\mathbf{P}[\tau_1(x) < \tau_2(x), \tau_1(y) < \tau_2(y)]\\ &\qquad - \mathbf{P}[\tau_1(x) < \tau_2(x)]\mathbf{P}[\tau_1(y) < \tau_2(y)]\big).\end{aligned}$$

Note that $\mathbf{P}[\tau_1(x) < \tau_2(x)]$ is approximately $\frac{1}{2}$. Let $H(x, y) = \{\tau_1(x) < \tau_1(y) \wedge \tau_2(x) \wedge \tau_2(y)\}$. Consequently, by symmetry, the above is approximately equal to

$$\sum_{x,y}\left(2\mathbf{P}[\tau_1(x) < \tau_2(x), \tau_1(y) < \tau_2(y) | H(x, y)]\mathbf{P}[H(x, y)] - \frac{1}{4}\right) + O(n^d).$$

The reason for the $O(n^d)$ term is that $\mathbf{P}[H(x, x)] = 0$, so all of the diagonal terms are ignored in the summation. Let $\widetilde{\pi}(\cdot; x, y)$ be the law of $X_2(\tau_1(x))$ conditional on $H(x, y)$. As $\mathbf{P}[H(x, y)]$ is approximately $\frac{1}{4}$, using the Markov property of (X_1, X_2) applied for the stopping time $\tau_1(x)$, we can rewrite the summation as

$$2\sum_{x,y,z}\left(\mathbf{P}_{x,z}[\tau_1(y) < \tau_2(y)] - \frac{1}{2}\right)\widetilde{\pi}(z; x, y)\mathbf{P}[H(x, y)].$$

Here, $\mathbf{P}_{x,z}$ denotes the joint law of X_1, X_2 with $X_1(0) = x$ and $X_2(0) = z$. Thus we need to estimate

$$\frac{1}{2}\sum_{x,y,z}\left(\mathbf{P}_{x,z}[\tau_1(y) < \tau_2(y)] - \frac{1}{2}\right)\widetilde{\pi}(z; x, y). \tag{1.8}$$

At this point, one is tempted to insert absolute values and then work on each of the summands separately. Since X_1 and X_2 are independent, note that $X_2(\tau_1(x)) \sim \pi(\cdot; \mathbf{Z}_n^d)$. Thus by Bayes' rule, we have

$$\widetilde{\pi}(z; x, y) = \frac{\mathbf{P}[H(x, y) | X_2(\tau_1(x)) = z]}{\mathbf{P}[H(x, y)]} \pi(z; \mathbf{Z}_n^d) \leq C_0 \pi(z; \mathbf{Z}_n^d);$$

see Theorem 4.1 for a much finer estimate. Hence the expression in (1.8) is bounded from above by

$$C_1 \sum_{x,y} \left| \mathbf{P}_{x,\pi}[\tau_1(y) < \tau_2(y)] - \frac{1}{2} \right|, \tag{1.9}$$

where $\mathbf{P}_{x,\pi}$ denotes the law of X_1, X_2 with $X_1(0) = x$ and $X_2(0) \sim \pi(\cdot; \mathbf{Z}_n^d)$.

It is a basic fact that $T_{\text{mix}}(\mathbf{Z}_n^d) = \Theta(n^2)$; one way to see this is to invoke the local central limit theorem ([12], Theorem 1.2.1). We can analyze $\mathbf{P}_{x,\pi}[\tau_1(y) < \tau_2(y)]$ as follows. We consider two different cases: either y is hit before time $t_c \equiv c T_{\text{mix}}(\mathbf{Z}_n^d) = c' n^2$ or afterward. The probability that X_2 hits y before t_c is of order n^{2-d} by a union bound since $X_2(t) \sim \pi(\cdot; \mathbf{Z}_n^d) = n^{-d}$ for all t. Second, by the local transience of random walk on $\mathbf{Z}_n^d$ for $d \geq 3$, the probability that X_1 hits y before t_c is, up to a multiplicative constant, well approximated by $g(x, y; \mathbf{Z}_n^d)$. We now consider the second case. By time t_c for $c > 0$ large enough, X_1 will have mixed. This means that if neither X_1 nor X_2 has hit y by this time, the probability that either one hits first is close to $1/2$. The careful reader who wishes to see precise, quantitative versions of these statements will find such in the lemmas we use to prove Theorem 1.3. Thus it is not difficult to see that there exists $C_2 > 0$ so that

$$|\mathbf{P}_{x,\pi}[\tau_1(y) < \tau_2(y)] - 1/2| \leq C_2 g(x, y; \mathbf{Z}_n^d).$$

This leads to an upper bound of

$$C_3 \sum_{x,y} g(x, y; \mathbf{Z}_n^d) \leq C_4 n^{d+2}.$$

A slightly more refined analysis leads to a lower bound of (1.9) with the same growth rate. As we will show in the next section, in every dimension this estimate is typically quite far from being sharp. The reason for the inaccuracy is that by moving the absolute value into the sum in (1.9) we are unable to take advantage of the cancellation that arises as $\mathbf{P}_{x,\pi}[\tau_1(y) < \tau_2(y)] > 1/2$ when x is close to y and $\mathbf{P}_{x,\pi}[\tau_1(y) < \tau_2(y)] < 1/2$ when x is far from y.

Outline. The remainder of this article is structured as follows. In the next section, we will deduce Theorem 1.1 and Corollary 1.4 from Theorem 1.3. In Section 3, we introduce some notation that will be used throughout in addition to collecting several basic random walk estimates. Next, in Section 4, we give a precise estimate of the Radon–Nikodym derivative of $\widetilde{\pi}(\cdot; x, y)$ with respect to π. In Section 5, we prove Theorem 1.3 and end in Section 6 with a list of related problems and discussion.

2. Proof of Theorem 1.1 and Corollary 1.4. The following proposition will be important for the proof of Theorem 1.1.

PROPOSITION 2.1. *Assume that $G_n = \mathbf{Z}_n^d$ for $d \geq 3$. For each $c > 1$, the limit*

$$\lim_{n\to\infty} \frac{1}{h_d(n)} \sum_{x,y} \big(f_{n,c}(x,y) - \overline{f}_{n,c}\big)^2 \tag{2.1}$$

exists. When $d \geq 4$, it is α_d as in (1.2), (1.3). *When $d = 3$, it is given by α_3^c where α_3^T is as in* (1.4).

The first step in the proof of the proposition is to reduce the existence of the limit to a computation involving Green's functions. Recall from (1.1) that $G(y; \mathbf{Z}^d)$ is the Green's function for lazy random walk on $\mathbf{Z}^d$. In order to keep the notation from becoming too heavy, throughout the rest of this section we will write T_{mix} for $T_{\text{mix}}(G_n)$ where G_n will be clear from the context. Let

$$g_c(x, y; G_n) = \mathbf{E}_x \sum_{t=0}^{cT_{\text{mix}}} \mathbf{1}_{\{X(t)=y\}}.$$

LEMMA 2.2. *Assume that $G_n = \mathbf{Z}_n^d$ for $d \geq 3$. For each $c > 1$, we have that*

$$\lim_{n\to\infty} \frac{1}{h_d(n)} \sum_{x,y} \big(f_{n,c}(x,y) - G^{-1}(0; \mathbf{Z}^d) g_c(x, y; \mathbf{Z}_n^d)\big)^2 = 0.$$

PROOF. Observe

$$g_c(x, y; \mathbf{Z}_n^d) \leq f_{n,c}(x, y) g_c(y, y; \mathbf{Z}_n^d).$$

We shall now prove a matching lower bound. Fix $0 < \widetilde{c} < c$. Then we have that

$$\begin{aligned} g_c(x, y; \mathbf{Z}_n^d) &\geq \mathbf{E}_x\left[\left(\sum_{t=\tau(y)}^{cT_{\text{mix}}} \mathbf{1}_{\{X(t)=y\}}\right) \mathbf{1}_{\{\tau(y) \leq (c-\widetilde{c})T_{\text{mix}}\}}\right] \\ &\geq f_{n,c-\widetilde{c}}(x, y) g_{\widetilde{c}}(y, y; \mathbf{Z}_n^d). \end{aligned} \tag{2.2}$$

Assuming $c - \widetilde{c} > 1$, by mixing considerations as well as a union bound (see Proposition 3.1) we have that

$$\begin{aligned} f_{n,c-\widetilde{c}}(x, y) &= f_{n,c}(x, y) - \mathbf{P}_x[(c - \widetilde{c})T_{\text{mix}} < \tau(y) \leq cT_{\text{mix}}] \\ &= f_{n,c}(x, y) + O(\widetilde{c} n^{2-d}). \end{aligned} \tag{2.3}$$

Since $\widetilde{c} > 0$, we have

$$\begin{aligned} g_{\widetilde{c}}(y, y; \mathbf{Z}_n^d) &= g_c(y, y; \mathbf{Z}_n^d) - \sum_z p^{\widetilde{c}T_{\text{mix}}}(y, z; \mathbf{Z}_n^d) g_{c-\widetilde{c}}(z, y; \mathbf{Z}_n^d) \\ &= g_c(y, y; \mathbf{Z}_n^d) + O\big((c - \widetilde{c})\widetilde{c}^{-d/2} n^{2-d}\big), \end{aligned} \tag{2.4}$$

where we used in the last line that $p^t(z, y; \mathbf{Z}_n^d) \le c_1 t^{-d/2}$ for some $c_1 > 0$ (see [12], Theorem 1.2.1) as well as the observation $\sum_z g_{c-\widetilde{c}}(z, y; \mathbf{Z}_n^d) = (c - \widetilde{c})T_{\mathrm{mix}}$. Combining (2.2), (2.3) and (2.4), we have thus proved the lower bound

$$g_c(x, y; \mathbf{Z}_n^d) \ge f_{n,c}(x, y) g_c(y, y; \mathbf{Z}_n^d) + O\big((c - \widetilde{c})\widetilde{c}^{-d/2}(\widetilde{c} + g_c(x, y; \mathbf{Z}_n^d))n^{2-d}\big).$$

Here, we used the bound $f_{n,c}(x, y) \le g_c(x, y; \mathbf{Z}_n^d)$. Theorem 1.5.4 of [12] implies $g_c(x, y; \mathbf{Z}_n^d) = \Theta(c|x - y|^{2-d})$ (it is actually stated for walks on $\mathbf{Z}^d$ which are not lazy, but the generalization is straightforward). Consequently,

$$\sum_y g_c^2(x, y; \mathbf{Z}_n^d) = \begin{cases} \Theta(n), & \text{if } d = 3, \\ \Theta(\log n), & \text{if } d = 4, \text{ and} \\ \Theta(1), & \text{if } d \ge 5. \end{cases}$$

Hence,

$$\begin{aligned} &\sum_{x,y}\big(f_{n,c}(x, y) g_c(y, y; \mathbf{Z}_n^d) - g_c(x, y; \mathbf{Z}_n^d)\big)^2 \\ &\qquad = \sum_{x,y}\big[O\big((c - \widetilde{c})\widetilde{c}^{-d/2}(\widetilde{c} + g(x, y; \mathbf{Z}_n^d))n^{2-d}\big)\big]^2 \\ &\qquad = O\big((c - \widetilde{c})^2\widetilde{c}^{-d}(\widetilde{c}^2 + o(1))n^4\big). \end{aligned}$$

Dividing both sides by $h_d(n)$, taking a limsup as $n \to \infty$, then as $\widetilde{c} \to 0$ yields

$$\lim_{n\to\infty} \frac{1}{h_d(n)} \sum_{x,y}\big(f_{n,c}(x, y) g_c(y, y; \mathbf{Z}_n^d) - g_c(x, y; \mathbf{Z}_n^d)\big)^2 = 0.$$

By (2.4) we know that $|g_c(y, y; \mathbf{Z}_n^d) - g_1(y, y; \mathbf{Z}_n^d)| = o(1)$, and, by local transience, it is not hard to see that $\lim_{n\to\infty} g_1(y, y; \mathbf{Z}_n^d) = G(0; \mathbf{Z}^d)$. $\square$

PROOF OF PROPOSITION 2.1. Lemma 2.2 implies that we may replace $f_{n,c}(x, y)$ by $G^{-1}(0; \mathbf{Z}^d) g_c(x, y; \mathbf{Z}_n^d)$ in (2.1). Letting $\overline{g}_{n,c} = cT_{\mathrm{mix}} n^{-d}$, we can likewise replace $\overline{f}_{n,c}$ in (2.1) by $G^{-1}(0; \mathbf{Z}^d)\overline{g}_{n,c}$. Consequently, to prove the proposition, it suffices to prove the existence of the limit

$$\lim_{n\to\infty} \frac{1}{h_d(n)} \sum_{x,y}\big(g_c(x, y; \mathbf{Z}_n^d) - \overline{g}_{n,c}\big)^2. \tag{2.5}$$

We will divide the proof into the cases $d \ge 4$ and $d = 3$.

Case 1: $d \ge 4$. As $\overline{g}_{n,c} = O(cn^{2-d})$, we have

$$\frac{1}{h_d(n)} \sum_{x,y}\big(\overline{g}_{n,c}^2 + 2\overline{g}_{n,c} g_c(x, y; \mathbf{Z}_n^d)\big) = o(1).$$

Thus it suffices to show in this case that

$$\lim_{n\to\infty} \frac{1}{\widetilde{h}_d(n)} \sum_y g_c^2(0, y; \mathbf{Z}_n^d)$$

exists, where $\widetilde{h}_4(n) = \log n$ and $\widetilde{h}_d(n) = 1$ for $d \geq 5$. This will be a consequence of two observations. First, note that

$$\sum_{|y| \geq \ell} g_c^2(0, y; \mathbf{Z}_n^d) = \sum_{|y| \geq \ell} O(c|y|^{4-2d}) = \sum_{m=\ell}^{n} O(cm^{4-2d} \cdot m^{d-1})$$
$$= \begin{cases} O(c\ell^{-1}), & \text{if } d \geq 5, \\ O\big(c \log(n/\ell)\big), & \text{if } d = 4. \end{cases}$$

Thus it suffices to show that, for $\ell = \ell(n, \varepsilon) = n^{1-\varepsilon}$ with $\varepsilon > 0$, the limit

$$\lim_{\varepsilon \to 0} \lim_{n \to \infty} \frac{1}{\widetilde{h}_d(n)} \sum_{|y| \leq \ell} g_c^2(0, y; \mathbf{Z}_n^d)$$

exists (we can even restrict to finite ℓ if $d \geq 5$). Our second observation is that

$$g_c(0, y; \mathbf{Z}_n^d) - G(y; \mathbf{Z}^d) = O(cn^{2-d}) \qquad \text{for } |y| \leq \ell.$$

This follows since we can couple the walks on $\mathbf{Z}_n^d$ and $\mathbf{Z}^d$ starting at 0 such that they are the same until the first time τ_0 they have reached distance $n/2$ from 0, then move independently thereafter. The expected number of visits each walk makes to y after time τ_0, where the former is stopped at time cT_{mix}, is easily seen to be $O(cn^{2-d})$. Thus,

$$\sum_{|y| \leq \ell} \big(g_c(0, y; \mathbf{Z}_n^d) - G(y; \mathbf{Z}^d)\big)^2 = o(1).$$

Therefore if $d \geq 5$, we have

$$\lim_{n \to \infty} \frac{1}{h_d(n)} \sum_{x,y} g_c^2(x, y; \mathbf{Z}_n^d) = \sum_{y \in \mathbf{Z}^d} G^2(y; \mathbf{Z}^d).$$

For $d = 4$,

$$\lim_{n \to \infty} \frac{1}{h_4(n)} \sum_{x,y} g_c^2(x, y; \mathbf{Z}_n^4) = \lim_{n \to \infty} \frac{1}{\log n} \sum_{y \in \mathbf{Z}^4, |y| \leq n} G^2(y; \mathbf{Z}^4).$$

Note that the limit on the right-hand side exists since by Theorem 1.5.4 of [12] (generalized to lazy walks)

$$G(y; \mathbf{Z}^d) = a_d |y|^{2-d} + o(|y|^{-\alpha}),$$

if $\alpha \in (0, d)$ is fixed.

Case 2: $d = 3$. The thrust of the previous argument was that random walk on $\mathbf{Z}_n^d$ for $d \geq 4$ is sufficiently transient so that pairs of points of distance $\Omega(n^{1-\varepsilon})$ make a negligible contribution to the variance, which in turn allowed us to make an accurate comparison between the Green's function for random walk on $\mathbf{Z}_n^d$ with that on $\mathbf{Z}^d$. The situation for $d = 3$ is more delicate since the opposite is true: pairs of distance $O(n^{1-\varepsilon})$ do not measurably affect the variance.

Theorem 1.2.1 of [12] (extended to the case of lazy random walk, see also Corollary 22.3 of [1]) implies the existence of constants $\beta_3, \gamma_3 > 0$ such that with $\overline{p}^t(x, y; \mathbf{Z}^3) = \frac{\beta_3}{t^{3/2}} \exp(-\frac{\gamma_3 |x-y|^2}{t})$, we have the estimate

$$|\overline{p}^t(x, y; \mathbf{Z}^3) - p^t(x, y; \mathbf{Z}^3)| = |x-y|^{-2} O(t^{-3/2}).$$

Hence letting $\overline{p}^t(x, y; \mathbf{Z}_n^3) = \sum_{k \in \mathbf{Z}^3} \overline{p}^t(x, y + kn; \mathbf{Z}^3)$, one can easily show that with

$$\Delta(x, y) \equiv \sum_{t=0}^{cT_{\text{mix}}} |\overline{p}^t(x, y; \mathbf{Z}_n^3) - p^t(x, y; \mathbf{Z}_n^3)|$$

we have that

$$\frac{1}{h_3(n)} \sum_{x,y} \Delta^2(x, y) = o(1). \tag{2.6}$$

By differentiating $\overline{p}$ in t, we see that for $1 \le t \le s \le t+1$, we have

$$\begin{aligned} &|\overline{p}^s(0, y; \mathbf{Z}_n^3) - \overline{p}^t(0, y; \mathbf{Z}_n^3)| \\ &\quad = O\biggl(\frac{\overline{p}^t(0, y; \mathbf{Z}_n^3)}{t} + \sum_k \frac{|y+kn|^2}{t^2} \overline{p}^t(0, y+kn; \mathbf{Z}_n^3)\biggr). \end{aligned}$$

We are now going to prove that

$$\sum_{y \in \mathbf{Z}_n^3} \biggl(\int_1^{cT_{\text{mix}}} \overline{p}^t(0, y; \mathbf{Z}_n^3) - \overline{p}^{\lfloor t \rfloor}(0, y; \mathbf{Z}_n^3)\, dt\biggr)^2 = O(1). \tag{2.7}$$

It suffices to bound

$$A \equiv \sum_{y \in \mathbf{Z}_n^3} \biggl(\int_1^{cT_{\text{mix}}} \frac{1}{t} \overline{p}^t(0, y; \mathbf{Z}_n^3)\, dt\biggr)^2,$$

$$B \equiv \sum_{y \in \mathbf{Z}_n^3} \biggl(\sum_k \int_1^{cT_{\text{mix}}} \frac{|y+kn|^2}{t^2} \overline{p}^t(0, y+kn; \mathbf{Z}_n^3)\, dt\biggr)^2.$$

For A, we apply Cauchy–Schwarz to the integral and invoke the integrability of $1/t^2$ over $[1, \infty)$ to arrive at

$$A \le C_2 \sum_{y \in \mathbf{Z}_n^3} \int_1^{cT_{\text{mix}}} [\overline{p}^t(0, y; \mathbf{Z}_n^3)]^2\, dt = O(1).$$

For B, we insert the formula for $\overline{p}$ into the integral, make the substitution $u = |y+kn|^2/t$ and then compute to see

$$B \le C_3 \sum_{y \in \mathbf{Z}_n^3} \frac{1}{|y|^6 + 1} = O(1).$$

This proves (2.7). Recall that $\mathbf{T}^3$ is the three-dimensional continuum torus. For $x, y \in \mathbf{T}^3$, let

$$g_c(x, y; \mathbf{T}^3) = \int_0^{cT_{\mathrm{mix}}} \overline{p}^t(nx, ny; \mathbf{Z}_n^3)\, dt = \frac{1}{n} \int_0^{cT} \overline{p}^u(x, y; \mathbf{T}^3)\, du, \tag{2.8}$$

where $T = T_{\mathrm{mix}}/n^2$. By (2.6), (2.7), we have that

$$\frac{1}{h_3(n)} \sum_{x,y \in \mathbf{Z}_n^3} \big(g_c(x, y; \mathbf{Z}_n^3) - g_c(x/n, y/n; \mathbf{T}^3)\big)^2 = o(1).$$

Therefore we may replace $g_c(x, y; \mathbf{Z}_n^3)$ in (2.5) with $g_c(x/n, y/n; \mathbf{T}^3)$. Note that $g_c(\cdot, \cdot; \mathbf{T}^3)$ is the product of n^{-1}, and the Green's function for $B_{t/2}$, where B_t is a Brownian motion on $\mathbf{T}^3$; roughly, the reason that the Brownian motion moves at 1/2-speed is that a lazy random walk moves at 1/2 the speed of a simple random walk. It is left to bound

$$n^2 \int_{\mathbf{T}^3} \int_{\mathbf{T}^3} \big(g_c(\lfloor nx \rfloor/n, \lfloor ny \rfloor/n; \mathbf{T}^3) - g_c(x, y; \mathbf{T}^3)\big)^2\, dx\, dy;$$

the reason for the pre-factor n^2 is that we need to multiply by $(n^3)^2$ in order to make the double integral comparable to the double summation, and we also divide by the normalization $h_3(n)$. From (2.8), we see that $g_c(x, y; \mathbf{T}^3)$ is $O(n^{-1})$-Lipschitz away from the diagonal $D_\varepsilon = \{(x, y) \in \mathbf{T}^3 \times \mathbf{T}^3 : |x - y| \le \varepsilon\}$. Thus since $|(x, y) - (\lfloor nx \rfloor/n, \lfloor ny \rfloor/n)| = O(n^{-1})$, the integrand is $O(n^{-4})$ on D_ε^c, hence the integral over D_ε^c is $O(n^{-2})$. Since both $ng_c(\lfloor nx \rfloor/n, \lfloor ny \rfloor/n; \mathbf{T}^3)$ and $ng_c(x, y; \mathbf{T}^3)$ are uniformly L^2-integrable over $\mathbf{T}^3 \times \mathbf{T}^3$, it follows that the contribution coming from D_ε can be made uniformly small in n by first fixing $\varepsilon > 0$ small enough. $\square$

We now deduce Theorem 1.1 from Theorem 1.3.

PROOF OF THEOREM 1.1. Suppose $G_n = \mathbf{Z}_n^d$ for $d \ge 3$. Recall that

$$T_{\mathrm{mix}}(\mathbf{Z}_n^d) = \Theta(n^2)$$

(see [13]) and there exists $c_d > 0$ so that $g(x, y; \mathbf{Z}_n^d) \le c_d |x - y|^{2-d} \wedge 1$ (see [12]). Consequently, the hypotheses of Theorem 1.3 are obviously satisfied, except for possibly (3). This is easy to see if x is sufficiently far from y, z so that $g(x, y; \mathbf{Z}_n^d) + g(x, z; \mathbf{Z}_n^d) \le 1/2$. Now suppose that $|x - y| \wedge |x - z| = r$ is small enough so that $g(x, y; \mathbf{Z}_n^d) + g(x, z; \mathbf{Z}_n^d) > 1/2$. We have the trivial bound that X starting at x will get to distance $r + s$ without hitting y, z in s steps with probability at least $(4d)^{-s}$ since in each step, X has probability at least $(4d)^{-1}$ of increasing its distance from y, z by 1. If s is large enough, then after such steps we will have $g(X_s, y; \mathbf{Z}_n^d) + g(X_s, z; \mathbf{Z}_n^d) \le 1/2$, which gives the desired result.

Proposition 2.1 implies that $F_{n,c} \sim \frac{1}{4}\alpha_{d,c}h_d(n)$ as $n \to \infty$. This is enough to dominate $T_{\mathrm{mix}}^2(\mathbf{Z}_n^d) = \Theta(n^4)$ except if $d = 3$. We shall now argue that, nevertheless, $F_{n,c}$ is still the dominant term in this case. Note that

$$\overline{f}_{n,c} \leq \frac{1}{n^3}\sum_y g_c(x, y; \mathbf{Z}_n^3) \leq A_0 c n^{-1}$$

for some $A_0 > 0$ and $c \geq 2$ fixed. Also, the transience of random walk on $\mathbf{Z}_n^3$ implies that there exists $A_1 > 0$ so that $f_{n,c}(x, y) \geq A_1|x - y|^{-1} \wedge 1$. Thus for

$$|x - y| \leq \left(\frac{A_1}{2A_0 c}\right)n \equiv A_2 n$$

we have that $f_{n,c}(x, y) - \overline{f}_{n,c} \geq \frac{A_1}{2}|x - y|^{-1} \wedge 1$. Consequently,

$$F_{n,c} \geq \frac{A_1^2}{4}\sum_{|x-y|\leq A_2 n}|x - y|^{-2} \wedge 1 = c^{-1}\Theta(n^4).$$

A matching upper bound, up to a multiplicative factor, is also not difficult to see.

Our lower bound for $F_{n,c}$ depends on c by a multiplicative factor of $1/c$ while the second term in (1.7) decays exponentially in c. Thus by taking $c \geq 2$ large enough, we see that $F_{n,c}$ is still dominant for $d = 3$. $\square$

We now turn to the proof of Corollary 1.4.

PROOF OF COROLLARY 1.4 FOR THE HYPERCUBE. For $\mathbf{Z}_2^n$, it is easier to work with the continuous time random walk (CTRW) since the types of estimates we require 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 [13]) which implies $\Omega(n) = T_{\mathrm{mix}}(\mathbf{Z}_2^n) = O(n^2)$ (see Theorem 12.3 of [13]). Consequently, the first hypothesis of Theorem 1.3 holds. If $|x - y| = r$, then it is easy to see there exists $C_\varepsilon, \rho_\varepsilon > 0$ so that

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

provided $\varepsilon > 0$ is sufficiently small. Thus it is not difficult to see that $g(x, y; \mathbf{Z}_2^n) \leq C'_\varepsilon n^{-r}$. Trivially,

$$\left|\{y \in \mathbf{Z}_2^n : |x - y| = r\}\right| = \binom{n}{r} \leq n^r.$$

Thus for x_0 fixed we have

$$\sum_{y \neq x_0} g^2(x_0, y; \mathbf{Z}_2^n) \leq O\left(\sum_{r=1}^{n} n^{-2r} \cdot n^r\right) = O\left(\frac{1}{n}\right),$$

so the second hypothesis of Theorem 1.3 is satisfied. The final hypothesis is obviously also satisfied. Now, a union bound implies that $\overline{f}_{n,c} = O(2^{-n} T_{\mathrm{mix}}(\mathbf{Z}_2^n))$, which implies $(f_{n,c}(x,x) - \overline{f}_{n,c})^2 = 1 + o(1)$. On the other hand,

$$\sum_{|x-y| \geq 1} f_{n,c}^2(x, y) = O\left(2^n \sum_{r=1}^{n} n^{-2r} \cdot n^r\right) = o(2^n).$$

Putting everything together, Theorem 1.3 implies

$$\mathrm{Var}(|\mathcal{A}_i|) = \tfrac{1}{4}\big(1 + o(1)\big)2^n. \qquad \square$$

PROOF OF COROLLARY 1.4 FOR THE CALEY GRAPH OF S_n. Let G_n be the Caley graph of S_n generated by transpositions. By work of Diaconis and Shashahani [7], the total variation mixing time of S_n is $\Theta(n \log n)$, which by Theorem 12.3 of [13] implies $T_{\mathrm{mix}}(G_n) = O(n(\log n)(\log n!)) = O(n^2(\log n)^2)$. 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 $1 \leq i \leq r$. Then the automorphism of G_n induced by conjugation by α satisfies $\alpha\tau\alpha^{-1} \neq \tau$. 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 $\binom{n}{2r} \geq 2^{-2r} n^{2r} ((2r)!)^{-1}$, assuming $n \geq 4r$. Therefore,

$$p^t(e, \tau; G_n) \leq \frac{2^{2r}(2r)!}{n^{2r}} \quad \text{and} \quad g(e, \tau; G_n) \leq C(2^{2r}(2r)!)(\log n)^2 n^{2-2r}. \tag{2.9}$$

This bound is good enough for $r \geq 2$, but does not quite suffice when $r = 1$. This case is not difficult to handle, however, since it is easy to see that the random walk has distance 3 from e with probability $1 - O(1/n)$ after its first three moves, hence with distance at least 2 from any permutation with distance 1 from e. Combining this with (2.9) gives a bound on $g(e, \tau; G_n)$ for all $\tau \in S_n$. From this is it clear that (G_n) satisfies the hypotheses of Theorem 1.3 and, arguing as in the case of the hypercube, that

$$\mathrm{Var}(|\mathcal{A}_i|) = \tfrac{1}{4}\big(1 + o(1)\big)n!. \qquad \square$$

3. Preliminaries.

3.1. *Notation.* Suppose that $G = (V, E)$ is a graph, and let X_1, X_2 be independent random walks on G. Recall that $a \wedge b = \min(a, b)$ for $a, b \in \mathbf{R}$. For $x, y \in V$, let

$$\tau_i(x,y) = \tau_i(x) \wedge \tau_i(y) \quad \text{and} \quad \tau(x,y) = \tau_1(x,y) \wedge \tau_2(x,y),$$

where $\tau_i(x) = \min\{t \geq 0 : X_i(t) = x\}$. Let

$$H(x,y) = \{\tau_1(x) < \tau_1(y) \wedge \tau_2(x,y)\}.$$

This is the event that x is hit by X_1 before X_2 as well as before both X_1, X_2 hit y. Let

$$\widetilde{\pi}(z;x,y) = \mathbf{P}[X_2(\tau_1(x,y)) = z | H(x,y)],$$

and let π be the uniform measure on V. Throughout, $\mathbf{P}_z[\cdot]$ denotes the law of random walk initialized at z (and the initial distribution is stationary whenever z is omitted). The proofs in this article will involve probabilities of complicated events. To keep the formulas succinct, it will be helpful for us to introduce the following notation: let

$$G_{ij}(x) = \{\tau_i(x) < \tau_j(x)\},$$
$$G_{ij}(x,y) = \{\tau_i(x,y) < \tau_j(x,y)\}$$

and

$$G_i(x,y) = \{\tau_i(x) < \tau_i(y)\}.$$

Throughout we will fix a sequence of graphs (G_n) satisfying Assumption 1.2. We let

$$\Gamma_n = c_0 T_{\text{mix}}(G_n) \log|V_n|, \qquad \Upsilon_n = \frac{T_{\text{mix}}(G_n)}{|V_n|},$$
$$\Delta_n = \Upsilon_n \log|V_n|, \qquad \mathcal{S}_n = \sum_{y \neq x_0} g^2(x_0, y; G_n),$$

where c_0 will be determined later, and x_0 is fixed. Note that $\mathcal{S}_n$ does not depend on x_0 by vertex transitivity. We will typically write T_{mix} for $T_{\text{mix}}(G_n)$, $p^t(\cdot,\cdot)$ for $p^t(\cdot,\cdot;G_n)$ and $g(\cdot,\cdot)$ for $g(\cdot,\cdot;G_n)$ in order to keep the notation light and, in general, suppress dependencies on n.

3.2. *Elementary estimates.* Recall that the total variation distance of probability measures μ, ν on V is

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

The following provides a bound on the rate of decay of the distance of $p^t(x,\cdot)$ to stationarity.

PROPOSITION 3.1. *For every* $s, t \in \mathbf{N}$,

$$\max_x \|p^{t+s}(x,\cdot) - \pi\|_{\mathrm{TV}} \le 4 \max_{x,y} \|p^t(x,\cdot) - \pi\|_{\mathrm{TV}} \|p^s(y,\cdot) - \pi\|_{\mathrm{TV}}, \tag{3.1}$$

$$\max_{x,y} \left| \frac{p^{t+s}(x,y)}{\pi(y)} - 1 \right| \le \max_{x,y} \frac{p^s(x,y)}{\pi(y)} \max_x \|p^t(x,\cdot) - \pi\|_{\mathrm{TV}}. \tag{3.2}$$

PROOF. The first part is a standard result; see, for example, Lemmas 4.11 and 4.12 of [13]. The second part is a consequence of the semigroup property:

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

Trivially,

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

Consequently, (3.1) and (3.2) give

$$\max_{x,y} \left| \frac{p^t(x,y)}{\pi(y)} - 1 \right| \le \overline{\gamma} e^{-\overline{\gamma}\alpha} \qquad \text{for } t \ge \alpha T_{\mathrm{mix}} \text{ and } \alpha > 0, \tag{3.3}$$

where $\overline{\gamma} > 0$ is a universal constant. We will often use (3.3) without reference.

Throughout the article, it will be important for us to have precise estimates of the Radon–Nikodym derivative of the law of random walk conditioned on various events with respect to the uniform measure. In the following, we are interested in the case of a random walk conditioned not to have hit a particular point. Let $T_k = kT_{\mathrm{mix}}$.

LEMMA 3.2. *There exists* $\gamma, p_0 > 0$ *so that for all* $k \ge 1$ *satisfying* $k\Upsilon_n \le p_0$ *and* $c \ge 2$, *we have*

$$\mathbf{P}_x[X(cT_k) = z | \tau(y) > cT_k] = \big[1 + O\big(e^{-\gamma ck} + ck\Upsilon_n + g(y,z)\big)\big]\pi(z).$$

Note that by part (1) of Assumption 1.2, this lemma applies if

$$k = O((\log|V_n|)^2).$$

PROOF OF LEMMA 3.2. Using that $\mathbf{P}_x[X(cT_k) = z] = (1 + O(e^{-\overline{\gamma}ck}))\pi(z)$, an application of Bayes' formula yields

$$\begin{aligned}\mathbf{P}_x&[X(cT_k) = z|\tau(y) > cT_k]\\&= \frac{\mathbf{P}_x[\tau(y) > cT_k|X(cT_k) = z]}{\mathbf{P}_x[\tau(y) > cT_k]}(1 + O(e^{-\overline{\gamma}ck}))\pi(z).\end{aligned}$$

The idea of the rest of the proof is to show it is unlikely that X hits y close to time cT_k, in which case we can use a mixing argument to show that conditioning on $X(cT_k) = z$ has little effect. For $1 \le \widetilde{c} \le \widetilde{c} + 1 \le c$, we have

$$\begin{aligned}\mathbf{P}_x&[\tau(y) > cT_k|X(cT_k) = z]\\&= \mathbf{P}_x[\tau(y) > \widetilde{c}T_k|X(cT_k) = z] - \mathbf{P}_x[cT_k \ge \tau(y) > \widetilde{c}T_k|X(cT_k) = z].\end{aligned}$$

By a time-reversal, we have that

$$\mathbf{P}_x[cT_k \ge \tau(y) > \widetilde{c}T_k|X(cT_k) = z] \le \mathbf{P}_z[\tau(y) < (c - \widetilde{c})T_k|X(cT_k) = x].$$

By mixing considerations and a union bound, we have

$$\mathbf{P}_z[\tau(y) \le (c - \widetilde{c})T_k|X(cT_k) = x] = O\big(g(y, z) + (c - \widetilde{c})k\Upsilon_n\big).$$

Applying Bayes' formula, observe

$$\begin{aligned}\mathbf{P}_x[\tau(y) > \widetilde{c}T_k|X(cT_k) = z] &= \frac{\mathbf{P}_x[X(cT_k) = z|\tau(y) > \widetilde{c}T_k]}{\mathbf{P}_x[X(cT_k) = z]}\mathbf{P}_x[\tau(y) > \widetilde{c}T_k]\\&= \big(1 + O\big(e^{-k\overline{\gamma}(c-\widetilde{c})}\big)\big)\mathbf{P}_x[\tau(y) > \widetilde{c}T_k].\end{aligned}$$

Similarly,

$$\mathbf{P}_x[\tau(y) > cT_k] = \mathbf{P}_x[\tau(y) > \widetilde{c}T_k] - \mathbf{P}_x[cT_k \ge \tau(y) > \widetilde{c}T_k].$$

By a union bound and mixing considerations, the second term on the right-hand side is of order $O((c - \widetilde{c})k\Upsilon_n)$. We now take $\widetilde{c} = c/2$ and $\gamma = \overline{\gamma}/2$. By part (3) of Assumption 1.2, we have that

$$\mathbf{P}_x[\tau(y) > cT_k] \ge 1 - \rho_0 - O(ck\Upsilon_n)$$

uniformly in n. In particular, there exists $p_0 > 0$ so that if $ck\Upsilon_n \le p_0$, then $\mathbf{P}_x[\tau(y) > cT_k]$ is uniformly positive in n. Putting everything together, for such k, we thus have

$$\begin{aligned}\mathbf{P}_x&[X(cT_k) = z|\tau(y) > cT_k]\\&= \frac{(1 + O(e^{-\gamma ck}))\mathbf{P}_x[\tau(y) > \widetilde{c}T_k] + O(g(y, z) + ck\Upsilon_n)}{\mathbf{P}_x[\tau(y) > \widetilde{c}T_k] + O(ck\Upsilon_n)}\\&\quad\times \big(1 + O(e^{-\overline{\gamma}ck})\big)\pi(z)\\&= \big(1 + O\big(e^{-\gamma ck} + g(y, z) + ck\Upsilon_n\big)\big)\pi(z)\end{aligned}$$

as desired. $\square$

In the following lemma, we will show that the difference in the probability that a random walk hits points y, z when started from x before time $\Gamma_n = c_0(\log|V_n|)T_{\text{mix}}$ is essentially determined by the corresponding difference except up to time cT_{mix}. The reason for the cancellation is that the previous lemma implies that conditional on not hitting a given point up to time cT_{mix}, the walk is well mixed and has long forgotten its starting point. Recall that $f_c(x, y) = \mathbf{P}_x[\tau(y) \le cT_{\text{mix}}]$ (we have suppressed n).

LEMMA 3.3. *There exists $\gamma > 0$ such that for all $c \ge 2$,*

$$\begin{aligned}\mathbf{P}_x[\tau(y) \le \Gamma_n] - \mathbf{P}_x[\tau(z) \le \Gamma_n] \\ = f_c(x, y) - f_c(x, z) + O\big(e^{-\gamma c}\Upsilon_n + \Delta_n[g(x, y) + g(x, z)]\big).\end{aligned}$$

PROOF. We observe

$$\begin{aligned}&\mathbf{P}_x[\tau(y) \le \Gamma_n] \\ &\quad = f_c(x, y) + \sum_k \mathbf{P}_x[cT_k < \tau(y) \le cT_{k+1} | \tau(y) > cT_k]\big(1 - \mathbf{P}_x[\tau(y) \le cT_k]\big),\end{aligned}$$

where, here and throughout the rest of this proof, the summation over k is from 1 to $\frac{c_0}{c}\log|V_n|$. We note that

$$\begin{aligned}&\mathbf{P}_x[cT_k < \tau(y) \le cT_{k+1} | \tau(y) > cT_k] \\ &\quad = \sum_w \mathbf{P}_x[cT_k < \tau(y) \le cT_{k+1} | \tau(y) > cT_k, X(cT_k) = w] \\ &\qquad \times \mathbf{P}_x[X(cT_k) = w | \tau(y) > cT_k] \\ &\quad = \sum_w \mathbf{P}_w[\tau(y) \le cT_1]\mathbf{P}_x[X(cT_k) = w | \tau(y) > cT_k].\end{aligned}$$

As the previous lemma is applicable for such choices of k and using $\mathbf{P}_w[\tau(y) \le cT_1] \le O(g(y, w) + c\Upsilon_n)$, we can rewrite the expression above as

$$\mathbf{P}_\pi[\tau(y) \le cT_1] + O\bigg(\sum_{w \ne y} g(y, w)\big(e^{-\gamma ck} + ck\Upsilon_n + g(y, w)\big)\pi(w)\bigg).$$

Performing the summation over w, we see that the latter term is of order

$$O(\Upsilon_n e^{-\gamma ck} + ck\Upsilon_n^2 + \mathcal{S}_n|V_n|^{-1}). \tag{3.4}$$

Recall from part (2) of Assumption 1.2 that $\mathcal{S}_n = o(T_{\text{mix}}/\log|V_n|)$, hence $(\log|V_n|)\mathcal{S}_n|V_n|^{-1} = o(\Upsilon_n)$. Consequently, summing (3.4) over k from 1 to $\frac{c_0}{c} \times \log|V_n|$ gives an error of

$$O(\Upsilon_n e^{-\gamma c} + \Delta_n^2).$$

By part (1) of Assumption 1.2 it is clear that $\Delta_n^2 = o(\Upsilon_n)$, hence the former is of order $O(\Upsilon_n e^{-\gamma c})$. This leaves

$$\sum_k \mathbf{P}_x[cT_k < \tau(y) \le cT_{k+1} | \tau(y) > cT_k] \mathbf{P}_x[\tau(y) \le cT_k]$$
$$= O\biggl(\sum_k \sum_z \mathbf{P}_z[\tau(y) \le cT_1]\pi(z)\bigl(g(x,y) + ck\Upsilon_n\bigr)\biggr).$$

Here, we used the previous lemma to get the crude estimate $\mathbf{P}_x[X(cT_k) = z | \tau(y) > cT_k] \le C\pi(z)$ for some $C > 0$. Summing everything up gives us an error of order $O(\Delta_n g(x,y) + \Delta_n^2)$. We also have another contribution of $O(\Delta_n g(x,z) + \Delta_n^2)$ coming from the corresponding estimate of $\mathbf{P}_x[\tau(z) \le \Gamma_n]$. Therefore our total error is $O(\Upsilon_n e^{-\gamma c} + \Delta_n[g(x,y) + g(x,z)])$, which proves the lemma. $\square$

4. The Radon–Nikodym derivative.

Recall

$$\widetilde{\pi}(z; x, y) = \mathbf{P}[X_2(\tau_1(x,y)) = z | H(x,y)].$$

The purpose of this section is to prove the following estimate of the Radon–Nikodym derivative of $\widetilde{\pi}(z; x, y)$ with respect to $\pi(z)$. Recall again $f_c(x,y) = \mathbf{P}_x[\tau(y) \le cT_{\mathrm{mix}}]$ and $\overline{f}_c = \sum_y f_c(x,y)\pi(y)$ (we are omitting the dependence on n).

THEOREM 4.1. *There exists a constant $\gamma > 0$ so that for all $c \ge 2$ and $x \ne y$, we have*

$$\frac{\widetilde{\pi}(z; x, y)}{\pi(z)} = 1 + \bigl(1 + O(\Delta_n)\bigr)\bigl(2\overline{f}_c - f_c(x,z) - f_c(y,z)\bigr)$$
$$+ O(e^{-\gamma c}\Upsilon_n) + O\bigl([g(x,z) + g(y,z) + \Delta_n][g(x,y) + \Delta_n]\bigr).$$

In particular,

$$\frac{\widetilde{\pi}(z; x, y)}{\pi(z)} = 1 + O\bigl(g(x,y) + g(y,z) + g(x,z)\bigr). \tag{4.1}$$

The setup for Theorem 4.1 is illustrated in Figure 1. Let $Y_2 = X_2(\tau_1(x,y))$. The idea of the proof is to observe that

$$\widetilde{\pi}(z; x, y) = \mathbf{P}[Y_2 = z | H(x,y)] = \frac{\mathbf{P}[H(x,y) | Y_2 = z]\pi(z)}{\mathbf{P}[H(x,y)]},$$

where we used $\mathbf{P}[Y_2 = z] = \pi(z)$ as X_1, X_2 are independent and the initial distribution of X_2 is stationary, then estimate the effect of conditioning on $\{Y_2 = z\}$ on the probability of $H(x,y)$. We will divide the proof into three lemmas. The first step in the proof is to express $\widetilde{\pi}(\cdot; x, y)/\pi$ in terms of the event

$$A(x,y) = \{\tau_2(x,y) > \tau_1(x,y) - \Gamma_n, G_1(x,y)\} \setminus H(x,y)$$
$$= \{\tau_1(x,y) \ge \tau_2(x,y) > \tau_1(x,y) - \Gamma_n, G_1(x,y)\}.$$

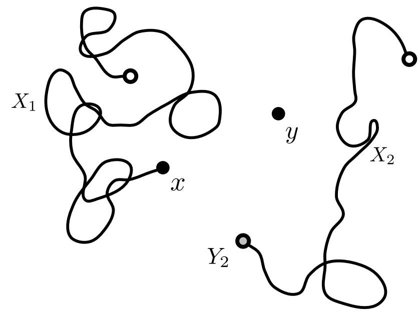

FIG. 1. *In Theorem 4.1, we give a precise estimate of the Radon–Nikodym derivative of the law of $Y_2 = X_2(\tau_1(x, y))$ with respect to the uniform measure on V_n conditional on the event that $H(x, y) = \{\tau_1(x) < \tau_1(y) \wedge \tau_2(x, y)\}$, that is, that the first point in $\{x, y\}$ hit by X_1, X_2 is x by X_1. The open circles indicate the starting points of X_1, X_2 and the shaded circle is Y_2.*

The event $A(x, y)$ is illustrated in Figure 2. Note that it is a slight abuse of notation to insert $G_1(x, y)$ into the braces defining $A(x, y)$ since $G_1(x, y)$ is itself an event. We will do this a number of times in the following lemma in order to lighten the notation.

LEMMA 4.2. *Uniformly in x, y, z, n,*

$$\frac{\widetilde{\pi}(z; x, y)}{\pi(z)} = 1 + \frac{\mathbf{P}[A(x, y)] - \mathbf{P}[A(x, y)|Y_2 = z]}{\mathbf{P}[H(x, y)]} + O(|V_n|^{-100}). \tag{4.2}$$

PROOF. Letting $R(x, y) = \{\tau_2(x, y) > \tau_1(x, y) - \Gamma_n, G_1(x, y)\}$, observe

$$\mathbf{P}[H(x, y)|Y_2 = z] = \mathbf{P}[R(x, y)|Y_2 = z] - \mathbf{P}[A(x, y)|Y_2 = z].$$

We will now manipulate the first term on the right-hand side. Let $\widetilde{R}(x, y) = G_1(x, y) \setminus R(x, y)$. We have

$$\mathbf{P}[R(x, y)|Y_2 = z] = \mathbf{P}[G_1(x, y)|Y_2 = z] - \mathbf{P}[\widetilde{R}(x, y)|Y_2 = z] \tag{4.3}$$

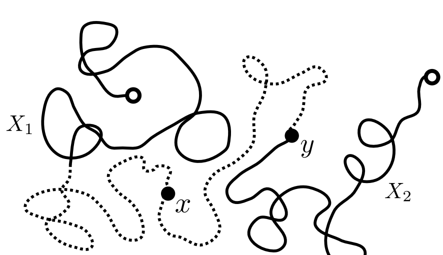

FIG. 2. *Illustration of the event $A(x, y)$. The solid lines are used to indicates the parts of X_1, X_2 up to the time $\tau_2(x, y)$, while the dashed line is used for the part of X_1 after $\tau_2(x, y)$. We have not indicated the part of X_2 after $\tau_2(x, y)$. Note that we have X_2 hitting y first, but $A(x, y)$ allows for X_2 to hit x first as well. On the other hand, $A(x, y)$ requires that X_1 does in fact hit x before y.*

and, since $Y_2 \sim \pi$, Bayes' rule implies

$$\mathbf{P}[\widetilde{R}(x,y)|Y_2=z] = \frac{1}{\pi(z)}\mathbf{P}[Y_2=z|\widetilde{R}(x,y)]\mathbf{P}[\widetilde{R}(x,y)].$$

Since the conditional probability on the right-hand side involves conditioning on the behavior of X_2 before $\tau_1(x,y)-\Gamma_n$; mixing considerations imply that this is equal to

$$[1+O(|V_n|^{-\overline{\gamma}c_0})]\mathbf{P}[\widetilde{R}(x,y)] = \mathbf{P}[\widetilde{R}(x,y)] + O(|V_n|^{-\overline{\gamma}c_0}). \tag{4.4}$$

As $\mathbf{P}[G_1(x,y)|Y_2=z] = \mathbf{P}[G_1(x,y)]$, combining (4.3) with (4.4) we thus have

$$\begin{aligned}\mathbf{P}[H(x,y)|Y_2=z] &= \mathbf{P}[R(x,y)] - \mathbf{P}[A(x,y)|Y_2=z] + O(|V_n|^{-\overline{\gamma}c_0})\\ &= \mathbf{P}[H(x,y)] + \mathbf{P}[A(x,y)] - \mathbf{P}[A(x,y)|Y_2=z]\\ &\quad + O(|V_n|^{-\overline{\gamma}c_0}).\end{aligned}$$

Assume that $\overline{\gamma}c_0 > 100$. Putting everything together, we see that

$$\frac{\widetilde{\pi}(z;x,y)}{\pi(z)} = \frac{\mathbf{P}[H(x,y)] + \mathbf{P}[A(x,y)] - \mathbf{P}[A(x,y)|Y_2=z]}{\mathbf{P}[H(x,y)]} + O(|V_n|^{-100}),$$

uniformly in x, y, z, n. $\square$

Note that if $V_n = \mathbf{Z}_n^d$ for $d \geq 3$, then $\mathbf{P}[G_1(x,y), G_{12}(x,y)] = \mathbf{P}[H(x,y)]$ does not change when x is swapped with y nor when 1 is swapped with 2 and hence is equal to $\frac{1}{4}$ up to negligible error (it is not exactly $\frac{1}{4}$ since it could be that X_1 hits x at the same time X_2 hits either x or y, though this is a rare event). This holds more generally if for every $x, y \in V_n$ distinct there exists an automorphism φ of G_n such that $\varphi(x) = y$ and $\varphi(y) = x$. The weaker hypothesis of vertex transitivity implies that we can always find an automorphism φ of G_n such that $\varphi(x) = y$ but not necessarily so that $\varphi(y) = x$ as well. Nevertheless, it is still true in this case that $\mathbf{P}[H(x,y)] \approx \frac{1}{4}$.

LEMMA 4.3. *If $x \neq y$, we have that*

$$\mathbf{P}[H(x,y)] = \frac{1}{4} + o\left(\frac{\Upsilon_n}{\log|V_n|}\right) + O\left(\frac{1}{|V_n|}\right).$$

PROOF. Let $\widetilde{A}(x,y) = \{\tau_1(x,y) \geq \tau_2(x,y) > \tau_1(x,y) - \Gamma_n\}$ and $\mu(z;x,y) = \mathbf{P}[Y_2 = z|\tau_1(x,y) \leq \tau_2(x,y)]$. Using exactly the same proof as the previous lemma, we have

$$\frac{\mu(z;x,y)}{\pi(z)} = 1 + O\big(\mathbf{P}[\widetilde{A}(x,y)] + \mathbf{P}[\widetilde{A}(x,y)|Y_2=z] + |V_n|^{-100}\big).$$

Using a time-reversal in the first step and a union bound in the second, we have that

$$\mathbf{P}[\widetilde{A}(x,y)|Y_2=z] \le \mathbf{P}_z[\tau_2(x,y) \le \Gamma_n] = O\big(\Delta_n + g(x,z) + g(y,z)\big)$$

(and similarly for $\mathbf{P}[\widetilde{A}(x,y)]$). Consequently,

$$\begin{aligned}\sum_z g(x,z)\mu(z;x,y) &= \Upsilon_n + \frac{1}{|V_n|}\sum_z g(x,z)O\big(\Delta_n + g(x,z) + g(y,z)\big)\\ &= \Upsilon_n + \frac{1}{|V_n|}\sum_z O\big(g(x,z)\Delta_n + g^2(x,z) + g^2(y,z)\big)\\ &= \Upsilon_n + \frac{1}{|V_n|}O(1+\mathcal{S}_n + \Delta_n T_{\mathrm{mix}}).\end{aligned}$$

By parts (1) and (2) of Assumption 1.2, we have that $\mathcal{S}_n + \Delta_n T_{\mathrm{mix}} = o(T_{\mathrm{mix}}/(\log|V_n|))$. Consequently, the above is equal to

$$\Upsilon_n + o\left(\frac{\Upsilon_n}{\log|V_n|}\right). \tag{4.5}$$

Let $p_x = \mathbf{P}[G_1(x,y), G_{12}(x,y)]$ and $p_y = \mathbf{P}[G_1(y,x), G_{12}(x,y)]$. Note that

$$p_x + p_y = \mathbf{P}[G_{12}(x,y)] = \frac{1}{2} + O\left(\frac{1}{|V_n|}\right) \tag{4.6}$$

since $\mathbf{P}[\tau_1(z) = \tau_2(w)] \le \mathbf{P}[X_2(\tau_1(z)) = w] = |V_n|^{-1}$ for any $z, w \in V_n$. Define stopping times as follows. Let

$$\tau_1 = \min\{t \ge 0 : X_1(t) \in \{x,y\} \text{ or } X_2(t) \in \{x,y\}\} = \tau(x,y).$$

For $j \ge 1$, inductively set

$$\tau_{j+1} = \min\{t \ge \tau_j + T_{\mathrm{mix}} + 1 : X_1(t) \in \{x,y\} \text{ or } X_2(t) \in \{x,y\}\}.$$

Let $\mathcal{T}_{j,z} = \sum_{t=\tau_j}^{\tau_j + T_{\mathrm{mix}}} \mathbf{1}_{\{X_1(t)=z\}}$, and, for $E \subseteq V_n$, set $A_{ij}(E) = \{X_i(\tau_j) \in E\}$. Note that the average amount of time spent at x by X_1 through time $\tau_k + T_{\mathrm{mix}}$ is given by the expression

$$\frac{1}{\tau_k + T_{\mathrm{mix}}}\sum_{j=1}^k \big(\mathbf{1}_{A_{1j}(x)}\mathbf{1}_{A_{2j}^c(x,y)}\mathcal{T}_{j,x} + \mathbf{1}_{A_{1j}(y)}\mathbf{1}_{A_{2j}^c(x,y)}\mathcal{T}_{j,x} + \mathbf{1}_{A_{2j}(x,y)}\mathcal{T}_{j,x}\big).$$

It is not difficult to see that the above quantity converges to $\pi(x)$ as $k \to \infty$. We can also define a similar quantity but replacing $\mathcal{T}_{j,x}$ with $\mathcal{T}_{j,y}$; this will converge to $\pi(y)$ as $k \to \infty$. Taking the ratio of these two quantities, we arrive at

$$1 = \lim_{k\to\infty} \frac{(1/k)\sum_{j=1}^k (\mathbf{1}_{A_{1j}(x)}\mathbf{1}_{A_{2j}^c(x,y)}\mathcal{T}_{j,x} + \mathbf{1}_{A_{1j}(y)}\mathbf{1}_{A_{2j}^c(x,y)}\mathcal{T}_{j,x} + \mathbf{1}_{A_{2j}(x,y)}\mathcal{T}_{j,x})}{(1/k)\sum_{j=1}^k (\mathbf{1}_{A_{1j}(x)}\mathbf{1}_{A_{2j}^c(x,y)}\mathcal{T}_{j,y} + \mathbf{1}_{A_{1j}(y)}\mathbf{1}_{A_{2j}^c(x,y)}\mathcal{T}_{j,y} + \mathbf{1}_{A_{2j}(x,y)}\mathcal{T}_{j,y})}$$

since $\pi(x) = \pi(y)$. It is not difficult to see that, almost surely,

$$\lim_{k\to\infty} \frac{1}{k} \sum_{j=1}^{k} \mathbf{1}_{A_{1j}(x)} \mathbf{1}_{A_{2j}^c(x,y)} \mathcal{T}_{j,x} = p_x g(x,x),$$

$$\lim_{k\to\infty} \frac{1}{k} \sum_{j=1}^{k} \mathbf{1}_{A_{1j}(y)} \mathbf{1}_{A_{2j}^c(x,y)} \mathcal{T}_{j,x} = p_y g(y,x),$$

$$\lim_{k\to\infty} \frac{1}{k} \sum_{j=1}^{k} \mathbf{1}_{A_{2j}(x,y)} \mathcal{T}_{j,x} = q_{xy} \sum_z g(z,x)\mu(z;x,y),$$

where $q_{xy} = 1 - p_x - p_y$. Analogous formulae hold for the terms in the denominator. Combining this with (4.5), we thus have

$$\begin{aligned} 1 &= \frac{p_x g(x,x) + p_y g(y,x) + q_{xy} \sum_z g(z,x)\mu(z;x,y)}{p_y g(y,y) + p_x g(x,y) + q_{xy} \sum_z g(z,y)\mu(z;x,y)} \\ &= \frac{p_x g(x,x) + p_y g(y,x) + q_{xy} \Upsilon_n}{p_y g(y,y) + p_x g(x,y) + q_{xy} \Upsilon_n} + o\left(\frac{\Upsilon_n}{\log|V_n|}\right). \end{aligned}$$

Rearranging and using that $g(x,y) = g(y,x)$ and $g(x,x) = g(y,y)$, this implies that

$$p_x = p_y + o\left(\frac{\Upsilon_n}{\log|V_n|}\right).$$

Combining this with (4.6) proves the lemma. $\square$

In order to complete the proof of Theorem 4.1 we need to estimate $\mathbf{P}[A(x,y)| Y_2 = z]$, which is the purpose of the following lemma. Though the proof will be computationally intensive, the basic idea is fairly simple. The main goal is to eliminate the conditioning on $Y_2 = z$. The first step is to perform a time reversal, which converts the terminal condition to an initial condition at the cost of making the event whose probability we are to compute a bit more complicated. The latter is easily mitigated, however, since the event can be greatly simplified at the cost of negligible error.

LEMMA 4.4. *There exists $\gamma > 0$ so that for all $c \geq 2$ we have*

$$\begin{aligned} \mathbf{P}[A(x,y)|Y_2 = z] &= \left(\tfrac{1}{4} + O(\Delta_n)\right)[f_c(x,z) + f_c(y,z)] + E_c(x,y) \\ &\quad + O\left(e^{-\gamma c}\Upsilon_n + [g(x,z) + g(y,z) + \Delta_n][g(x,y) + \Delta_n]\right), \end{aligned}$$

where $E_c(x,y)$ is some constant which does not depend on z.

Note that the lemma implies

$$\mathbf{P}[A(x,y)] - \mathbf{P}[A(x,y)|Y_2 = z] = O\left(g(x,y) + g(y,z) + g(x,z) + e^{-\gamma c}\Upsilon_n\right).$$

PROOF OF LEMMA 4.4. Let

$$B(x,y)=\{X_2(t)\notin\{x,y\}\text{ for all }t\in(\Gamma_n,\tau_1(x,y)],G_1(x,y)\},$$

and let $\mathbf{P}_{\pi,z}$ be the law of (X_1,X_2) where $X_1(0)\sim\pi$ and $X_2(0)=z$. We compute

$$\begin{aligned}&\mathbf{P}[A(x,y)|Y_2=z]\\&\quad=\frac{1}{\pi(z)}\mathbf{P}[A(x,y),Y_2=z]\\&\quad=\sum_w\mathbf{P}_{\pi,w}[\tau_1(x,y)\ge\tau_2(x,y)>\tau_1(x,y)-\Gamma_n,G_1(x,y),Y_2=z].\end{aligned}$$

By reversing the time of X_2 (but not X_1), we see that this is equal to

$$\begin{aligned}&\sum_w\mathbf{P}_{\pi,z}[\tau_2(x,y)\le\Gamma_n\wedge\tau_1(x,y),B(x,y),Y_2=w]\\&\qquad=\mathbf{P}_{\pi,z}[\tau_2(x,y)\le\Gamma_n\wedge\tau_1(x,y),B(x,y)].\end{aligned}\tag{4.7}$$

We will now work toward approximating this event with a simpler event. We begin by eliminating the "minimum" operation using the observation that it is unlikely for both X_1,X_2 to hit $\{x,y\}$ quickly. Indeed, as

$$\mathbf{P}_{\pi,z}[\tau_1(x,y)\le\Gamma_n,\tau_2(x,y)\le\Gamma_n]=O([g(x,z)+g(y,z)+\Delta_n]\Delta_n),$$

we see by setting $\widetilde{B}(x,y)=B(x,y)\cap\{\tau_1(x,y)>\Gamma_n\}$ that (4.7) is equal to

$$\mathbf{P}_{\pi,z}[\tau_2(x,y)\le\Gamma_n,\widetilde{B}(x,y)]+O([g(x,z)+g(y,z)+\Delta_n]\Delta_n).$$

We would now like to eliminate the dependence of the probability on z, the starting point of X_2. We accomplish this by considering two possible cases. Either X_2 hits x or y within some multiple of the mixing time or it does not. Conditional on the latter, the walk will have mixed, so the relevant probability does not depend on z. We implement this strategy as follows:

$$\begin{aligned}&\mathbf{P}_{\pi,z}[\tau_2(x,y)\le\Gamma_n,\widetilde{B}(x,y)]\\&\quad=\mathbf{P}_{\pi,z}[\tau_2(x,y)<cT_{\text{mix}},\widetilde{B}(x,y)]\\&\qquad+\mathbf{P}_{\pi,z}[\tau_2(x,y)\le\Gamma_n,\widetilde{B}(x,y)|\tau_2(x,y)\ge cT_{\text{mix}}]\\&\qquad\quad\times(1-\mathbf{P}_z[\tau_2(x,y)<cT_{\text{mix}}]).\end{aligned}$$

Using the same proof as Lemma 3.2, except in the case that the random walk is conditioned not to hit two points rather than just one, implies $\mu(w;x,y,z)=\mathbf{P}_z[X_2(cT_{\text{mix}})=w|\tau_2(x,y)\ge cT_{\text{mix}}]\le C\pi(w)$ for some constant $C>0$. Consequently,

$$\begin{aligned}&\mathbf{P}_{\pi,z}[\tau_2(x,y)\le\Gamma_n,\widetilde{B}(x,y)|\tau_2(x,y)\ge cT_{\text{mix}}]\mathbf{P}_z[\tau_2(x,y)<cT_{\text{mix}}]\\&\quad\le C\mathbf{P}_\pi[\tau_2(x,y)\le\Gamma_n]\mathbf{P}_z[\tau_2(x,y)<cT_{\text{mix}}]\\&\quad=O([g(x,z)+g(y,z)+c\Upsilon_n]\Delta_n).\end{aligned}$$

We are left with two terms to estimate

$$\mathbf{P}_{\pi,z}[\tau_2(x,y) < cT_{\mathrm{mix}}, \widetilde{B}(x,y)], \tag{4.8}$$

$$\mathbf{P}_{\pi,z}[\tau_2(x,y) \le \Gamma_n, \widetilde{B}(x,y) | \tau_2(x,y) \ge cT_{\mathrm{mix}}]. \tag{4.9}$$

We will first deal with (4.8) which, using the independence of $\tau_1(x,y)$ and $\tau_2(x,y)$, we can rewrite as

$$\begin{aligned}&\mathbf{P}_{\pi,z}[B(x,y)|\tau_1(x,y) > \Gamma_n, \tau_2(x,y) < cT_{\mathrm{mix}}]\\ &\quad \times \mathbf{P}_z[\tau_2(x,y) < cT_{\mathrm{mix}}]\mathbf{P}[\tau_1(x,y) > \Gamma_n].\end{aligned}$$

Since $B(x,y)$ depends on $X_2(t)$ only for $t \ge \Gamma_n$, from mixing considerations it is easy to see that

$$\begin{aligned}&\mathbf{P}_{\pi,z}[B(x,y)|\tau_1(x,y) > \Gamma_n, \tau_2(x,y) < cT_{\mathrm{mix}}]\\ &\quad = \mathbf{P}[B(x,y)|\tau_1(x,y) > \Gamma_n] + O(|V_n|^{-100}).\end{aligned}$$

Consequently, (4.8) is equal to

$$\mathbf{P}_z[\tau_2(x,y) < cT_{\mathrm{mix}}]\mathbf{P}[\widetilde{B}(x,y)] + O(|V_n|^{-100}).$$

Note that

$$\widetilde{B}(x,y) = \big(H(x,y) \cap \{\tau(x,y) > \Gamma_n\}\big) \cup \big(\widetilde{B}(x,y) \cap \{\tau(x,y) \le \Gamma_n\}\big).$$

Using $\mathbf{P}[\tau(x,y) \le \Gamma_n] = O(\Delta_n)$, the previous lemma thus implies

$$\mathbf{P}[\widetilde{B}(x,y)] = \mathbf{P}[H(x,y)] + O(\Delta_n) = \tfrac{1}{4} + O(\Delta_n).$$

Observe

$$\begin{aligned}&\mathbf{P}_z[\tau_2(x) < cT_{\mathrm{mix}}, \tau_2(y) < cT_{\mathrm{mix}}]\\ &\quad = \mathbf{P}_z[\tau_2(x) < \tau_2(y) < cT_{\mathrm{mix}}] + \mathbf{P}_z[\tau_2(y) < \tau_2(x) < cT_{\mathrm{mix}}]\\ &\quad = O\big([g(x,z) + g(y,z) + c\Upsilon_n][g(x,y) + c\Upsilon_n]\big).\end{aligned}$$

Consequently,

$$\begin{aligned}&\mathbf{P}_z[\tau_2(x,y) < cT_{\mathrm{mix}}]\\ &\quad = f_c(x,z) + f_c(y,z) - \mathbf{P}_z[\tau_2(x) < cT_{\mathrm{mix}}, \tau_2(y) < cT_{\mathrm{mix}}]\\ &\quad = f_c(x,z) + f_c(y,z) + O\big([g(x,z) + g(y,z) + c\Upsilon_n][g(x,y) + c\Upsilon_n]\big).\end{aligned}$$

Arguing as in the proof of Lemma 3.3, we can estimate (4.9) as follows:

$$\begin{aligned}&\mathbf{P}_{\pi,z}[\tau_2(x,y) \le \Gamma_n, \widetilde{B}(x,y)|\tau_2(x,y) \ge cT_{\mathrm{mix}}]\\ &\quad = \mathbf{P}[\tau_2(x,y) \le \Gamma_n, \widetilde{B}(x,y)|\tau_2(x,y) \ge cT_{\mathrm{mix}}]\\ &\qquad + O\big(e^{-\gamma c}\Upsilon_n + \big(g(x,z) + g(y,z)\big)\Delta_n\big).\end{aligned}$$

Taking $E_c(x, y) = \mathbf{P}[\tau_2(x, y) \leq \Gamma_n, \widetilde{B}(x, y) | \tau_2(x, y) \geq cT_{\mathrm{mix}}]$ and noting that $c\Upsilon_n = O(\Delta_n)$ finishes the proof of the lemma. $\square$

By combining the three lemmas, we can now complete the proof of Theorem 4.1.

PROOF OF THEOREM 4.1. Recalling that $\mathbf{P}[Y_2 = z] = \pi(z)$, observe

$$\mathbf{P}[A(x, y)] = \sum_z \mathbf{P}[A(x, y) | Y_2 = z] \mathbf{P}[Y_2 = z] = \sum_z \mathbf{P}[A(x, y) | Y_2 = z] \pi(z).$$

Hence by Lemma 4.4, we have

$$\mathbf{P}[A(x, y)] = \left(\tfrac{1}{4} + O(\Delta_n)\right)(2\overline{f}_c) + E_c(x, y) + O\left(e^{-\gamma c}\Upsilon_n + \Delta_n[g(x, y) + \Delta_n]\right).$$

Here, we used that

$$\sum_z \left(g(x, z) + g(y, z)\right)\pi(z) = O(\Upsilon_n) = O(\Delta_n).$$

In particular,

$$\begin{aligned} &\mathbf{P}[A(x, y)] - \mathbf{P}[A(x, y) | Y_2 = z] \\ &\quad = \left(\tfrac{1}{4} + O(\Delta_n)\right)[2\overline{f}_c - f_c(x, z) - f_c(y, z)] \\ &\qquad + O\left(e^{-\gamma c}\Upsilon_n + [g(x, z) + g(y, z) + \Delta_n][g(x, y) + \Delta_n]\right). \end{aligned}$$

Inserting this expression along with the estimate of $\mathbf{P}[H(x, y)]$ from Lemma 4.3 into into the equation for $\widetilde{\pi}(z; x, y)/\pi(z)$ from Lemma 4.2 gives the theorem. $\square$

5. The variance. We will complete the proof of Theorem 1.3 in this section. The general theme is to eliminate asymmetry wherever possible. We first apply this idea by considering

$$\mathcal{B} = \sum_x \mathbf{1}_{G_{12}(x)} - \sum_x \mathbf{1}_{G_{21}(x)}$$

in place of $|\mathcal{A}_1|$. In addition to being symmetric in X_1, X_2, note that $\mathcal{B}$ also differs from $|\mathcal{A}_1|$ in that we have eliminated those sites whose mark is determined by the flip of a fair coin. These, however, do not make a significant contribution to the variance since it is a rare event that both walks hit a particular point for the first time simultaneously. In particular, we will show in Lemma 5.1 that $\mathrm{Var}(\mathcal{B}) \approx 4\,\mathrm{Var}(|\mathcal{A}_1|)$, up to negligible error. Consequently, to prove Theorem 1.3 it suffices to show

$$\mathrm{Var}(\mathcal{B}) = \sum_{x,y} \left(f_c(x, y) - \overline{f}_c\right)^2 + O(e^{-\gamma c}(T_{\mathrm{mix}})^2).$$

It is convenient to work with $\mathcal{B}$ as the expansion of its variance takes on the following form:

$$\operatorname{Var}(\mathcal{B}) = 2\sum_{x,y}\big(\mathbf{P}[G_{12}(x), G_{12}(y)] - \mathbf{P}[G_{12}(x), G_{21}(y)]\big)$$

$$= 4\sum_{x\neq y\neq z}\big(\mathbf{P}_{x,z}[G_{12}(y)] - \mathbf{P}_{x,z}[G_{21}(y)]\big)\widetilde{\pi}(z; x, y)\mathbf{P}[H(x, y)] \tag{5.1}$$

$$+ 2\sum_{x}\mathbf{P}[G_{12}(x)]. \tag{5.2}$$

The reason the summation in (5.1) is over $x \neq y \neq z$ is $\mathbf{P}[H(x, x)] = 0$ and $\widetilde{\pi}(z; x, y) = 0$ if $z \in \{x, y\}$; the summation in (5.2) contains the diagonal terms. We will now focus on (5.1) and handle (5.2) at the end of the section. Applying Lemma 4.3, we can rewrite (5.1) as

$$\sum_{x\neq y\neq z}\big(\mathbf{P}_{x,z}[G_{12}(y)] - \mathbf{P}_{x,z}[G_{21}(y)]\big)\widetilde{\pi}(z; x, y) \tag{5.3}$$

$$+ \sum_{x\neq y\neq z}|\mathbf{P}_{x,z}[G_{12}(y)] - \mathbf{P}_{x,z}[G_{21}(y)]| \times\left[o\left(\frac{T_{\mathrm{mix}}}{|V_n|^2\log|V_n|}\right) + O\left(\frac{1}{|V_n|^2}\right)\right]. \tag{5.4}$$

We will show at the end of this section that (5.4) is negligible. Note that

$$\begin{aligned}\mathbf{P}_{x,z}[G_{ij}(y)] &= \mathbf{P}_{x,z}[\tau_i(y) < \tau_j(y) \leq \Gamma_n] \\ &\quad + \mathbf{P}_{x,z}[\tau_i(y) \leq \Gamma_n, \tau_j(y) > \Gamma_n] + \mathbf{P}_{x,z}[\Gamma_n < \tau_i(y) < \tau_j(y)] \\ &\equiv A + B + C.\end{aligned} \tag{5.5}$$

In Section 5.2, we break the sum in (5.3) into three different cases based on the time decomposition in (5.5) and bound each in a given lemma. It will turn out that the contributions to the variance coming from the terms corresponding to A and C are negligible (Lemmas 5.3 and 5.4). The reason for the former is that it is unlikely for both X_1 and X_2 to hit y quickly and the latter follows as, conditional on having not hit y by time Γ_n, both walks have long forgotten their initial conditions and are well mixed. This leaves B, which, along with the diagonal, dominates the variance. Its asymptotics will be computed (Lemma 5.2) by reducing the estimate to a computation involving $\widetilde{\pi}(z; x, y)$, whose Radon–Nikodym derivative with respect to the uniform measure has already been estimated precisely in Theorem 4.1.

5.1. *Symmetrization.*

LEMMA 5.1. *We have*

$$\operatorname{Var}(\mathcal{B}) = 4\operatorname{Var}(|\mathcal{A}_1|) + O\big(\sqrt{T_{\mathrm{mix}}\operatorname{Var}(|\mathcal{A}_1|)} + T_{\mathrm{mix}}\big) \qquad \text{as } n \to \infty.$$

PROOF. Let $(\xi_n(x) : x \in V_n)$ be i.i.d. random variables independent of X_1, X_2 with $\mathbf{P}[\xi_n(x) = 1] = \mathbf{P}[\xi_n(x) = 2] = \frac{1}{2}$, and let $A(x,i) = \{\tau_1(x) = \tau_2(x), \xi_n(x) = i\}$. By definition,

$$\begin{aligned}\operatorname{Var}(\mathcal{B}) &= \operatorname{Var}\biggl(|\mathcal{A}_1| - (|V_n| - |\mathcal{A}_1|) - \sum_x (\mathbf{1}_{A(x,1)} - \mathbf{1}_{A(x,2)})\biggr) \\ &= \operatorname{Var}\biggl(2|\mathcal{A}_1| + \sum_x (\mathbf{1}_{A(x,2)} - \mathbf{1}_{A(x,1)})\biggr).\end{aligned}$$

Observe

$$\mathbf{E}\biggl(\sum_x \mathbf{1}_{A(x,1)}\biggr)^2 \le \sum_{x,y} \mathbf{P}[\tau_1(x) = \tau_2(x), \tau_1(y) = \tau_2(y)].$$

By the strong Markov property and independence of X_1, X_2, the above is bounded by twice

$$\begin{aligned}&\sum_{x,y} \mathbf{P}_{x,x}[\tau_1(y) = \tau_2(y)]\mathbf{P}[\tau_1(x) = \tau_2(x)] \\ &\qquad \le \sum_{x,y} \sum_t (\mathbf{P}_x[\tau(y) = t])^2 \mathbf{P}[X_2(\tau_1(x)) = x].\end{aligned}$$

Using that $\mathbf{P}_x[\tau(y) = t] \le \mathbf{P}_x[X(t) = y]$ and $X_2(\tau_1(x)) \sim \pi$ when $X(0) \sim \pi$, we have the further bound

$$\frac{1}{|V_n|} \sum_{x,y} \Biggl(\sum_{t=0}^{4T_{\text{mix}}} \mathbf{P}_x[X(t) = y] + \sum_{t > 4T_{\text{mix}}} (\mathbf{P}_x[\tau(y) = t])^2 \Biggr). \tag{5.6}$$

Summing the first term over x, y, t plainly yields $4T_{\text{mix}}$. For the second term, note there exists $C > 0$ so that for $t > 4T_{\text{mix}}$, we have

$$\mathbf{P}_x[\tau(y) = t] \le \mathbf{P}_x[X(t) = y] \le \frac{C}{|V_n|}$$

hence

$$\sum_{t > 4T_{\text{mix}}} (\mathbf{P}_x[\tau(y) = t])^2 \le \frac{C}{|V_n|} \sum_{t > 4T_{\text{mix}}} \mathbf{P}_x[\tau(y) = t] \le \frac{C}{|V_n|}.$$

Therefore the second term in the summation in (5.6) is $O(1)$. The lemma now follows from Cauchy–Schwarz. $\square$

5.2. *Time decomposition.* We begin by estimating the part of (5.3) corresponding to "B" from (5.5).

LEMMA 5.2. *We have*

$$\sum_{x\neq y\neq z} \big(\mathbf{P}_{x,z}[\tau_1(y)\le \Gamma_n, \tau_2(y) > \Gamma_n] - \mathbf{P}_{x,z}[\tau_2(y)\le \Gamma_n, \tau_1(y) > \Gamma_n]\big)\widetilde{\pi}(z; x, y)$$
$$= \big(1+O(\Delta_n)\big)\sum_{x\neq y}\big(f_c(x,y)-\overline{f}_c\big)^2 + O(e^{-\gamma c}(T_{\mathrm{mix}})^2).$$

PROOF. Note that

$$\mathbf{P}_{x,z}[\tau_1(y)\le \Gamma_n, \tau_2(y) > \Gamma_n] - \mathbf{P}_{x,z}[\tau_2(y)\le \Gamma_n, \tau_1(y) > \Gamma_n]$$
$$= \mathbf{P}_x[\tau_1(y)\le \Gamma_n] - \mathbf{P}_z[\tau_2(y)\le \Gamma_n].$$

Let $\delta_1(x,y,z) = O(e^{-\gamma c}\Upsilon_n + (g(x,z)+g(y,z)+\Delta_n)(g(x,y)+\Delta_n))$ be the error term from Theorem 4.1 and $\delta_2(x,y,z) = O(e^{-\gamma c}\Upsilon_n + \Delta_n(g(x,y)+g(x,z)))$ be the error term from Lemma 3.3. Then we can rewrite the summation in the statement of the lemma as

$$\sum_{x\neq y\neq z} \big(\mathbf{P}_x[\tau_1(y)\le \Gamma_n] - \mathbf{P}_z[\tau_2(y)\le \Gamma_n]\big)\widetilde{\pi}(z; x, y)$$
$$= \frac{1}{|V_n|}\sum_{x\neq y\neq z}\big(\mathbf{P}_x[\tau_1(y)\le \Gamma_n] - \mathbf{P}_z[\tau_2(y)\le \Gamma_n]\big)\big(1+\varepsilon(x,y,z)\big)$$
$$+ \frac{1}{|V_n|}\sum_{x\neq y\neq z}\big(|f_c(x,y)-f_c(y,z)| + \delta_2(x,y,z)\big)\delta_1(x,y,z)$$
$$\equiv B_1 + B_2,$$

where, by Theorem 4.1,

$$\varepsilon(x,y,z) = \big(1+O(\Delta_n)\big)\big(2\overline{f}_c - f_c(x,z) - f_c(y,z)\big).$$

Applying Assumption 1.2 repeatedly, it is tedious but not difficult to see that $B_2 = O(e^{-\gamma c}T_{\mathrm{mix}}^2)$. By Lemma 3.3,

$$B_1 = \big(1+O(\Delta_n)\big)\frac{1}{|V_n|}\sum_{x\neq y\neq z}\big(f_c(x,y)-f_c(y,z)+\delta_2(x,y,z)\big)$$
$$\times \big(2\overline{f}_c - f_c(x,z) - f_c(y,z)\big).$$

Multiplying through, using the symmetry of f in its arguments and canceling many terms, this becomes

$$\big(1+O(\Delta_n)\big)\sum_{x\neq y}\big(f_c(x,y)-\overline{f}_c\big)^2 + O(e^{-\gamma c}(T_{\mathrm{mix}})^2).$$
□

We will now show that the part of (5.3) coming from "A" of (5.5) is negligible. Roughly, the reason for this is that it is unlikely for both walks to hit y quickly, though in order to get a sufficiently good bound we will need to take advantage of some more cancellation. This will in turn require us to invoke (4.1), which is a rough estimate of the Radon–Nikodym derivative of $\widetilde{\pi}$ with respect to π.

LEMMA 5.3. *Uniformly in n, we have*

$$\sum_{x\neq y\neq z}\big(\mathbf{P}_{x,z}[\tau_1(y)<\tau_2(y)\le\Gamma_n]-\mathbf{P}_{x,z}[\tau_2(y)<\tau_1(y)\le\Gamma_n]\big)\widetilde{\pi}(z;x,y)$$
$$=o(T_{\text{mix}}^2).$$

PROOF. By (4.1), the summation in the statement of the lemma is equal to

$$\frac{1}{|V_n|}\sum_{x\neq y\neq z}\big(\mathbf{P}_{x,z}[\tau_1(y)<\tau_2(y)\le\Gamma_n]-\mathbf{P}_{x,z}[\tau_2(y)<\tau_1(y)\le\Gamma_n]\big)$$
$$\times\big(1+O\big(g(x,y)+g(y,z)+g(x,z)\big)\big).$$

By symmetry, we see that this is equal to

$$\frac{1}{|V_n|}\sum_{x\neq y\neq z}\big(\mathbf{P}_{x,z}[\tau_1(y)<\tau_2(y)\le\Gamma_n]-\mathbf{P}_{x,z}[\tau_2(y)<\tau_1(y)\le\Gamma_n]\big)$$
$$\times\big(O\big(g(x,y)+g(y,z)+g(x,z)\big)\big).$$

Using

$$\mathbf{P}_{x,z}[\tau_1(y)<\tau_2(y)\le\Gamma_n]\le\mathbf{P}_{x,z}[\tau_1(y)\le\Gamma_n,\tau_2(y)\le\Gamma_n]$$

and the independence of X_1, X_2, we have the further bound

$$\frac{1}{|V_n|}\sum_{x\neq y\neq z}O\big(\big[g(x,y)+\Delta_n\big)\big(g(y,z)+\Delta_n\big)\big)O\big(g(x,y)+g(y,z)+g(x,z)\big).$$

By the symmetry of g in its arguments, we can rewrite this as

$$\frac{1}{|V_n|}\sum_{x\neq y\neq z}O\big(g(x,y)g^2(y,z)+g^2(x,y)\Delta_n+g(x,y)\Delta_n^2$$
$$+g(x,y)g(x,z)\Delta_n+g(x,y)g(x,z)g(y,z)\big).$$

The terms in the summation are of order

$$\mathcal{S}_nT_{\text{mix}},\qquad \mathcal{S}_nT_{\text{mix}}\log|V_n|,\qquad \frac{T_{\text{mix}}^3(\log|V_n|)^2}{|V_n|},\qquad \frac{T_{\text{mix}}^3\log|V_n|}{|V_n|},\qquad \frac{T_{\text{mix}}^3}{|V_n|},$$

respectively. Assumption 1.2 implies that all of these are $o(T_{\text{mix}}^2)$, which gives the lemma. $\square$

We complete this subsection by proving that "C" from (5.5) is also negligible in comparison to the bound we seek to prove. The intuition for this is that by time Γ_n, both walks are very well mixed hence given that both have not hit y, the difference in the probability that one hits before the other is of smaller order than any negative power of $|V_n|$ (though we choose to write -100). The proof will be in a slightly different spirit than the previous lemmas.

LEMMA 5.4. *For any fixed x, z, we have*

$$\mathbf{P}_{x,z}[\Gamma_n < \tau_1(y) < \tau_2(y)] - \mathbf{P}_{x,z}[\Gamma_n < \tau_2(y) < \tau_1(y)] = O(|V_n|^{-100}).$$

PROOF. We may assume without loss of generality that $x, z \neq y$. The idea of the proof is to use a standard coupling argument to show that, conditional on $\{\tau_1(y) \wedge \tau_2(y) \geq \Gamma_n\}$, the laws of $X_1(\Gamma_n)$ and $X_2(\Gamma_n)$ have total variation distance $O(|V_n|^{-100})$ independent of x, z. To this end, we set $\mu(z; x, y) = \mathbf{P}_x[X(\Gamma_n) = z | \tau(y) \geq \Gamma_n]$. Let $Y(t)$ be the process given by $X(t)$ conditioned on the event $\{\tau(y) \geq \Gamma_n\}$. Then $Y(t)$ is Markov (though time-inhomogeneous) as

$$\begin{aligned}
&\mathbf{P}[Y(t) = z | Y(0) = z_0, \ldots, Y(t-1) = z_{t-1}] \\
&\quad = \mathbf{P}[X(t) = z | X(0) = z_0, \ldots, X(t-1) = z_{t-1}, \tau(y) \geq \Gamma_n] \\
&\quad = \frac{\mathbf{P}[X(t) = z, \tau(y) \geq \Gamma_n | X(0) = z_0, \ldots, X(t-1) = z_{t-1}]}{\mathbf{P}[\tau(y) \geq \Gamma_n | X(0) = z_0, \ldots, X(t-1) = z_{t-1}]} \\
&\quad = \frac{\mathbf{P}_{z_{t-1}}[X(1) = z, \tau(y) \geq \Gamma_n - (t-1)]}{\mathbf{P}_{z_{t-1}}[\tau(y) \geq \Gamma_n - (t-1)]}
\end{aligned}$$

depends only on z, z_{t-1}. Recall that $T_k = kT_{\text{mix}}$. For $t = cT_k$, note that

$$\begin{aligned}
\nu(z; t, x) \equiv \mathbf{P}_x[Y(t) = z] &= \frac{\mathbf{P}_x[X(t) = z, \tau(y) \geq \Gamma_n | \tau(y) \geq t]}{\mathbf{P}_x[\tau(y) \geq \Gamma_n | \tau(y) \geq t]} \\
&= \frac{\mathbf{P}_z[\tau(y) \geq \Gamma_n - t]\mathbf{P}_x[X(t) = z | \tau(y) \geq t]}{\mathbf{P}_x[\tau(y) \geq \Gamma_n | \tau(y) \geq t]}.
\end{aligned}$$

Combining part (3) of Assumption 1.2 with Lemma 3.2, we have that

$$\frac{\mathbf{P}_z[\tau(y) \geq \Gamma_n - t]}{\mathbf{P}_x[\tau(y) \geq \Gamma_n | \tau(y) \geq t]} = \Theta(1) \qquad \text{for } z \neq y.$$

Also, since $\sum_z g(y, z) = T_{\text{mix}}$ and $T_{\text{mix}} = o(|V_n|)$, it follows that for each $\varepsilon > 0$ fixed, with $A = \{z : g(y, z) \leq \varepsilon\}$ we have $|A|/|V_n| = 1 - o(1)$. Lemma 3.2 also implies that $\mathbf{P}_x[X(t) = z | \tau(y) \geq t] = \Theta(1)\pi(z)$ on A uniformly in n large provided that k is large enough and $\varepsilon > 0$ is sufficiently small. This implies that we can couple together the laws of $Y_u(cT_k), Y_v(cT_k)$ starting at u, v distinct so that with probability $\rho > 0$, we have $Y_u(cT_k) = Y_v(cT_k)$. If we iterate this procedure $c_1 = \frac{c_0}{\eta} \log|V_n|$ times, $\eta = \eta(c, k, \rho)$, we get that with probability $1 - O(|V_n|^{-c_1})$, we have $Y_u(\Gamma_n) = Y_v(\Gamma_n)$. Consequently, we may assume that c_0 is sufficiently large so that

$$\max_{u,v} \|\nu(\cdot;\Gamma_n,u) - \nu(\cdot,;\Gamma_n,v)\|_{\mathrm{TV}} = O(|V_n|^{-500}).$$

Let $\overline{\nu}$ be a measure so that $\max_u \|\nu(\cdot;\Gamma_n,u) - \overline{\nu}\|_{\mathrm{TV}} = O(|V_n|^{-500})$. Let $D = \{\tau_1(y) \wedge \tau_2(y) \geq \Gamma_n\}$. Then we have that

$$\begin{aligned}
&\mathbf{P}_{x,z}[\Gamma_n < \tau_1(y) < \tau_2(y)] - \mathbf{P}_{x,z}[\Gamma_n < \tau_2(y) < \tau_1(y)] \\
&\quad = \big(\mathbf{P}_{x,z}[G_{12}(y)|D] - \mathbf{P}_{x,z}[G_{21}(y)|D]\big)\mathbf{P}_{x,z}[D] \\
&\quad = \sum_{u,v}\big(\mathbf{P}_{u,v}[G_{12}(y)] - \mathbf{P}_{u,v}\mathbf{P}[G_{21}(y)]\big)\overline{\nu}(u;\Gamma_n,x)\overline{\nu}(v;\Gamma_n,z)\mathbf{P}_{x,z}[D] \\
&\qquad + O(|V_n|^{-200}) \\
&\quad = O(|V_n|^{-200}). \qquad \square
\end{aligned}$$

PROOF OF THEOREM 1.3. To finish the proof of Theorem 1.3, we need to estimate the diagonal (5.2) and take care of the term in (5.4). Observe that (5.2) is equal to

$$2\sum_x \mathbf{P}[G_{12}(x)] = |V_n| + O\Big(\sum_x \mathbf{P}[\tau_1(x) = \tau_2(x)]\Big). \tag{5.7}$$

We can estimate the sum on the right-hand side using

$$\begin{aligned}
\sum_x \mathbf{P}[\tau_1(x) = \tau_2(x)] &= \sum_x\sum_t \mathbf{P}[\tau_1(x) = t]\mathbf{P}[\tau_2(x) = t] \\
&\leq \sum_x\sum_t \mathbf{P}[\tau_1(x) = t]\mathbf{P}[X_2(t) = x] = 1.
\end{aligned}$$

On the other hand, note

$$\sum_x \big(f_c(x,x) - \overline{f}_c\big)^2 = \sum_x (1 - 2\overline{f}_c + \overline{f}_c^2). \tag{5.8}$$

By a union bound, we have $\overline{f}_c = O(c\Upsilon_n)$. Thus the diagonal term in (5.7) and (5.8) differ by $O(cT_{\mathrm{mix}}) = o(e^{-\gamma c}T_{\mathrm{mix}}^2)$ (recall from Assumption 1.2 that $T_{\mathrm{mix}} \to \infty$ as $n \to \infty$). This takes care of (5.2).

We now turn to (5.4). The previous lemma implies

$$\begin{aligned}
&\sum_{x\neq y\neq z} |\mathbf{P}_{x,z}[G_{12}(y)] - \mathbf{P}_{x,z}[G_{21}(y)]| \\
&\quad = \sum_{x\neq y\neq z} |\mathbf{P}_{x,z}[G_{12}(y), \tau_1(y) \leq \Gamma_n] - \mathbf{P}_{x,z}[G_{21}(y), \tau_2(y) \leq \Gamma_n]| \\
&\qquad + O(|V_n|^{-50}).
\end{aligned}$$

Observe $\{G_{ij}(y), \tau_i(y) \le \Gamma_n\} \subseteq \{\tau_i(y) \wedge \tau_j(y) \le \Gamma_n\} \cup \{\tau_i(y) \le \Gamma_n < \tau_j(y)\}$. Thus we can bound from above the previous expression by

$$\sum_{x \ne y \ne z} \bigl(2\mathbf{P}_{x,z}[\tau_1(y) \wedge \tau_2(y) \le \Gamma_n] + |\mathbf{P}_x[\tau_1(y) \le \Gamma_n] - \mathbf{P}_z[\tau_2(y) \le \Gamma_n]|\bigr)$$
$$\equiv E_1 + E_2.$$

The term corresponding to E_1 can be bounded in a similar manner as "A" in the proof of Lemma 5.3. Indeed, by the independence of X_1, X_2, we have that

$$\mathbf{P}_{x,z}[\tau_1(y) \wedge \tau_2(y) \le \Gamma_n] \le \bigl(g(x,y) + \Delta_n\bigr)\bigl(g(y,z) + \Delta_n\bigr),$$

which, when summed over x, y, z, is of order $O((\log|V_n|)^2|V_n|T_{\mathrm{mix}}^2)$. We can estimate E_2 using techniques similar to the proof of Lemma 5.2 since by Lemma 3.3,

$$|\mathbf{P}_x[\tau_1(y) \le \Gamma_n] - \mathbf{P}_z[\tau_2(y) \le \Gamma_n]| = O\bigl(g(x,y) + g(y,z) + \delta_2(x,y,z)\bigr),$$

where, as in the proof of Lemma 5.2, $\delta_2(x,y,z)$ corresponds to the error from Lemma 3.3. When summed over x, y, z, this is of order $O(|V_n|^2 T_{\mathrm{mix}})$. Therefore

$$(E_1 + E_2)\left(o\left(\frac{T_{\mathrm{mix}}}{|V_n|^2 \log|V_n|}\right) + O\left(\frac{1}{|V_n|^2}\right)\right) = o(T_{\mathrm{mix}}^2)$$

as desired. $\square$

6. Further questions. (1) The first step in proving a sequence of random variables (X_n) has a Gaussian limit after appropriate normalization is the determination of the asymptotic mean and variance. We remarked in the beginning that, in our case, the expected number of sites painted 1 is $|V_n|/2$, and Theorem 1.3 gives the limiting variance. Figure 3 shows Q–Q plots of the empirical distribution of the number of sites painted 1 in the final coloring against an appropriately fitted normal for three different base graphs. Based on these plots, we conjecture that

$$\frac{|\mathcal{A}_i| - \mathbf{E}|\mathcal{A}_i|}{\sqrt{\mathrm{Var}(|\mathcal{A}_i|)}}$$

has a normal limit for all graphs satisfying Assumption 1.2.

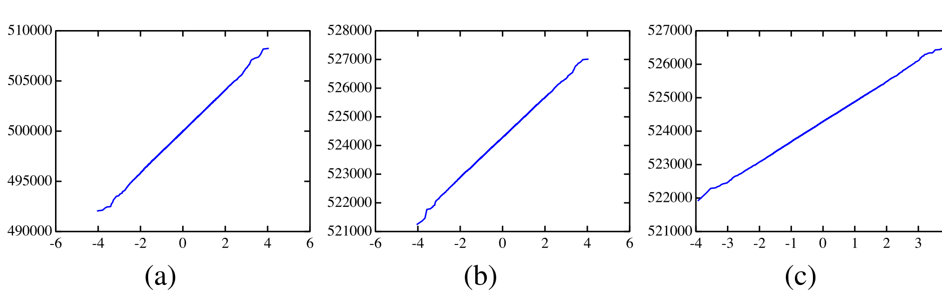

FIG. 3. *Q–Q plots based on* 20,000 *simulations of the number of sites visited by X_1 before X_2 against an appropriately fitted normal distribution, supporting the conjecture of asymptotic normality, where* (a) $\mathbf{Z}_{100}^3$; (b) $\mathbf{Z}_{32}^4$ *and* (c) $\mathbf{Z}_2^{20}$.

(2) Our derivation of the variance ignores the time aspect of the problem in the sense that it gives no indication of at what point in the process of coverage the variance is "created." Does it come in bursts or continuously? Does it come sooner than any multiple of the cover time or perhaps in $[\varepsilon T_{\mathrm{cov}}, T_{\mathrm{cov}}]$? More generally, when normalized appropriately, does the the process $t \mapsto \sum_x \mathbf{1}_{\{\tau_1(x)<\tau_2(x)\le t\}}$ have a scaling limit?

(3) We make repeated used of the symmetry afforded by the fact that we consider two random walks moving at the same speed on vertex transitive graph. It would be interesting to see if a similar result holds when the various degrees of symmetry are broken. Starting points for exploring this problem include considering continuous time walks moving at various speeds, multiple walks and graphs which are not vertex transitive.

(4) Theorem 1.1 only holds for tori of dimension $d \ge 3$ as the case $d = 2$ falls just outside of the scope of Theorem 1.3. It would be interesting to see a more refined analysis carried out to handle this case.

(5) That the variance computed in Theorem 1.1 for $d = 3, 4$ is significantly larger than in the i.i.d. case suggests that the clusters which have an unusually large number of sites painted a given color are either larger or more dense than in an i.i.d. marking. How large and frequent are such clusters? What is their geometric structure?

(6) Another interesting quantity is the size $\mathcal{B}$ of the boundary separating the sites painted 1 and 2, as studied in [9]. It is not difficult to see that there exists a constant $\beta_d > 0$ such that $\mathbf{E}|\mathcal{B}| \sim \beta_d n^d$ when $d \ge 3$ as $n \to \infty$. Indeed, this follows since the probability that $\{\tau_1(y) < \tau_2(y)\}$ for $y \sim x$ given $\{\tau_1(x) < \tau_2(x)\}$ converges to a limit $p_d \in (0, 1)$. Note that this is of the same order of magnitude as $\mathbf{E}|\mathcal{A}_1|$. Is it also true that $\mathrm{Var}(|\mathcal{B}|) = \Theta(\mathrm{Var}(|\mathcal{A}_1|))$ or do these quantities differ significantly?

Acknowledgments. I am grateful to Yuval Peres for introducing me to this problem as well as for helpful discussions on different methods of proof over the course of a summer internship in the Theory Group at Microsoft Research. The idea of considering the symmetrized problem described in the beginning of Section 5 was crucial and motivated the use of symmetrization whenever possible elsewhere in the proof. In addition, I thank Robin Pemantle for communicating the conjecture that led to Theorem 1.1, which was formulated jointly with Peres [16]. Having such an accurate target was rather helpful for crafting the proof, in particular for achieving a precise lower bound. I thank Amir Dembo for pointing out that the original proof of Theorem 1.1 gave both the constant in front of the leading term and held in much greater generality, which yielded Theorem 1.3. Finally, I thank both Amir Dembo and Perla Sousi for many useful comments on an earlier draft of this article.

REFERENCES

[1] BHATTACHARYA, R. N. and RANGA RAO, R. (1976). *Normal Approximation and Asymptotic Expansions*. Wiley, New York. MR0436272

[2] 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

[3] 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

[4] 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

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

[6] DEMBO, A. and SZNITMAN, A.-S. (2006). On the disconnection of a discrete cylinder by a random walk. *Probab. Theory Related Fields* **136** 321–340. MR2240791

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

[8] DICKER, L. (2006). Coloring a $d \geq 3$ dimensional lattice with two independent random walks. M.S. thesis, Univ. Pennsylvania. Available at http://www.math.upenn.edu/~pemantle/papers/Student-theses/Masters/Dicker060421.pdf.

[9] GOMES, S. R. JR., LUCENA, L. S., DA SILVA, L. R. and HILHORST, H. J. (1996). Coloring of a one-dimensional lattice by two independent random walkers. *Phys. A* **225** 81–88. MR1379999

[10] JAIN, N. and OREY, S. (1968). On the range of random walk. *Israel J. Math.* **6** 373–380. MR0243623

[11] JAIN, N. C. and PRUITT, W. E. (1970). The central limit theorem for the range of transient random walk. *Bull. Amer. Math. Soc.* (*N.S.*) **76** 758–759. MR0258135

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

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

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

[15] MILLER, J. and PERES, Y. (2012). Uniformity of the uncovered set of random walk and cutoff for lamplighter chains. *Ann. Probab.* **40** 535–577.

[16] PEMANTLE, R. and PERES, Y. Private communication.

[17] SZNITMAN, A.-S. (2010). Vacant set of random interlacements and percolation. *Ann. of Math.* (2) **171** 2039–2087. MR2680403

[18] WINDISCH, D. (2008). Random walk on a discrete torus and random interlacements. *Electron. Commun. Probab.* **13** 140–150. MR2386070

[19] WINDISCH, D. (2010). Random walks on discrete cylinders with large bases and random interlacements. *Ann. Probab.* **38** 841–895. MR2642893

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