Random walks on graphs: new bounds on hitting, meeting, coalescing and returning*

Roberto I. Oliveira †   Yuval Peres ‡

Abstract

We prove new results on lazy random walks on finite graphs. To start, we obtain new estimates on return probabilities $P^t(x,x)$ and the maximum expected hitting time $t_{\rm hit}$, both in terms of the relaxation time. We also prove a discrete-time version of the first-named author's "Meeting time lemma" that bounds the probability of a random walk hitting a deterministic trajectory in terms of hitting times of static vertices. The meeting time result is then used to bound the expected full coalescence time of multiple random walks over a graph. This last theorem is a discrete-time version of a result by the first-named author, which had been previously conjectured by Aldous and Fill. Our bounds improve on recent results by Lyons and Oveis-Gharan; Kanade et al; and (in certain regimes) Cooper et al.

1 Introduction.

Random walks on graphs and other reversible Markov chains are fundamental mathematical objects, with applications in algorithms and beyond. In this paper, we derive new bounds on return probabilities, hitting times, and meeting deterministic trajectories for these processes. We also bound the expected full coalescence time of multiple random walks on the same graph.

We state our main results in terms of lazy random walks on graphs, although some of them hold for more general chains (cf. Section 2). In the paper $G = (V,E)$ is always a finite, unoriented, connected graph with vertex set V, with no loops or parallel edges and $n := |V| \geq 2$ vertices. The degree of $x \in V$ is denoted by d_x. Minimum, maximum, and average degrees in G are denoted by $d_{\min}$, $d_{\max}$ and $d_{\rm avg}$, respectively. The parameter t always denotes discrete time and takes values in $\mathbf{N} = \{0,1,2,\dots\}$. We write $a \wedge b$ to denote the minimum of $a, b \in \mathbf{R}$.

Definition 1. *Lazy random walk (LRW) on G is the Markov chain on V with transition matrix P given by:*

$$P(x,y) := \begin{cases} \frac{1}{2}, & y = x; \\ \frac{1}{2d_x}, & xy \in E; \\ 0, & \text{otherwise} \end{cases} \qquad (x,y \in V).$$

We let $(X_t)_{t=0}^{+\infty}$ denote trajectories of the chain and $\mathbf{P}_x$, $\mathbf{E}_x$, $\mathbf{P}_\mu$, $\mathbf{E}_\mu$ denote probabilities and expectations when $X_0 = x$ or X_0 has law μ (respectively). Finally, we denote by π the stationary distribution:

$$\pi(x) := \frac{d_x}{d_{\rm avg}\, n}\ (x \in V).$$

As discussed in Section 2, P has nonnegative real spectrum $1 = \lambda_1 > \lambda_2 \geq \lambda_3 \geq \dots \lambda_n \geq 0$. We define the relaxation time $t_{\rm rel}$ as the reciprocal of the spectral gap:

$$t_{\rm rel} := \frac{1}{1-\lambda_2}.$$

Recall that the hitting time of $A \subset V$ is

$$\tau_A := \inf\{t \geq 0 \,:\, X_t \in A\}.$$

We write $\tau_x := \tau_{\{x\}}$ for $x \in V$. The maximum expected hitting time of LRW on G,

$$t_{\rm hit} := \max_{y,x \in V} \mathbf{E}_y\left[\tau_x\right],$$

is a natural parameter for a Markov chain, with connections to electrical network theory, cover times, mixing times and many other aspects of such processes. See [7] (especially Chapter 10) for more information.

1.1 Hitting and returning. Our first two results estimate $t_{\rm hit}$ and return probabilities $P^t(x,x)$ in terms of $t_{\rm rel}$. Note that hitting times and return probabilities are related by the formula [7, Proposition 10.26]:

$$\pi(x)\mathbf{E}_\pi\left[\tau_x\right] = \sum_{s=0}^{+\infty}(P^s(x,x) - \pi(x)). \tag{1.1}$$

Theorem 1.1. (Proof in Section 3) *Under Definition 1,*

$$t_{\rm hit} \leq \frac{20\, d_{\rm avg}}{d_{\min}}\, n\, \sqrt{t_{\rm rel}+1}.$$

*RIO was supported by a *Bolsa de Produtividade em Pesquisa* from CNPq and a *Cientista do Nosso Estado* grant from FAPERJ.

†IMPA. Rio de Janeiro, RJ, Brazil.
‡Microsoft Research. Redmond, WA, USA.

Theorem 1.2. (Proof in Section 3) *Under Definition 1, for any $x \in V$ and $t \geq 0$:*

$$P^t(x,x) - \pi(x) \quad \leq \quad \frac{10\, d_x}{d_{\min}} \left(\frac{1}{\sqrt{t+1}} \wedge \frac{\sqrt{t_{\mathrm{rel}}+1}}{t+1} \right).$$

These Theorems are specific to random walks on graphs. As we explain in Section 2.1, the following bounds are essentially best possible for general reversible chains.

$$(1.2) \qquad t_{\mathrm{hit}} \quad \leq \quad 2 \max_x \left(\frac{1-\pi(x)}{\pi(x)} \right) t_{\mathrm{rel}}$$
$$\left(\approx \frac{d_{\mathrm{avg}}\, n}{d_{\min}}\, t_{\mathrm{rel}} \text{ for } G; \right)$$
$$(1.3) \quad P^t(x,x) - \pi(x) \quad \leq \quad (1 - t_{\mathrm{rel}}^{-1})^t\, (1 - \pi(x)).$$

Kanade et al. [5] obtained the following improvement to (1.2) in the graph setting. Their theorem shows that t_{hit} is at most

$$Kn \left(\left(\frac{d_{\mathrm{avg}}}{d_{\min}} \right)^2 \sqrt{t_{\mathrm{rel}} \log t_{\mathrm{rel}}} \right) \wedge \left(\left(\frac{d_{\mathrm{avg}}}{d_{\min}} \right)^{5/2} t_{\mathrm{mix}} \right),$$

where $t_{\mathrm{mix}} = \Omega(t_{\mathrm{rel}})$ is the mixing time [7, Theorem 12.5] and $K > 0$ is universal. Theorem 1.1 has better dependence on t_{rel} and on the ratio $d_{\mathrm{avg}}/d_{\min}$. The family of lollipop graphs shows that our bound is sharp up to constant factors; see the Appendix for details.

The best previous result for return probabilities on graphs is due to Lyons and Oveis-Gharan [8, Theorem 4.9][1]:

$$(1.4) \qquad P^t(x,x) - \pi(x) < \frac{13\, d_{\max}}{d_{\min}} \frac{1}{\sqrt{t+1}}.$$

Theorem 1.2 improves the constant, replaces $d_{\max}$ with d_x, and gives asymptotic improvements over both (1.4) and (1.3) when $1 \ll t/t_{\mathrm{rel}} \ll \log t_{\mathrm{rel}}$.

Remark 1. *Anna Ben-Hamou (personal communication) improved our original constants in Theorems 1.1 and 1.2. We thank her for her permission to use these improvements.*

1.2 Results on meeting and coalescing.

Our next result bounds the probability that LRW hits a moving target in terms of t_{hit}.

Theorem 1.3. (Proof in Section 4) *Under Definition 1, for any $t \in \mathbf{N}\backslash\{0\}$ and any sequence $h_0, h_1, \ldots, h_t \in V$,*

$$\mathbf{P}_\pi\, (\forall 0 \leq s \leq t \,:\, X_s \neq h_s) \leq \left(1 - \frac{1}{t_{\mathrm{hit}}} \right)^t.$$

This is a discrete-time version of the "Meeting Time Lemma" in [10]. Notice that this Theorem fails for non-lazy random walk on a bipartite graph. On the other hand, Theorem 1.3 can be extended to all reversible Markov chains with nonnegative spectrum (cf. Section 2 below).

The original application of the continuous-time version of Theorem 1.3 was to the study of coalescing random walks. Following the argument in [10], we may obtain a discrete-time version of the main result of that paper, which had been previously conjectured by Aldous and Fill [1].

Theorem 1.4. (Proof in section 5) *Under Definition 1, consider n coalescing random walks that evolve according to LRW on G. Let t_{coal} denote the expected time until these walks have all coalesced into one. Then $t_{\mathrm{coal}} \leq K\, t_{\mathrm{hit}}$ with $K > 0$ a universal constant.*

Kanade et al. [5, Theorem 1.3] obtained a partial form of Theorem 1.4 where $K = O(\log(d_{\max}/d_{\min}))$ grows slowly with the ratio $d_{\max}/d_{\min}$. In particular, their result implies Theorem 1.4 when G is regular. They also obtained sharp results relating coalescing and meeting times in general [5, Theorems 1.1 and 1.2] that do not follow from our methods.

1.3 Additional comments.

We finish the introduction with additional comments on our results. Theorem 1.1 gives a bound on t_{hit} that may be viewed through the lens of electrical network theory. The relation between effective resistances and commute times [7, Chapter 9] gives the following immediate corollary.

Corollary 1.1. (Proof omitted) *Let G be a finite connected graph. Place unit resistances on each edge of G. Then the efffective resistance between any two vertices $x, y \in V$ is bounded by $K\sqrt{t_{\mathrm{rel}}}/d_{\min}$, where $K > 0$ is universal.*

Intriguingly, we do not know of a direct proof of this result via the Dirichlet form or a network reduction.

Theorem 1.2 on return probabilities may be used to tighten results on random walk intersections by Peres et al. [11]. Let t_I be the maximum expected intersection time of two independent lazy random walks over G from worst-case initial states. Proposition 1.8 in [11] gives the following bound for regular graphs:

$$t_I \leq K\sqrt{n}\, [t_{\mathrm{unif}} \wedge (t_{\mathrm{rel}} \log(t_{\mathrm{rel}}+1))]^{3/4},$$

with $K > 0$ universal and t_{unif} the *uniform* mixing time. With Anna Ben-Hamou, we have used Theorem 1.2 to obtain an improved bound that holds for all graphs:

$$t_I \leq K \sqrt{\frac{d_{\mathrm{avg}}\, n}{d_{\min}}}\, (t_{\mathrm{rel}})^{3/4}, \text{ with } K > 0 \text{ universal},$$

[1] That result is for regular graphs, but the authors explain right after the proof that it can be extended to all graphs in the form presented here.

which can be shown to be sharp. This result will appear in the full version of [2].

Finally, the combination of Theorems 1.4 and 1.1 leads to the following corollary.

COROLLARY 1.2. (PROOF OMITTED) *Under Definition 1, let $t_{\rm coal}$ be as in Theorem 1.4. Then:*

$$t_{\rm coal} \leq K \frac{d_{\rm avg}}{d_{\rm min}} n \sqrt{t_{\rm rel}}, \text{ with } K > 0 \text{ universal.}$$

This may be compared to [4, Theorem 1] by Cooper et al.:

$$t_{\rm coal} \leq K \log^4 n \, t_{\rm rel} + K \left(\frac{d_{\rm avg}^2}{\frac{1}{n}\sum_{v\in V} d_v^2}\right) n \, t_{\rm rel}.$$

Our bound improves on this result when G is regular or (more generally) when

$$\frac{1}{n}\sum_{v\in V} d_v^2 \ll d_{\rm avg}\, d_{\rm min} \sqrt{t_{\rm rel}}.$$

On the other hand, Cooper et al.'s bound is stronger when degrees are not balanced (eg. in a star graph). See [4, Section 5] for even stronger bounds when degrees are imbalanced.

2 Preliminaries.

Before we prove our results, we discuss basic properties of LRW. The books [1, 7] contain much more material on reversible chains and random walks on graphs.

LRW as in Definition 1 is an irreducible chain because G is connected. LRW is also reversible: $\pi(x)\, P(x,y) = \pi(y)\, P(y,x)$ for all $x, y \in V$. This means that P, as an operator over $\mathbf{R}^V$, is self-adjoint with respect to the inner product:

$$\langle f, g\rangle := \sum_{x\in V} \pi(x)\, f(x)\, g(x) \ (f, g \in \mathbf{R}^V).$$

So P has real spectrum

$$\lambda_1 = 1 > \lambda_2 \geq \lambda_3 \geq \dots \lambda_n \geq -1,$$

and the fact that $P(x,x) \geq 1/2$ for all $x \in V$ implies $\lambda_n \geq 0$. That is, LRW on a finite connected graph G is a *reversible and irreducible finite Markov chain with nonnegative spectrum*. In particular, P is positive semidefinite and has a positive semidefinite square root $\sqrt{P}$. As in Definition 1, the relaxation time of such a chain is the reciprocal of its spectral gap, $t_{\rm rel} := (1-\lambda_2)^{-1}$.

Let $\mathbf{1} \in \mathbf{R}^V$ denote the function that is equal to 1 everywhere. To each λ_i we may associate an eigenfunction Ψ_i (with $\Psi_1 = \mathbf{1}$) so that $\{\Psi_i\}_{i=1}^n$ is an orthonormal basis of $(\mathbf{R}^V, \langle\cdot,\cdot\rangle)$. We use $\|\cdot\|$ to denote the norm corresponding to the inner product in $\mathbf{R}^V$.

2.1 General bounds for hitting and returning.

Using the above notation, we explain why the bounds on hitting times (1.2) and return probabilities (1.3) are nearly optimal for a reversible chain P with nonnegative spectrum.

Set $\gamma := 1/t_{\rm rel}$, so that $\lambda_2 = 1-\gamma$. Let Π be the matrix with entries $\Pi(x,y) = \pi(y)$ $(x, y \in V)$. Let I denote the identity matrix. The Markov chain that, at each step, stays put with probability $1-\gamma$ and jumps to a point picked from π with probability γ, has transition matrix:

$$P_* := (1-\gamma)\, I + \gamma\, \Pi.$$

P_* has the same eigenbasis as P, with eigenvalues 1 (with multiplicity 1, eigenvector $\mathbf{1}$) and $\lambda_2 = 1-\gamma$ (multiplicity $n-1$). Thus P and P_* have nonnegative spectrum, the same relaxation time, and moreover $\langle f, P^t f\rangle \leq \langle f, P_*^t f\rangle$ for all $t \geq 0$ and $f \in \mathbf{R}^V$. In particular, for all $x \in V$ and $t \geq 0$,

$$\begin{aligned} P^t(x,x) - \pi(x) &\leq P_*^t(x,x) - \pi(x) \\ &= (1-\gamma)^t\, (1-\pi(x)) \end{aligned}$$

and (with $\mathbf{E}_*$ denoting expectation for P_*)

$$\pi(x)\mathbf{E}_\pi\, [\tau_x] \leq \pi(x)\mathbf{E}_{*,\pi}[\tau_x] = (1-\pi(x))\, \gamma^{-1}.$$

In particular, (1.3) is exact for P_*, whereas (1.2) is sharp up to a constant factor.

REMARK 2. *As noted in the introduction, Theorem 1.1 improves upon these bounds in the graph setting. In the Appendix we give a family of examples that achieve Theorem 1.1 up to constant factors.*

3 Return probabilities and hitting times.

In this section we prove Theorem 1.1 on $t_{\rm hit}$ and Theorem 1.2 on $P^t(x,x)$. We will need two lemmas on "Green's function":

$$g_t(x,x) := \sum_{s=0}^{t} P^s(x,x). \tag{3.5}$$

The first lemma relates the Green's function to transition probabilities and shows how one can control $g_t(x,x)$ for $t \geq t_{\rm rel}$ in terms of $g_{\lceil t_{\rm rel}\rceil - 1}(x,x)$.

LEMMA 3.1. (PROOF IN §3.1) *Consider P, a reversible and irreducible finite Markov chain with nonnegative spectrum, stationary distribution π and state space V. Fix $x \in V$ and $t \geq 0$. Then:*

$$0 \leq P^t(x,x) - \pi(x) \leq \frac{g_t(x,x) - (t+1)\pi(x)}{t+1}$$

and

$$g_t(x,x) - (t+1)\pi(x) \leq \left(\frac{e}{e-1}\right) g_{(\lceil t_{\rm rel}\rceil - 1)\wedge t}(x,x).$$

The next lemma gives upper bounds for Green's function for $t \le t_{\mathrm{rel}}$.

LEMMA 3.2. (PROOF IN §3.2) *Under Definition 1, for $x \in V$ and $0 \le t \le \lceil t_{\mathrm{rel}} \rceil$:*

$$\frac{g_t(x,x)}{\pi(x)} \le 6\,\frac{d_{\mathrm{avg}}\,n}{d_{\min}}\,\sqrt{t+1}.$$

Let us see how Theorem 1.1 and Theorem 1.2 follow from the propositions.

Proof. [of Theorem 1.1] For $x \in V$, formula (1.1) may be rewritten as:

$$\mathbf{E}_\pi\left[\tau_x\right] = \frac{\lim_{t\to+\infty}(g_t(x,x) - (t+1)\pi(x))}{\pi(x)}.$$

Lemma 3.1, Lemma 3.2 and the formula for $\pi(x)$ give:

$$\mathbf{E}_\pi\left[\tau_x\right] \le \frac{e}{e-1}\,\frac{g_{(\lceil t_{\mathrm{rel}}\rceil-1)}(x,x)}{\pi(x)},$$

and this last quantity is upper bounded by

$$\left(\frac{6e}{e-1}\right)\frac{d_{\mathrm{avg}}\,n}{d_{\min}}\,\sqrt{t_{\mathrm{rel}}+1}$$

To finish, we bound $6e/(e-1) \le 10$ and use $t_{\mathrm{hit}} \le 2\max_{x\in V}\mathbf{E}_\pi\left[\tau_x\right]$ [7, Lemma 10.2].

Proof. [of Theorem 1.2] Lemma 3.1 implies

$$P^t(x,x) - \pi(x) \le \left(\frac{e}{e-1}\right)\frac{g_{(\lceil t_{\mathrm{rel}}\rceil-1)\wedge t}(x,x)}{t+1}.$$

We bound the Green's function via Lemma 3.2 and obtain:

$$P^t(x,x) - \pi(x) \le \left(\frac{6e}{e-1}\right)\frac{d_x}{d_{\min}}\,\frac{\sqrt{(t+1)\wedge(t_{\mathrm{rel}}+1)}}{t+1}.$$

The result follows from estimating the constant by 10.

3.1 Sum up to the relaxation time and stop.

We prove here Lemma 3.1.

Proof. [of Lemma 3.1] We use the notation in Section 2. Note that there exists a $f \in \mathbf{R}^V$ such that, for all $t \ge 0$:

$$P^t(x,x) - \pi(x) = \langle f, P^t f\rangle - \langle f, \mathbf{1}\rangle^2,$$

and the spectral decomposition gives:

$$P^t(x,x) - \pi(x) = \sum_{i=2}^n \lambda_i^t\,\langle \Psi_i, f\rangle^2. \tag{3.6}$$

Since $0 \le \lambda_i \le \lambda_2 < 1$ for each $i \in [n]\backslash\{1\}$, $P^t(x,x) - \pi(x)$ decreases with t. This gives the first statement in the Lemma. Adding the λ_i^t in (3.6), we also obtain:

$$g_t(x,x) - (t+1)\pi(x) = \sum_{i=2}^n\left(\frac{1-\lambda_i^{t+1}}{1-\lambda_i}\right)\langle \Psi_i, f\rangle^2.$$

The terms with λ_i are $\le (1-\lambda_i)^{-1}$ for all t. On the other hand, if $t = \lceil t_{\mathrm{rel}}\rceil - 1$, $0 \le \lambda_i^{t+1} \le \lambda_2^{t_{\mathrm{rel}}} \le e^{-1}$, and

$$g_{\lceil t_{\mathrm{rel}}\rceil-1}(x,x) - \lceil t_{\mathrm{rel}}\rceil\pi(x) \ge \sum_{i=2}^n\left(\frac{1-e^{-1}}{1-\lambda_i}\right)\langle \Psi_i, f\rangle^2.$$

The proof finishes by combining the two last displays.

3.2 Green's function and electrical network theory.

We will prove Lemma 3.2 by adapting an old argument of Aldous and Fill [1, Proposition 6.16] to non-regular graphs. The proof is based on the fact that $g_t(x,x)$ counts the expected number of returns of LRW to x up to time t. Fixing $\alpha > 1$, we define the set

$$A_\alpha := \{y \in V \,:\, g_t(y,x) \le \alpha\,\pi(x)\,(t+1)\}. \tag{3.7}$$

This set is defined so that LRW makes few returns to x after hitting A_α. We also note that, by reversibility, $\pi(y)g_t(y,x) = \pi(x)g_t(x,y)$, so:

$$1 - \pi(A_\alpha) \le \sum_{y\in V}\frac{\pi(y)\,g_t(y,x)}{\alpha\,(t+1)\pi(x)} = \frac{\sum_{y\in V} g_t(x,y)}{\alpha\,(t+1)} = \frac{1}{\alpha}.$$

So A_α is large for α large. The idea, then, is to show that LRW will make few returns to x *before* hitting A_α (cf. Proposition 3.2 below).

The proof of Lemma 3.2 will be finished at the end of this subsection. We first need some estimates that we obtain via the theory of electrical networks [7, Chapter 9]. We assign conductance $c(a,b) = 1 = c(b,a)$ to each edge $ab \in E$ and set $c(a,a) = d_a$ for each $a \in V$. All other pairs a', b' (with $a' \ne b'$ and $a'b' \notin E$) have zero conductance. In particular, $c(a) = 2d_a$ for each $a \in V$. We will also use a standard *path fact* that follows from the argument in the end of the proof of [7, Proposition 10.16, part (b)].

FACT 3.1. (PATH FACT) *If $x_0, \ldots, x_\ell$ is a geodesic path in G, then $\ell \le 3n/d_{\min} - 1$.*

We need an a priori bound on the relaxation time.

PROPOSITION 3.1. *Under Definition 1,*

$$t_{\mathrm{rel}} \le \max_{x,y\in V}(\mathbf{E}_x\left[\tau_y\right] + \mathbf{E}_y\left[\tau_x\right]) \le 6\left(\frac{d_{\mathrm{avg}}}{d_{\min}}\right)n^2 - 4.$$

Proof. The first inequality follows from [7, Lemma 12.17]. To bound the maximum commute time, we note that, by a simple network reduction argument,

$$\max_{x,y\in V}(\mathbf{E}_x[\tau_y] + \mathbf{E}_y[\tau_x])$$

is at most the diameter of G times $2d_{\rm avg}\,n$. Since the diameter is $\leq 3n/d_{\min} - 1$ by the Path Fact and $d_{\rm avg}\,n = 2|E| \geq 2$, the Proposition follows.

The next proposition is the key step where we show that LRW returns few times to x before hitting a large set A. It will be used with $A = A_\alpha$.

PROPOSITION 3.2. *Under Definition 1, let $A \subset V$ be nonempty. Take $x \in V\backslash A$ and consider the number of returns to x up to time $\tau_A - 1$:*

$$g_{\tau_A-1}(x,x) := \mathbf{E}_x\left[\sum_{s=0}^{\tau_A-1}\mathbf{I}_{\{X_s=x\}}\right].$$

Then:

$$\frac{g_{\tau_A-1}(x,x)}{\pi(x)} \leq 9\left(\frac{d_{\rm avg}\,n}{d_{\min}}\right)^2(1-\pi(A)).$$

Proof. Thanks to [7, Lemma 9.6] and a standard network reduction:

$$\frac{g_{\tau_A-1}(x,x)}{\pi(x)} = (2d_{\rm avg}\,n)\,R_{\rm eff}(x\leftrightarrow A)$$

where $R_{\rm eff}$ denotes effective resistance (note that $\pi(x) = c(x)/2d_{\rm avg}n$). Letting $B := V\backslash A$, our main goal will be to show:

$$\textbf{Goal:}\quad R_{\rm eff}(x\leftrightarrow A) \leq \frac{9|B|}{2d_{\min}}$$

as the result then follows from the fact that:

$$|B| \leq \sum_{x\in B}\frac{d_x}{d_{\min}} = \frac{d_{\rm avg}\,n}{d_{\min}}\,\pi(B).$$

To bound the effective resistance, consider first the case $|B| \leq 2d_{\min}/3$. In that case, x has at least $d_x/3$ neighbors in $A = V\backslash B$. Using that $|B| \geq 1$ we obtain that $R_{\rm eff}(x\leftrightarrow A) \leq (\sum_{a\in A}c(x,a))^{-1}$ satisfies:

$$R_{\rm eff}(x\leftrightarrow A) \leq \frac{3}{d_x} \leq \frac{9|B|}{2d_{\min}}.$$

We now consider the case $|B| > 2d_{\min}/3$. Let $y \in A$ be as close as possible to x (in the graph distance). Since each edge ab has resistance 1, we have:

$$R_{\rm eff}(x\leftrightarrow A) \leq \mathrm{dist}(x,y). \tag{3.8}$$

We now bound the graph distance $k := \mathrm{dist}(x,y)$. Let $B := V\backslash A$. If $x = x_0,\ldots,x_k = y$ is a shortest path from x to y, then each of the vertices $x_0,x_1,\ldots,x_{k-2} \in B$ has all of their neighbors in B, because all points of A are at distance $\geq k$ from x. We apply the Path Fact to the geodesic path $x_0,x_1,\ldots,x_{k-2}$ in the induced subgraph $G[B]$ and obtain:

$$k \leq \frac{3|B|}{d_{\min}} + 1 < \frac{9|B|}{2d_{\min}}.$$

because we are assuming $|B| > 2d_{\min}/3$.

We may now complete the proof of Lemma 3.2.

Proof. (of Lemma 3.2). Take $\alpha > 1$. We use the definition of A_α in (3.7) and the estimate $1 - \pi(A_\alpha) \leq 1/\alpha$ obtained at the beginning of the subsection. In particular, $A_\alpha \neq \emptyset$ when $\alpha > 1$. We now *claim* that:

$$\textbf{Claim:}\quad \frac{g_t(x,x)}{\pi(x)} \leq \frac{9}{\alpha}\left(\frac{d_{\rm avg}\,n}{d_{\min}}\right)^2 + \alpha\,(t+1). \tag{3.9}$$

In fact, assuming this we obtain Lemma 3.2 by setting

$$\alpha = \frac{3\,d_{\rm avg}\,n}{d_{\min}\sqrt{t+1}},$$

noting that $\alpha > 1$ when $t \leq \lceil t_{\rm rel}\rceil$ thanks to Proposition 3.1.

To prove (3.9), we may assume $x \notin A_\alpha$. Note that the number of returns to x up to time t is at most the sum of returns up to time $\tau_{A_\alpha} - 1$ with those occurring at times $\tau_{A_\alpha} \leq s \leq \tau_{A_\alpha}+t$. The strong Markov property gives:

$$\frac{g_t(x,x)}{\pi(x)} \leq \frac{g_{\tau_{A_\alpha}-1}(x,x)}{\pi(x)} + \mathbf{E}_x\left[\frac{g_t(X_{\tau_{A_\alpha}},x)}{\pi(x)}\right],$$

and inequality (3.9) follows from Proposition 3.2 (applied to the first term in the RHS) and the definition of A_α (applied to the second term in the RHS).

4 The Meeting Time Theorem in general form.

We now present a more general version of the Meeting Time Theorem (Theorem 1.3 above). We use the same notation and definitions for trajectories, hitting times and other quantities that we defined for LRW. We also take the material and notation from Section 2 for granted.

THEOREM 4.1. *Consider a reversible and irreducible finite Markov chain with nonnegative spectrum P which is defined over a finite set $V \neq \emptyset$ and has stationary measure π. Then for any $t \geq 0$ and any choice of $h_0,\ldots,h_t \in V$:*

$$\mathbf{P}_\pi(\forall 0 \leq s \leq t\,:\,X_s \neq h_s) \leq \left(1 - \frac{1}{t_{\rm hit}}\right)^t.$$

Proof. Given a linear operator $A : \mathbf{R}^V \to \mathbf{R}^V$, we denote its operator norm by:

$$\|A\|_{\rm op} := \sup\{\langle g, Af\rangle \,:\, f, g \in \mathbf{R}^V,\, \|f\| = \|g\| = 1\}.$$

For $h \in V$, let D_h be a $|V| \times |V|$ diagonal matrix that that has 1's at entries (x, x) with $x \neq h$ and a zero at the entry (h, h). D_h is self-adjoint with respect to $\langle\cdot,\cdot\rangle$. Defining

$$M_t := D_{h_0}\, P\, D_{h_1}\, P\, D_{h_2}\, \ldots D_{h_{t-1}}\, P\, D_{h_t},$$

we see that

$$\mathbf{P}_\pi\left(\forall 0 \le s \le t \,:\, X_s \neq h_s\right) = \langle \mathbf{1}, M_t\mathbf{1}\rangle \le \|M_t\|_{\rm op}.$$

Thus it suffices to estimate the norm of M_t. We do this in two steps.

1: $\|M_t\|_{\rm op} \le (\max_{h\in V} \|D_h\, P\, D_h\|_{\rm op})^t$;

2: For any $h \in V$, $\|D_h\, P\, D_h\|_{\rm op} \le 1 - t_{\rm hit}^{-1}$.

For step 1 we employ $\sqrt{P}$ (defined in Section 2) and the fact that $D_{h_s} = D_{h_s}^2$ for each s, and the fact that $\|\cdot\|_{\rm op}$ is submultiplicative:

$$\begin{aligned}
\|M_t\|_{\rm op} &= \left\|\prod_{s=0}^{t-1}\left(D_{h_s}\, P\, D_{h_{s+1}}\right)\right\|_{\rm op}\\
&= \left\|\prod_{s=0}^{t-1}\left(D_{h_s}\sqrt{P}\,\sqrt{P}\, D_{h_{s+1}}\right)\right\|_{\rm op}\\
&\le \prod_{s=0}^{t-1}\left(\|D_{h_s}\sqrt{P}\|_{\rm op}\, \|\sqrt{P}\, D_{h_{s+1}}\|_{\rm op}\right).
\end{aligned}$$

This last product of norms contains $2t$ terms of the form $\|D_h\sqrt{P}\|_{\rm op}$ or $\|\sqrt{P}D_h\|_{\rm op}$ for elements $h \in V$. Now, for *any* linear operator A with adjoint $A^\dagger$, $\|A^\dagger\|_{\rm op}$ and $\|A\|_{\rm op}$ are both equal to $\sqrt{\|A^\dagger A\|_{\rm op}}$. Apply this to $A := \sqrt{P}\, D_h$ and obtain:

$$\|D_h\sqrt{P}\|_{\rm op} = \|\sqrt{P}D_h\|_{\rm op} = \sqrt{\|D_h\, P\, D_h\|_{\rm op}}.$$

Thus step 1 follows from the inequality for $\|M_t\|_{\rm op}$.

For step 2, we fix $h \in V$. Notice that $D_h\, P\, D_h$ is self-adjoint and positive semidefinite (because P is), so $\|D_h\, P\, D_h\|_{\rm op}$ equals the largest eigenvalue of $D_h\, P\, D_h$. According to [1, Section 3.6.5 and Theorem 3.33], the largest eigenvalue is equal to $1 - \mathbf{E}_{q_h}\left[\tau_h\right]^{-1}$ for some quasistationary distribution q_h over $V\backslash\{h\}$. Since $\mathbf{E}_{q_h}\left[\tau_h\right] \le t_{\rm hit}$, the result follows.

5 Coalescing random walks.

In this section we will present a generalization of Theorem 1.4. To state it, we first define the process.

Let P be a Markov chain over a finite set V. Following [10, Section 3.3], we use a definition in terms of "killed particles". That is, when several particles meet on the same vertex at the same time, only the particle with lower index survives (one can alternatively think that these particles coalesce).

To make this formal, write $V = \{v_1, \ldots, v_n\}$ and let

$$(X_t(a))_{t\ge 0} \,:\, a = 1, 2, 3, \ldots, n$$

be n independent trajectories on V evolving according to P, each with initial state $X_0(i) = v_i$. Now define killed trajectories $Y_t(a)$ as follows. Let $\partial \notin V$ be a "coffin state". We set $Y_t(1) = X_t(1)$ for all t. Given $1 < a \le n$, assume $Y_s(1), \ldots, Y_s(a-1)$ have been defined. Now let[2]:

$$\kappa_a := \inf\{t \ge 0 \,:\, X_t(a) = Y_t(b) \text{ for some } b < a\}.$$

For $t \ge 0$, we set

$$Y_t(a) := \begin{cases} X_t(a), & t < \kappa_a;\\ \partial, & t \ge \kappa_a.\end{cases}$$

One can check that the set valued process

$$S_t := \{Y_t(a) \,:\, 1 \le a \le n, Y_t(a) \neq \partial\}$$

is a time-homogeneous Markov chain on $2^V\backslash\{\emptyset\}$. It is this process that we call coalescing random walks evolving according to P. The *full coalescence time*[3] is:

$$\tau_{\rm coal} := \inf\{t \ge 0 \,:\, |S_t| = 1\}.$$

THEOREM 5.1. *Consider a reversible and irreducible finite Markov chain with nonnegative spectrum P which is defined over a finite set $V \neq \emptyset$. Let $\tau_{\rm coal}$ denote the full coalescence time of a system of coalescing random walks that evolve according to P. Let $t_{\rm coal} := \mathbf{E}\left[\tau_{\rm coal}\right]$. Then:*

$$t_{\rm coal} \le K\, t_{\rm hit}, \text{ where } K > 0 \text{ is universal.}$$

Proof. We explain how the proof of [10] can be modified to work in the discrete time setting. As in [10, Proposition 4.1], it suffices to prove that:

$$\mathbf{P}\left(\tau_{\rm coal} \ge c\,(t_{\rm mix} + t_{\rm hit})\right) \le 1 - \eta$$

for some universal $c, \eta > 0$. We do this by considering a process where less particles die. That is, assume that for each $t \ge 0$ we have a set $\mathcal{A}_t \subset [n]^2$ of 'allowed killings".

[2] This time κ_a is called τ_a in [10].
[3] Called C in [10].

Define a new killed process $Y_\cdot^{\mathcal{A}}(i)$, $i \in [n]$, by setting $Y_t^{\mathcal{A}}(1) = X_t(1)$ for all i; redefining κ_a as

$$\kappa_s^{\mathcal{A}} := \inf \left\{ t \geq 0 \ : \ \begin{array}{c} \exists (b,a) \in \mathcal{A}_t \text{ with } b < a \\ \text{and } X_t(a) = Y_t(b) \end{array} \right\};$$

and defining $Y_t^{\mathcal{A}}(a)$ accordingly. As in [10, Proposition 3.4], we may observe that the full coalescence time $\tau_{\mathrm{coal}}^{\mathcal{A}}$ of this modified process dominates τ_{coal} in the sense that $\mathbf{P}(\tau_{\mathrm{coal}} \geq t) \leq \mathbf{P}\left(\tau_{\mathrm{coal}}^{\mathcal{A}} \geq t\right)$ for all $t \geq 0$.

We then apply the same strategy as in [10, Section 4.2]. We partition $[n] = A_0 \cup A_1 \cup \ldots \cup A_m$ so that $A_0 = \{1\}$,

$$A_i = \{\max A_{i-1} + k \ : \ k = 1, 2, \ldots, 2^i\}$$

for each $1 \leq i \leq m$, and $|A_m| \leq 2^m$ (so m is of order $\log n$). We define a set of epochs by setting $t_\infty = 0$ and defining t_m, t_{m-1}, t_{m-2}, …, t_0 as:

$$\begin{aligned} t_m &:= 1 + \lceil 2t_{\mathrm{mix}} \rceil, \\ t_j &= 1 + t_{j-1} + \left\lceil \frac{2^4 \ln 5}{2^j}\, t_{\mathrm{hit}} \right\rceil. \end{aligned}$$

This follows the definition in the paper except for "$1 + \lceil\cdot\rceil$" terms. The sets of allowed killings $\mathcal{A}_t$ are defined in the same way as in that paper:

1. *Epoch # ∞:* $\mathcal{A}_t = \emptyset$ for $t < t_m$;
2. *Epochs # m through # 1:* for $t_j \leq t < t_{j-1}$, $\mathcal{A}_t := A_{j-1} \times \cap_{i=j}^m A_i$.
3. *Epoch # 0:* set $\mathcal{A}_t = [n]^2$

Letting c_0, c denote universal constants, we obtain:

$$t_0 \leq 2t_{\mathrm{mix}} + 2(m+1) + c_0\, t_{\mathrm{hit}} \leq c\,(t_{\mathrm{mix}} + t_{\mathrm{hit}})$$

where we have $m = O(\log n)$ whereas t_{hit} is at least linear in n. The proofs of [10, Propositions 4.2 - 4.4] all go through as in that paper when one uses our version of the Meeting Time Lemma, Theorem 4.1. One can then finish the proof as in that article.

A Sharpness of the hitting time bound.

In this Appendix, we give a family of graphs where Theorem 1.1 is sharp. We note without proof that one may also construct regular graphs with this property. In fact, the graphs in [2, Proposition 4.1] provide such an example.

Definition 2. (Lollipop) *Choose integers $n \geq 3$ and k. A (n,k)-lollipop is a graph G built as follows. Take a clique of size $n-k$ on vertices $V_0 := \{x_0, x_1, \ldots, x_{n-k-1}\}$ and add k additional vertices $x_1^*, \ldots, x_k^*$ and edges $x_0x_1^*$, $x_1^*x_2^*$, …, $x_{k-1}^*x_k^*$. So G is a clique with $n-k$ vertices with a path of length k attached to some vertex.*

This family of graphs is known to contain the graph with largest t_{hit} over all n-vertex graphs [3]. We show that all lollipops with k not too close to n achieve the bound in Theorem 1.1 up to constant factors.

Theorem A.1. *There exists a universal $c > 0$ such that, if G is a (n,k)-lollipop with $n - k \geq \sqrt{6n} + 1$, then:*

$$t_{\mathrm{hit}} \geq c\, \frac{d_{\mathrm{avg}}}{d_{\min}} \sqrt{t_{\mathrm{rel}}}\, n.$$

Proof. For simplicity, we use $O/\Theta/\Omega$ asymptotic notation; all implied constants are uniform over $1 \leq k \leq n - \sqrt{6n} + 1$.

We need to compute the relevant parameters of G. It is immediate that $d_{\min} = 1$ and the average degree is

$$d_{\mathrm{avg}} = \frac{(n-k-1)(n-k) + 2k}{n}.$$

Notice that $(n-k-1)(n-k) \geq 6n$ under our assumptions. This term is always larger than $2k \leq 2n$, so

$$\frac{d_{\mathrm{avg}}}{d_{\min}} = \Theta\left(\frac{(n-k)^2}{n}\right).$$

We now argue that $t_{\mathrm{hit}} = \Omega((n-k)^2\, k)$ via a lower bound on $\mathbf{E}_{x_0}[\tau_{x_k}]$. If LRW starts from x_0, its only chance of hitting x_k before returning to x_0 is that it jumps to x_1 and then hits x_k before returning x_0. The chance of jumping to x_1 is $1/2(n-k)$. The probability of x_k being hit first is $1/k$ (this is a classical calculation on a path [7, Proposition 2.1]). This implies:

$$\mathbf{P}_{x_0}(X_t \text{ hits } x_k \text{ before returning to } x_0) = \frac{1}{2(n-k)\,k}.$$

Since random walk trajectories between returns to x_0 are independent, this means that $2(n-k)k$ returns to x_0 are needed on average before x_k is hit. Now, for any irreducible and aperiodic Markov chains, returning to x_0 takes $\pi(x_0)^{-1}$ steps on average [7, §1.5.3]. In our case, $\pi(x_0)^{-1} = \Theta(n-k)$. Using the again independence of the trajectories between returns, we obtain:

$$\mathbf{E}_{x_0}[\tau_{x_k}] \geq \frac{2\,(n-k)k}{\pi(x_0)} = \Omega((n-k)^2\, k).$$

We now argue that t_{rel} is of order k^2, noting that this suffices to finish the proof.

We need some preliminary facts. We need to consider the mixing time t_{mix} of graph G [7, page 54], which is an upper bound for t_{rel} up to a universal constant factor [7, Theorem 12.5]. The exact definition of t_{mix} does not matter to us; what we do need is that "mixing times are hitting times of large sets" (see [7,

§24.6 and 24.7] for an exposition). Given $0 < \alpha \le 1/2$, if we define

$$t_{\mathrm{hit}}(\alpha) := \max\{\mathbf{E}_x[\tau_A] \,:\, x \in V,\, A \subset V,\, \pi(A) \ge \alpha\},$$

then [12, 9, 6] show that

$$\frac{t_{\mathrm{mix}}}{C(\alpha)} \le t_{\mathrm{hit}}(\alpha) \le C(\alpha)\, t_{\mathrm{mix}}$$

for some $C(\alpha) > 0$ independent of G. The upshot is that an upper bound on $t_{\mathrm{hit}}(\alpha)$ is also an upper bound on t_{rel} and t_{mix} (up to universal constant factors).

We now argue that, in our concrete example, $t_{\mathrm{hit}}(1/2) = O(k^2)$. To see this, observe that, under our assumptions, the clique part of the graph and the path part of the graph satisfy:

$$\frac{\pi(V_0)}{\pi(\{x_1^*, \ldots, x_k^*\})} \ge \frac{(n-k-1)^2}{2k} \ge 3 \Rightarrow \pi(V_0) \ge \frac{3}{4}.$$

Now fix a starting point $x \in V$ and a subset $A \subset V$ with $\pi(A) \ge 1/2$. It follows from the above that, letting $A_0 := A \cap V_0$, we have $\pi(A_0) \ge 1/4$. In particular, once LRW reaches the clique V_0, it has probability $\ge 1/8$ of hitting A_0 at the next step (we lose a factor of $1/2$ because the random walk is lazy). Since the expected time to leave $\{x_1^*, \ldots, x_k^*\}$ is $\le (k+1)^2/4$ in expectation [7, Proposition 2.1], it follows that in time $(k+1)^2 + 1$ steps there is a positive probability $p > 0$, independent of all problem parameters, that A is hit, no matter the starting point. That is,

$$\forall x \in V \,:\, \mathbf{P}_x\left(\tau_A \ge (k+1)^2 + 1\right) \le 1 - p.$$

Let $f(r) := \mathbf{P}_x\left(\tau_A \ge r\,[(k+1)^2+1]\right)$. The Markov property implies that for all $r \in \mathbf{N}\backslash\{0\}$,

$$\frac{f(r)}{f(r-1)} \le \max_x \mathbf{P}_x\left(\tau_A \ge (k+1)^2 + 1\right) \le 1 - p,$$

so $f(r) \le (1-p)^r$. From this one easily obtains $\mathbf{E}_x[\tau_A] = O(k^2)$ uniformly over $x \in V$ and $A \subset V$ with measure $\ge 1/2$. So $t_{\mathrm{hit}}(1/2) = O(k^2)$ in our family of graphs.

References

[1] D. Aldous and J. A. Fill, *Reversible Markov Chains and Random Walks on Graphs*, manuscript available from www.stat.berkeley.edu/∼aldous/RWG/book.html (1994).

[2] A. Ben-Hamou, R. I. Oliveira, and Y. Peres. *Estimating graph parameters via random walks with restarts,* in Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms (2018), pp. 1702–1714.

[3] G. Brightwell and P. Winkler, *Maximum hitting time for random walks on graphs,* Random Structures and Algorithms, 1 (1990), pp. 263 – 276.

[4] C. Cooper, R. Elssser, H. Ono, and T. Radzik, *Coalescing random walks and voting on connected graphs,* SIAM Journal of Discrete Mathematics, 27 (2013), pp. 1748 – 1758.

[5] V. Kanade, F. Mallman-Trenn, T. Sauerwald, *On coalescence time in graphs – When is coalescing as fast as meeting?,* arXiv:1611.02460 (2016), 78 pp.

[6] S. Griffiths, R. Kang, R. I. Oliveira, V. Patel, *Tight inequalities among set hitting times in Markov chains,* Proceedings of the American Mathematical Society, 142 (2014), pp. 3285 – 3298.

[7] D. Levin and Y. Peres, *Markov chains and mixing times (2nd edition)*, American Mathematical Society (2017).

[8] R. Lyons, S. Oveis-Gharan, *Sharp bounds on random walk eigenvalues via spectral embedding,* International Mathematics Research Notices, rnx082 (2017).

[9] R. I. Oliveira, *Mixing and hitting times for finite Markov chains,* Electronic Journal of Probability 17 (2012), paper no. 70, 12 pp.

[10] R. I. Oliveira, *On the coalescence time of reversible random walks,* Transactions of the American Mathematical Society, 364 (2012), pp. 2109 – 2128.

[11] Y. Peres, T. Sauerwald, P. Sousi, and A. Stauffer, *Intersection and mixing times for reversible chains,* Electronic Journal of Probability, 22 (2017), paper no. 12, 16 pp.

[12] Y. Peres and P. Sousi, *Mixing times are hitting times of large sets,* Journal of Theoretical Probability, 28 (2015), pp. 488 – 519 .