Reversible random walks on dynamic graphs

Nobutaka Shimizu¹ | Takeharu Shiraga²

¹Department of Industrial Engineering and Economics, Tokyo Institute of Technology, Tokyo, Japan
²Department of Information and System Engineering, Chuo University, Tokyo, Japan

Abstract
This paper discusses random walks on edge-changing dynamic graphs. We prove general and improved bounds for mixing, hitting, and cover times for a random walk according to a sequence of irreducible and reversible transition matrices with the time-invariant stationary distribution. An interesting consequence is the tight bounds of the lazy Metropolis walk on any dynamic connected graph. We also prove bounds for multiple random walks on dynamic graphs. Our results extend previous upper bounds for simple random walks on dynamic graphs and give improved and tight upper bounds in several cases. Our results reinforce the observation that time-inhomogeneous Markov chains with an invariant stationary distribution behave almost identically to a time-homogeneous chain.

KEYWORDS
dynamic graph, Markov chain, random walk

1 | INTRODUCTION

A random walk is a basic stochastic process on an undirected graph. A walker starts from a specific vertex of a graph. At each discrete time step, a walker moves to a random neighbor. The probability that the walker moves from $u$ to $v$ is given by $P(u, v)$, where the matrix $P \in [0, 1]^{V \times V}$ is called a transition matrix. Because of their locality, simplicity, and low memory requirements, random walks have a wide range of applications including network analysis, computational complexity, and distributed algorithms [14, 26]. The efficiency of a random walk can be measured by the rate of diffusion, which has been formalized by several notions including mixing time, hitting time, and cover time. The mixing time is the time it takes for the distribution of the walker to converge to a limit distribution (called a stationary distribution). The hitting time is the maximum expected time for the walker to visit a target vertex, where the maximum is taken over the starting and target vertex. The cover time is the expected time for the walker to visit all vertices starting from the worst vertex. The mixing, hitting, and cover times are important quantities in the study of random walks.
Our results

In this paper, we support the insight that a random walk on a dynamic graph with the same stationary distribution behaves almost identically to that on a static graph by studying a random walk according to time-inhomogeneous transition matrices: Given a sequence of transition matrices $P = (P_t)_{t \geq 1}$ where $P_t \in [0, 1]^{V \times V}$ for all $t$, we consider a random walk according to $P = (P_t)_{t \geq 1}$ that is a sequence of random variables $(X_t)_{t \geq 0}$ satisfying $\Pr[X_t = v_t | X_0 = v_0, \ldots, X_{t-1} = v_{t-1}] = \Pr[X_t = v_t | X_{t-1} = v_{t-1}] = P_t(v_{t-1}, v_t)$ for any $t \geq 1$ and $(v_0, \ldots, v_t) \in V^{t+1}$ satisfying $\Pr[X_0 = v_0, \ldots, X_{t-1} = v_{t-1}] > 0$. In other words, at the $t$th time step ($t \geq 1$), the walker at vertex $u$ randomly selects a vertex according to the distribution $P_t(u, \cdot)$. Our interest is to bound the mixing, hitting, cover, and coalescing times under the assumption that all $P_t$ have the same stationary distribution and are irreducible.

In order to present our results, we briefly introduce some important terminology. A transition matrix $P$ is irreducible if for any $u, v \in V$ there exists a $t > 0$ such that $P^t(u, v) > 0$ holds and reversible if there exists $\pi \in \mathbb{R}^V$ such that $\pi(u)P(u, v) = \pi(v)P(v, u)$ holds for any $u, v \in V$. times on static graphs have been studied extensively for decades [3, 9, 23, 30, 35, 36]. For example, Aleliunas, Karp, Lipton, Lovász, and Rackoff [3] proved that the cover time of the simple random walk on any $n$-vertex connected graph is $O(n^3)$.

Recently, there has been a growing interest in a random walk on a dynamic graph since real-world networks change their structure over time [5, 11, 15, 32, 33, 41, 46]. In this setting, at the beginning of the $r$th round, the walker moves to a random neighbor on the current graph and then the edge set of the graph changes (we assume that the vertex set is static). A central interest is the difference between random walks on a dynamic graph and a static one. Indeed, while the (lazy) simple random walk has an $O(n^3)$ hitting time for any $n$-vertex static graphs, there is a sequence $(G_t)_{t \geq 1}$ of connected graphs called the Sisyphus Wheel (Figure 1) on which the hitting time is exponential [5].

On the other hand, several researchers have observed that random walks on dynamic graphs behave almost identically to those on static graphs when the stationary distribution of a random walk does not change over time [5, 19, 43, 46]. Avin, Koucký, and Lotker [5] considered a random walk called $d_{\text{max}}$-lazy walk that has the uniform stationary distribution on any graph. They proved that the cover time of this walk on any dynamic connected graph is $O(n^3 \log^2 n)$, which was later improved by Denysyuk and Rodrigues [19]. Sauerwald and Zanetti [46] considered the lazy simple random walk on a dynamic connected graph with the same time-invariant degree distribution. They obtained tight or nearly tight bounds for the mixing and hitting times. For example, they showed that both the mixing and hitting times are $O(n^2)$ on any dynamic regular graph. These bounds are tight up to a constant factor even on static regular graphs. See Section 1.5 for more details on previous work.

\begin{figure}
\centering
\includegraphics[width=\textwidth]{fig1.png}
\caption{The Sisyphus Wheel of five vertices. The Sisyphus Wheel $G = (G_t)_{t \geq 1}$ is defined as follows: For each $t \geq 1$, let $V = V(G_t) = \{0, \ldots, n-1\}$, $v(t) = t \mod (n-1)$, and $E(G_t) = \{(v(t), i) : i \in V \setminus \{v(t)\}\}$. The lazy simple random walk starting from the vertex 0 of $G_t$ has to choose the self-loop for $\Omega(n)$ consecutive times in order to reach the vertex $n - 1$. Note that the hitting time of the simple random walk on the Sisyphus Wheel is unbounded.}
\end{figure}
A probability distribution \( \pi \) is a **stationary distribution** of \( P \) if \( \pi P = \pi \) holds. Let \( \lambda_*(P) \) denote the second largest eigenvalue in absolute value. Note that \( \lambda_*(P) < 1 \) if \( P \) is *aperiodic* or *lazy*. We say that \( P \) is *aperiodic* if for any \( v \in V \), \( \gcd\{t \geq 0 : P^t(v, v) > 0\} = 1 \) holds and \( P \) is *lazy* if \( P(v, v) \geq 1/2 \) holds for any \( v \in V \). Let \( t_{\text{rel}}(P) := (1 - \lambda_*(P))^{-1} \) be the relaxation time of \( P \). Let \( P = (P_t)_{t \geq 1} \) be a sequence of transition matrices. Suppose that all \( P_t \) have the same stationary distribution \( \pi \) and let \( t_{\text{hit}}(P) := \max_{u \in V} E \left[ \min\{t \geq 0 : X_t = w\} \mid X_0 = u \right] \) be the worst-case expected hitting time of \( P \). We sometimes identify \( P \) with the sequence \( (P_t)_{t \geq 1} \) of transition matrices with \( P_t = P \) for all \( t \geq 1 \). For example, \( t_{\text{hit}}(P) \) denotes the hitting time of \( P = (P_t)_{t \geq 1} \) with all \( P_t = P \). Let 
\[
t_{\text{REL}}(P) := \max_{t \geq 1} t_{\text{rel}}(P_t) \quad \text{and} \quad t_{\text{HIT}}(P) := \max_{t \geq 1} t_{\text{hit}}(P_t).
\]

### 1.1.1 Mixing time

Our first result concerns the mixing time. Specifically, for a sequence \( P = (P_t)_{t \geq 1} \) of transition matrices, a positive probability vector \( \pi \in (0, 1)^V \), and a parameter \( \epsilon > 0 \), we define the **uniform mixing time** \( t_{\text{mix}}^{(\infty, \pi)}(P, \epsilon) \) by

\[
t_{\text{mix}}^{(\infty, \pi)}(P, \epsilon) = \min \left\{ t \geq 0 : \max_{s \geq 0, u, v \in V} \left| P_{s+1} P_{s+2} \cdots P_{s+t}(u, v) \pi(v) \right| - 1 \right\} \leq \epsilon \text{ \quad (1)},
\]

and let \( t_{\text{mix}}^{(\infty, \pi)}(P) := t_{\text{mix}}^{(\infty, \pi)}(P, 1/2) \). The uniform mixing time \( t_{\text{mix}}^{(\infty)}(P, \epsilon) \) for a static Markov chain \( P \) can be seen as the mixing time using the \( \ell^\infty \)-norm metric and has been well studied (see, e.g., section 4.7 in [34]). The intuition behind our definition of \( t_{\text{mix}}^{(\infty, \pi)}(P, \epsilon) \) is that the walk mixes well at any consecutive \( t_{\text{mix}}^{(\infty, \pi)}(P, \epsilon) \) steps.

As a consequence of the previous work (e.g., (3.13) in [43]), we can easily obtain the following uniform mixing time bound.

**Proposition 1.1.** Let \( P = (P_t)_{t \geq 1} \) be a sequence of irreducible, aperiodic, and reversible transition matrices. Suppose that all \( P_t \) have the same stationary distribution \( \pi \). Then, for any \( \epsilon > 0 \), \( t_{\text{mix}}^{(\infty, \pi)}(P, \epsilon) = O \left( t_{\text{REL}}(P) \log(\pi_{\text{min}}^{-1} \epsilon^{-1}) \right) \), where \( \pi_{\text{min}} := \min_{v \in V} \pi(v) \).

Proposition 1.1 can be seen as an extension of the well-known mixing time bound \( t_{\text{mix}}^{(\infty)}(P, \epsilon) \leq \left[ t_{\text{rel}}(P) \log(\pi_{\text{min}}^{-1} \epsilon^{-1}) \right] \) for static \( P \) (theorem 12.4 in [34]) to the dynamic \( P \).

Our first result is the extension of the following well-known mixing time bound for static \( P \) (theorem 10.22 in [34]):

\[
t_{\text{mix}}^{(\infty)}(P, 1/4) \leq 4t_{\text{hit}}(P) + 1.
\]

**Theorem 1.2** (Main result 1). Let \( P = (P_t)_{t \geq 1} \) be a sequence of irreducible, reversible, and lazy transition matrices. Suppose that all \( P_t \) have the same stationary distribution \( \pi \).

Then, for any \( 0 < \epsilon < 1 \), \( t_{\text{mix}}^{(\infty, \pi)}(P, \epsilon) = O \left( t_{\text{HIT}}(P) + t_{\text{REL}}(P) \log \epsilon^{-1} \right) \).

Note that \( t_{\text{rel}}(P) \leq t_{\text{hit}}(P) \) holds for any irreducible, reversible and lazy \( P \) (lemma 4.24 in [2]) and thus Theorem 1.2 implies \( t_{\text{mix}}^{(\infty, \pi)}(P, 1/4) = O(t_{\text{HIT}}(P)) \). Compared to Proposition 1.1, Theorem 1.2 eliminates the dependency of \( \pi_{\text{min}}^{-1} \) in the mixing time bound at the cost of additional \( t_{\text{HIT}} \) term and the laziness assumption. Proposition 1.1 gives a better bound if all \( P_t \) have a small relaxation time (e.g., random walks on expanders). On the other hand, for \( P \) with \( t_{\text{REL}} \approx t_{\text{HIT}} \) (e.g., lazy simple random walk on dynamic cycles, on which both \( t_{\text{HIT}} \) and \( t_{\text{REL}} \) are \( \Theta(n^2) \)), Theorem 1.2 provides a better bound.

### 1.1.2 Hitting time and cover time

For the hitting and cover times, we recall an exponential lower bound on the Sisyphus Wheel (Figure 1), which implies the following.
**Proposition 1.3** ([5]). There is a sequence of irreducible, reversible, and lazy transition matrices \( P = (P_t)_{t \geq 1} \) of \( t_{\text{hit}}(P) = 2^{\Omega(n)} \).

Note that, the sequence \( P \) in Proposition 1.3 has a time-varying stationary distribution. Our second result concerns the hitting and cover times of multiple random walks according to \( P = (P_t)_{t \geq 1} \), where all \( P_t \) have the same stationary distribution. For \( k \in \mathbb{N} \), let \( t_{\text{hit}}^{(k)}(P) \), \( t_{\text{cov}}^{(k)}(P) \) denote the worst-case expected hitting and cover times of \( k \) independent random walks where each walker moves according to \( P \) (see Section 2 for details). For \( P = (P_t)_{t \geq 1}, \pi \in (0, 1)^V \) and a parameter \( \epsilon > 0 \), we define the separation time \( t_{\text{sep}}^{(\pi, \epsilon)}(P, \epsilon) \) by

\[
  t_{\text{sep}}^{(\pi, \epsilon)}(P, \epsilon) := \min \left\{ t \geq 0 : \max_{s \geq 0, u, v \in V} \left( 1 - \frac{(P_{s+1}P_{s+2} \cdots P_{s+t})(u, v)}{\pi(v)} \right) \leq \epsilon \right\},
\]

and let \( t_{\text{sep}}^{(\pi)}(P) = t_{\text{sep}}^{(\pi)}(P, 1/2) \). If \( P_t = P \) for all \( t \geq 1 \), the definition (2) coincides with well-known definition of the separation time for static \( P \) in the literature (see, e.g., section 4.3 in [2]). Note that \( t_{\text{sep}}^{(\pi)}(P) \leq t_{\text{sep}}^{(\infty, \epsilon)}(P) \) by definition.

**Theorem 1.4** (Main result 2). Let \( P = (P_t)_{t \geq 1} \) be a sequence of irreducible and reversible transition matrices. Suppose that all \( P_t \) have the same stationary distribution \( \pi \). Then, the following holds.

(i) \( t_{\text{hit}}^{(k)}(P) = O \left( t_{\text{sep}}^{(\pi)}(P) + \frac{\text{hrt}(P)}{k} \right) \) for any \( k \geq 1 \). In particular, \( t_{\text{hit}}(P) = O(t_{\text{HIT}}(P)) \) if \( P_t \) is lazy for all \( t \geq 1 \).

(ii) \( t_{\text{cov}}^{(k)}(P) = O \left( t_{\text{sep}}^{(\pi)}(P) + \frac{\text{hrt}(P) \log n}{k} \right) \) for any \( k \geq 1 \). In particular, \( t_{\text{cov}}(P) = O(t_{\text{HIT}}(P) \log n) \) if \( P_t \) is lazy for all \( t \geq 1 \).

Theorem 1.4 is the first result concerning multiple random walks on dynamic graphs. For \( k \)-independent random walks according to a static \( P \), the following bounds are known: \( t_{\text{hit}}^{(k)}(P) = O \left( t_{\text{sep}}^{(\pi)}(P) + \frac{\text{hrt}(P)}{k} \right) \) (theorem 8 in [21]) and \( t_{\text{cov}}^{(k)}(P) = O \left( t_{\text{sep}}^{(\pi)}(P) + \frac{\text{hrt}(P) \log n}{k} \right) \) (theorem 3.2 in [22]). Hence, Theorem 1.4 can be seen as a generalization of these previous bounds. Furthermore, for the case of \( k = 1 \), Theorem 1.4 improves various previous bounds of \( t_{\text{hit}}(P) \) and \( t_{\text{cov}}(P) \). See Section 1.2 for details.

### 1.1.3 Meeting time and coalescing time

Our third result concerns the *meeting time* \( t_{\text{meet}}(P) \) and the *coalescing time* \( t_{\text{coal}}(P) \) of random walks on dynamic graphs. Consider two independent random walks according to the same transition matrix sequence \( P \). The meeting time \( t_{\text{meet}}(P) \) is the expected time for the two walkers to meet starting from the worst initial positions. In the coalescing random walk, we consider \(|V|\) independent random walks according to \( P \) starting from \(|V|\) different initial positions. As soon as two or more walkers meet at the same position, the walkers are merged into a single walker. The coalescing time \( t_{\text{coal}}(P) \) is the expected time for the \(|V|\) walkers to merge into one. By definition, we have \( t_{\text{meet}}(P) \leq t_{\text{coal}}(P) \) in general.

For any irreducible, reversible, and lazy \( P \), it is known that \( t_{\text{coal}}(P) = O(t_{\text{hit}}(P)) \) (theorem 1.4 in [40]). Similarly to the hitting time, we observe an exponential gap between static and dynamic settings.

**Proposition 1.5.** There is a sequence of irreducible, reversible, and lazy transition matrices \( P = (P_t)_{t \geq 1} \) satisfying \( t_{\text{meet}}(P) = 2^{\Omega(n)} \).

Note that Proposition 1.5 also gives an exponential lower bound of the coalescing time since \( t_{\text{coal}}(P) \geq t_{\text{meet}}(P) \). The random walk we consider in Proposition 1.5 is the lazy simple random walk.
on the graph sequence presented by Olshevsky and Tsitsiklis [41], which has an exponential meeting time.

The following main result presents further evidence that a time-inhomogeneous coalescing walk with a common stationary distribution behaves almost identically to that on a static graph.

**Theorem 1.6** (Main result 3). Let $\mathcal{P} = (P_t)_{t \geq 1}$ be a sequence of irreducible, reversible, and lazy transition matrices. Suppose that all $P_t$ have the same stationary distribution $\pi$. Then, $t_{\text{coal}}(\mathcal{P}) = O(t_{\text{HIT}}(\mathcal{P}))$.

Theorem 1.6 is the first result for the coalescing time on dynamic graphs. Theorem 1.6 generalizes the aforementioned bound of $t_{\text{coal}}(P) = O(t_{\text{hit}}(P))$ for static $P$ (Theorem 1.4 in [40]). Furthermore, Theorem 1.6 plays a key role to bound the consensus time of the pull voting on dynamic graphs (Section 1.3).

### 1.2 Examples

Our general results Theorems 1.2, 1.4, and 1.6 yield new and improved bounds for mixing, hitting, cover, and coalescing times for various concrete random walks. These consequences are summarized in Table 1. Throughout this section, unless otherwise stated, we assume that a graph $G$ is connected. For a graph $G$ and a vertex $v \in V(G)$, let $N(G, v)$ be the set of neighboring vertices of $v$ (excluding $v$) and deg($G, v$) = $|N(G, v)|$ be the degree of $v$.

#### 1.2.1 Lazy simple random walk

The transition matrix $P_{\text{LS}}(G)$ of the lazy simple random walk$^1$ on a graph $G$ is defined by

\[
P_{\text{LS}}(G)(u, v) = \begin{cases} 
1/2 \deg(G, u) & \text{if } v \in N(G, u), \\
1/2 & \text{if } u = v, \\
0 & \text{otherwise}.
\end{cases}
\]

It is well known that $t_{\text{hit}}(P_{\text{LS}}(G)) = O(n^{3})$ for any $G$ [3], $t_{\text{hit}}(P_{\text{LS}}(G)) = O(n^{2})$ for any regular $G$ [30], and $t_{\text{hit}}(P_{\text{LS}}(G)) = O(n)$ for any regular expander$^2$ $G$ [2].

---

$^1$The laziness does not change the order of the hitting and cover times. On the other hand, on connected bipartite graphs, the mixing and coalescing times of the simple random walk are unbounded, while these are bounded for the lazy simple random walk. Hence, we assume laziness in many cases.

$^2$A graph is expander if $t_{\text{rel}}(P_{\text{LS}}(G)) \leq C$ for some constant $C > 0$. 

---

| Type           | Examples | $t_{\text{mix}}(n, \lambda)$ | $t_{\text{hit}}$ | $t_{\text{cov}}$ |
|---------------|----------|-------------------------------|-----------------|-----------------|
| General       |          | $O(t_{\text{IT}} \log t_{\text{IT}})$ | Theorem 1.2     | Theorem 1.4     |
| LS RW         |          | $O(n^{3})$ | Theorem 1.4 | $O(t_{\text{IT}} \log n)$ | Theorem 1.4 |
| Regular       |          | $O(n^{2})$ | Theorem 1.4 | $O(t_{\text{IT}} \log n)$ | Theorem 1.4 |
| $d_{\text{max}}$-lazy | Any graph | $O(n^{3})$ | Theorem 1.4 | $O(t_{\text{IT}} \log n)$ | Theorem 1.4 |
| LMW           | Any graph | $O(n^{2})$ | Theorem 1.4 | $O(t_{\text{IT}} \log n)$ | Theorem 1.4 |

Notes: The general results from Proposition 1.1 and Theorems 1.2 and 1.4 are given in General row. The examples include bounds for a lazy simple random walk (LSRW) on graphs with a time-homogeneous degree distribution, regular graphs, a $d_{\text{max}}$-lazy walk, and a lazy Metropolis walk (LMW), where we assume that the underlying graph is always connected. Gray cells mean new or improved bounds.
Let $G = (G_t)_{t \geq 1}$ be a sequence of graphs with a time-homogeneous degree distribution $(d(v))_{v \in V}$, that is, $\text{deg}(G_t, v) = d(v)$ holds for all $t \geq 1$ and $v \in V$. Then, $P = (P_{LS}(G_t))_{t \geq 1}$ has the common stationary distribution $\pi = \left( \frac{d(v)}{\sum_{v \in V} d(u)} \right)_{v \in V}$. Hence, we can apply Theorems 1.2, 1.4, and 1.6. In particular, Theorem 1.4 implies $t_{\text{hit}}(P) = O(t_{\text{HTT}}(P)) = O(n^3)$. This improves the $O(n^3 \log n)$ bound (Main Result 1(3)) of [46]. Another interesting example is the sequence $G = (G_t)_{t \geq 1}$ of regular expander graphs. Let $P = (P_{LS}(G_t))_{t \geq 1}$. For $k = O(n)$, the cover time of $k$ independent lazy simple random walks satisfies $t_{\text{cov}}^{(k)}(P) = O\left( \frac{n \log n}{k} \right)$ from Proposition 1.1 and Theorem 1.4. This bound is tight since $t_{\text{cov}}^{(k)}(P_{LS}(G)) = \Omega\left( \frac{n \log n}{k} \right)$ holds for any (static) $G$ and $k = O(n \log n)$ [42].

1.2.2 $d_{\text{max}}$-lazy walk

Let $d_{\text{max}} = d_{\text{max}}(G) := \max_{v \in V(G)} \text{deg}(v)$ denote the maximum degree of $G$. The transition matrix $P_{DM}(G)$ of the $d_{\text{max}}$-lazy walk on a graph $G$ is defined by $P_{DM}(u, v) = \frac{1}{2d_{\text{max}}}$ if $\{u, v\} \in E(G)$, $P_{DM}(u, u) = 1 - \frac{\text{deg}(G, u)}{2d_{\text{max}}}$, and $P_{DM}(G)(u, v) = 0$ otherwise. It is known that $t_{\text{hit}}(P_{DM}(G)) = O(n^3)$ holds for any $G$ ([5, 19]).

Note that $P_{DM}(G)$ has the uniform stationary distribution for any $G$ since $P_{DM}(G)$ is symmetric. Hence, for any sequence of connected graphs $G = (G_t)_{t \geq 1}$ and $P = (P_{DM}(G_t))_{t \geq 1}$, we can apply Theorems 1.2, 1.4, and 1.6. For example, $t_{\text{hit}}(P) = O(t_{\text{HTT}}(P)) = O(n^3)$ holds from Theorem 1.4. This improves the previous $O(n^3 \log n)$ bound in [19].

1.2.3 Metropolis walk

We saw in the previous paragraph that the $d_{\text{max}}$-lazy walk has a polynomial cover time for any dynamic graph. However, as mentioned in [5], the $d_{\text{max}}$-lazy random walk requires knowledge of the maximum degree, which is a global information of $G_t$ at each $t \geq 1$. Regarding this issue, we consider the lazy Metropolis walk of Nonaka, Ono, Sadakane, and Yamashita [38]. The transition matrix $P_{LM}(G)$ of the lazy Metropolis walk on a graph $G$ is defined by

$$P_{LM}(G)(u, v) = \begin{cases} \frac{1}{2 \max\{\text{deg}(G, u), \text{deg}(G, v)\}} & \text{if } \{u, v\} \in E(G), \\ 1 - \frac{1}{2 \max\{\text{deg}(G, u), \text{deg}(G, v)\}} & \text{if } u = v, \\ 0 & \text{otherwise.} \end{cases}$$

Note that a lazy Metropolis walk uses local degree information around the walker. For any $G$, Nonaka et al. [38] showed $t_{\text{hit}}(P_{LM}(G)) = O(n^3)$.

Since the transition matrix of the lazy Metropolis walk is symmetric, the stationary distribution is uniform for any underlying graph. Hence, we can apply Theorems 1.2, 1.4, and 1.6. Interestingly, our bounds for Metropolis walks in Table 1 are tight up to a constant factor: On the (static) cycle graph, the lazy Metropolis walk has $O(n^2)$ mixing time and $O(n^2)$ hitting time [2]. On the glitter star graph of [38], the lazy Metropolis walk has an $\Omega(n^2 \log n)$ cover time.

1.2.4 Metropolis walks on edge-Markovian graphs

In this paper, we also study random walks on the edge-Markovian graph $(G_t)_{t \geq 1}$ defined as follows: Let $p, q \in [0, 1]$ be parameters and $G_1$ be an arbitrary fixed graph. The graph $G_{t+1}$ is obtained by adding
each $e \in \binom{V}{2} \setminus E(G_i)$ independently with probability $p$ and removing each $e \in E(G_i)$ independently with probability $q$.

The model of edge-Markovian graph was introduced by Clementi, Macci, Monti, Pasquale, and Silvestri [12] as a wide generalization of time-independent dynamic random graphs. Since then, several properties including the flooding [7, 12], rumor spreading [13], and mixing time [11] have been investigated. In this paper, we focus on the Metropolis walk on the edge-Markovian graph $(G_i)_{i \geq 1}$ and obtain the following result. See Section 7 for the proof.

**Theorem 1.7.** Let $c \geq 1$ be arbitrary. Consider $G(n, p, q) = (G_i)_{i \geq 1}$ the edge-Markovian graph satisfying $\frac{p}{p+q} \geq 32(c+1)\log \frac{n}{q}$ and $0 < p + q \leq 1$. Let $P = (P_{\text{LM}}(G_i))_{i \geq 1}$. Then, for any $k \geq 1$, $G(n, p, q)$ satisfies the following with probability $1 - n^{-\Omega(1)}$: 

$$t_{\text{hit}}(P) = O\left(\frac{\max\{1/p\}}{p+q} \log n + \frac{n}{k}\right), \quad t_{\text{cov}}(P) = O\left(\frac{\max\{1/p\}}{p+q} \log n + \frac{n \log n}{k}\right), \quad \text{and} \quad t_{\text{coal}}(P) = O\left(\frac{\max\{1/p\}}{p+q} + n\right).$$

1.3 Pull voting on dynamic graphs

In the (weighted) pull voting, we consider an $n$-vertex graph $G = (V, E)$ where each vertex $v \in V$ holds an opinion $\sigma_v \in \Sigma$ for a finite set $\Sigma \subseteq \{0, \ldots, n-1\}$ of possible opinions. Let $P \in [0, 1]^{V \times V}$ be a transition matrix. At each discrete time step, each vertex $u$ chooses a random neighbor according to the distribution $P(u, \cdot)$ and then updates its opinion with the neighbor’s opinion. The goal of the protocol is to reach a consensus in which every vertex holds the same opinion. The consensus time is the time of the process to reach consensus.

We consider the pull voting on dynamic graphs. Specifically, let $P = (P_t)_{t \geq 1}$ be a transition matrix sequence. In the pull voting according to $P$, at the $t$th round, vertices perform the one-round pull voting according to $P_t$. We denote by $t_{\text{cons}}(P)$ the consensus time of the pull voting according to $P$ and consider $t_{\text{cons}}(P) := E[t_{\text{cons}}(P)]$. Combining Theorem 1.6 and the idea of well-known duality between the coalescing random walk and pull voting [26] with some additional argument, we prove the following.

**Theorem 1.8** (Consensus time). Let $P = (P_t)_{t \geq 1}$ be a sequence of irreducible, lazy, and reversible transition matrices. Suppose that all $P_t$ have the same stationary distribution $\pi$. Then, $t_{\text{cons}}(P) \leq O(t_{\text{HIT}}(P))$.

On the other hand, the consensus time can be exponential in general.

**Proposition 1.9.** There is a sequence $(G_i)_{i \geq 1}$ of $n$-vertex connected graphs such that $t_{\text{cons}}(P) = 2^{\Omega(n)}$ for $P = (P_{\text{LS}}(G_i))_{i \geq 1}$.

Another important question concerning pull voting is the probability that the process finally agrees with a specific opinion $\sigma \in \Sigma$. Using the voting martingale argument (e.g., [16]), we obtain the following result.

**Proposition 1.10** (Winning probability). Let $P = (P_t)_{t \geq 1}$ be a sequence of irreducible, lazy, and reversible transition matrices. Suppose that all $P_t$ have the same stationary distribution $\pi$. Consider the pull voting over opinion set $\Sigma$ according to $P$. Then, the process finally agrees with $\sigma \in \Sigma$ with probability $\sum_{v \in V_\sigma} \pi(v)$, where $V_\sigma$ is the set of vertices initially holding opinion $\sigma$. 

1.4 Proof overview

We give an overview of the proof of the main theorems (Theorems 1.2, 1.4, and 1.6). Throughout this section, we assume that \( P = (P_t)_{t \geq 1} \) is a sequence of irreducible and reversible transition matrices, where all \( P_t \) have the same stationary distribution \( \pi \in (0, 1)^V \). For \( f, g \in \mathbb{R}^V \), we define the inner product \( \langle u, v \rangle_\pi = \sum_{v \in V} f(v)g(v)\pi(v) \) and the induced norm \( \|f\|_\pi = \sqrt{\sum_{v \in V} f(v)^2 \pi(v)} \). Define \( f^L \in \mathbb{R}^V \) by \( (f^L)(v) = (f(v))_\pi \). Let \( \mathcal{L} \) denote the \(|V|\)-dimensional all-one vector.

For a vertex \( w \in V \), let \( D_w \in \{0, 1\}^{V \times V} \) be the diagonal matrix defined by \( D_w(v, v) = 1 \) if \( v \neq w \) and \( D_w(w, w) = 0 \). We are interested in the substochastic matrix \( D_wPD_w \). Note that \( D_wPD_w \) is the matrix obtained by replacing elements of \( P(w, \cdot) \) and \( P(\cdot, w) \) by 0. It is known that \( \rho(D_wPD_w) \leq 1 - \frac{1}{\text{hit}(P)} \), where \( \rho(M) \) is the spectral radius of \( M \) (see Lemma A.1 or section 3.6.5 in [2]).

1.4.1 Mixing time (Section 3)

Let \( \mu_0 \in [0, 1]^V \) be an initial distribution and \( \mu_T = \mu_0 \prod_{t=1}^T P_t \) be the distribution of \( X_T \). Let \( d(\cdot, \cdot)(\mu_T) := \|\mu_T - \frac{1}{\pi}\|_{\pi} \), that is, the \( \ell^2 \)-distance between \( \mu_T/\pi \) and \( \frac{1}{\pi} \). It is known that \( \max_{v \in V} \frac{\mu_0(v)}{\pi(v)} - 1 \) can be bounded in terms of \( d(\cdot, \cdot)(\mu_T) \) (see, e.g., [34, 43] or (8)). Henceforth, we focus on bounding \( d(\cdot, \cdot)(\mu_T) \).

The main part of the proof of Theorem 1.2 is to prove the following inequality: For any \( t \geq 0 \) and lazy \( P = (P_t)_{t \geq 1} \), \( d(\cdot, \cdot)(\mu_{t+1})^2 \leq d(\cdot, \cdot)(\mu_t)^2 \left( 1 - \frac{d(\cdot, \cdot)(\mu_t)^2}{\text{hit}(P)} \right) \) holds. Applying this inequality repeatedly, we obtain \( d_{\text{mix}}(P) = O(\text{hit}(P)) \) (see Lemma C.4 or [46] for detail).

The aforementioned inequality comes from a variant of Mihail’s identity (Lemma A.4), which claims that \( d(\cdot, \cdot)(\mu)^2 \leq d(\cdot, \cdot)(\mu)^2 - \mathcal{E}_{P, \pi}(\mu/\pi) \) for any reversible and lazy \( P \) and any probability vector \( \mu \). Here, \( \mathcal{E}_{P, \pi}(f) = \frac{1}{2} \sum_{u, v \in V} \pi(u)P(u, v)f(u) - \langle f, \pi \rangle \) is the Dirichlet form of \( P \) and \( \pi \). In [46], the authors consider a lazy simple random walk and give lower bounds on \( \mathcal{E}_{P, \pi}(\mu/\pi) \) in terms of some graph parameters, for example, \( \mathcal{E}_{P, \pi}(\mu/\pi) \geq d(\cdot, \cdot)(\mu)^4/n^2 \) for regular graphs. This lower bound means that \( d(\cdot, \cdot)(\mu_{t+1})^2 \leq d(\cdot, \cdot)(\mu_t)^2 \left( 1 - \frac{d(\cdot, \cdot)(\mu_t)^2}{n^2} \right) \). Our technical contribution is to generalize the previous lower bound on \( \mathcal{E}_{P, \pi}(\mu/\pi) \) for any reversible and lazy \( P \) in terms of the hitting time (Lemma 3.2): \( \mathcal{E}_{P, \pi}(\mu/\pi) \geq d(\cdot, \cdot)(\mu)^4/\text{hit}(P) \) holds. This implies the desired inequality, \( d(\cdot, \cdot)(\mu_{t+1})^2 \leq d(\cdot, \cdot)(\mu_t)^2 \left( 1 - \frac{d(\cdot, \cdot)(\mu_t)^2}{\text{hit}(P)} \right) \).

The proof of the key lemma (Lemma 3.2) consists of four steps. First, observe that \( \mathcal{E}_{P, \pi}(f) = \mathcal{E}_{P, \pi}(f - \mathcal{E}_{P, \pi}(f)) \) for any \( f \in \mathbb{R}^V \) and \( c \in \mathbb{R} \). Hence, \( \mathcal{E}_{P, \pi}(f) = \mathcal{E}_{P, \pi}(f_{\text{max}} - f) \) holds for \( f_{\text{max}} := \max_{v \in V} f(v) \). Second, let \( g := f_{\text{max}} \mathbb{1} - f \). Note that \( g(w) = 0 \) for \( w \in V \) satisfying \( f(w) = f_{\text{max}} \). Therefore, for any \( u, v \in V \), we have \( g(v)P(v, u)g(u) = g(v)(D_wPD_w)g(u) + g(u)D_wPD_wg(v) \pi \) holds. Hence, we have \( \mathcal{E}_{P, \pi}(g) = \mathcal{E}_{P, \pi}(g) - \langle g, P\pi \rangle \pi = \mathcal{E}_{D_wPD_w, \pi}(g) \). Third, from the known inequality \( \langle g, D_wPD_wg \pi \rangle \pi \leq \rho(D_wPD_w)\|g\|_{\pi}^2 \) for the spectral radius (Lemma A.2) and \( \rho(D_wPD_w) \leq 1 - \frac{1}{\text{hit}(P)} \) (Lemma A.1), we have \( \mathcal{E}_{P, \pi}(g) = \mathcal{E}_{D_wPD_w, \pi}(g) \geq \|g\|_{\pi}^2 \). Finally, from a carefully calculation, it is not difficult to see that \( \|g\|_{\pi}^2 \geq d(\cdot, \cdot)(\mu)^4 \) for \( f = \mu/\pi \).

1.4.2 Hitting and cover times (Section 4)

Let \( (X_t)_{t \geq 0} \) be the random walk according to \( P = (P_t)_{t \geq 1} \). To obtain an upper bound on the hitting time, it suffices to bound the probability \( \Pr \left[ X_T \neq w \right] \) for any fixed vertex \( w \).

Lemma 1.11. Consider a random walk \( (X_t)_{t \geq 0} \) according to a sequence \( (P_t)_{t \geq 1} \) of (not necessarily be lazy) reversible transition matrices such that all \( P_t \) have the same stationary
distribution \( \pi \) and \( X_0 \) is sampled from \( \pi \). Then, for any vertex \( w \in V \) and \( T \geq 0 \),

\[
\Pr \left[ \bigwedge_{t=0}^{T} \{ X_t \neq w \} \right] \leq \prod_{t=1}^{T} \left( 1 - \frac{1}{\ell_{\text{hit}}(P_t)} \right) \leq \exp \left( - \frac{T}{\ell_{\text{HIT}}(P)} \right).
\]

We will prove Lemma 1.11 in Section 4.1. If the walker starts according to the stationary distribution, Lemma 1.11 immediately gives the bounds on the hitting and cover times since the event \( \bigwedge_{t=0}^{T} \{ X_t \neq w \} \) means that the walk does not hit \( w \) until time step \( T \). By the union bound, the probability that the cover time is larger than \( 2\ell_{\text{HIT}}(P) \log n \) is upper bounded by \( n \left( 1 - \frac{1}{\ell_{\text{HIT}}(P)} \right)^{2\ell_{\text{HIT}}(P) \log n} \leq 1/n \).

This argument can be easily extended to the case of \( k \) independent random walks starting from positions according to the stationary distribution; the expected cover time is bounded by \( \sum_{t=0}^{\infty} \left( 1 - \frac{1}{\ell_{\text{HIT}}(P)} \right)^{kt} = O(\ell_{\text{HIT}}(P)/k) \).

Lemma 1.11 and the bounds on the separation distance (Theorem 1.2) enable us to obtain upper bounds on the hitting and cover times of \( k \) walkers from the worst initial positions. Here is the proof sketch. Let \((X_t(1), \ldots, X_t(k))_{t \geq 0} \) be \( k \) independent walks starting from the worst initial positions. By definition of the separation distance (2), the probability that the walker \( i \in [k] \) is on \( u \in V \) at time \( t_{\text{sep}}^{(i)}(P) \) is \( P_{[1, t_{\text{sep}}^{(i)}(P)]}(X_t(1), u) \geq \frac{1}{2} \pi(u) \). Therefore, in expectation, half of the walkers are distributed according to \( \pi \) after \( t_{\text{sep}}^{(i)}(P) \) steps. From Lemma 1.11, the probability that all of such walks do not hit a specific vertex \( w \) within \( T \) steps is at most roughly \( (1 - 1/\ell_{\text{HIT}}(P))^{Tk/2} \). Taking \( T = O(\ell_{\text{HIT}}(P)/k) \), this probability can be bounded by some constant probability. This implies \( \ell_{\text{hit}}(P) = O(\ell_{\text{sep}}(P)/t_{\text{HIT}}(P)/k) \) (see Lemma B.1 for details). The proof for the cover time proceeds in a similar way. By the union bound over \( w \in V \), the probability that there is a vertex \( w \) such that all the \( \pi \)-distributed walkers do not hit \( w \) within \( T \) steps is at most roughly \( n(1 - 1/\ell_{\text{HIT}}(P))^{Tk/2} \).

The proof of Lemma 1.11 goes as follows. For a fixed \( w \in V \), consider the sequence of substochastic matrices \((D_wP_tD_w)_{t \geq 1}\). By definition of \( D_w \) and the Cauchy–Schwarz inequality, we have

\[
\Pr \left[ \bigwedge_{t=0}^{T} \{ X_t \neq w \} \right] = \sum_{x, y \in V} \pi(x) \left( \prod_{t=1}^{T} (D_wP_tD_w) \right) (x, y) \leq \left\| \left( \prod_{t=1}^{T} (D_wP_tD_w) \right) \right\|_{2, \pi}.
\]

Furthermore, from the reversibility of \( P_t \), all \( D_wP_tD_w \) are reversible. Thus, we can apply a variant of the Courant–Fischer theorem for the \( \pi \)-inner product (Lemma A.2) repeatedly and obtain

\[
\left\| \left( \prod_{t=1}^{T} (D_wP_tD_w) \right) \right\|_{2, \pi} \leq \prod_{t=1}^{T} \rho(D_wP_tD_w). \]

Since \( \rho(D_wPD_w) \leq 1 - 1/\ell_{\text{hit}}(P) \) (Lemma A.1), we obtain Lemma 1.11.

### 1.4.3 Meeting and coalescing times (Section 5)

To give an upper bound of the coalescing time, we recall the powerful Meeting Time Lemma given by Oliveira [39] (for continuous-time walks) and Oliveira and Peres [40] (for discrete-time lazy walks) in the context of time-homogeneous random walks. Our key observation is that Meeting Time Lemma indeed holds for time-inhomogeneous random walks with the same stationary distribution.

**Lemma 1.12.** Consider a random walk \((X_t)_{t \geq 0}\) according to a sequence \((P_t)_{t \geq 1}\) of lazy reversible transition matrices such that all \( P_t \) have the same stationary distribution \( \pi \) and \( X_0 \) is sampled from \( \pi \). Then, for any sequence of vertices \((w_t)_{t \geq 0}\) and \( T \geq 0 \),

\[
\Pr \left[ \bigwedge_{t=0}^{T} \{ X_t \neq w_t \} \right] \leq \prod_{t=1}^{T} \left( 1 - \frac{1}{\ell_{\text{hit}}(P_t)} \right) \leq \exp \left( - \frac{T}{\ell_{\text{HIT}}(P)} \right).
\]
In comparison to Lemma 1.11, Lemma 1.12 deals with the hitting probability of a sequence \((w_t)\) of vertices at the cost of the laziness assumption, while Lemma 1.11 consider the special case of hitting a fixed vertex \(w\) without the laziness assumption. We will prove Lemma 1.12 in Section 4.1, which essentially relies on the laziness assumption.

Consider two independent random walks \((X_t(1))_{t\geq 0}\) and \((X_t(2))_{t\geq 0}\) such that the initial positions \(X_0(1)\) and \(X_0(2)\) are sampled according to the stationary distribution independently. Since 
\[
\Pr \left[ \bigwedge_{t=0}^{T} (X_t(1) \neq X_t(2)) \right] \leq \Pr \left[ \bigwedge_{t=0}^{T} (X_t(1) \neq w_t) \right]
\]
for some sequence of vertices \((w_t)_{t\geq 0}\), we have 
\[
\Pr \left[ \bigwedge_{t=0}^{T} (X_t(1) \neq X_t(2)) \right] \leq \exp \left( -\frac{T}{t_{\text{hit}}(P)} \right)
\]
by Lemma 1.11.

The proof of our \(O(t_{\text{HIT}}(P))\) coalescing time bound essentially consists of three parts. In the first part, we bound the probability that coalescing time is larger than \(Ct_{\text{HIT}}(P)\) for a suitable constant \(C > 0\) in terms of the probability of a suitable event of independent multiple random walks (Lemma 5.1). This part proceeds similarly to the argument of static setting in [40] by coupling arguments. In the second part, similar to the arguments of multiple random walks, we show that the positions of an appropriate group of walkers at \(t = t_{\text{sep}}(P)\) are according to the stationary distribution with constant probability. In the third part, we apply Lemma 1.12 to such group of walkers (Lemma 5.3). Since Lemma 1.12 works in the dynamic setting, we can complete this part in a similar way to the static setting [40].

In the proof of Lemma 1.12, the laziness assumption of walkers is essential as well as that of [40]. Since \(P\) is positive semidefinite by the laziness, we rewrite \(D_xPD_x = D_x\sqrt{P}\sqrt{PD_x}\), where \(\sqrt{P}\) is the square root of \(P\). Furthermore, the reversibility of \(P\) implies that \(D_x\sqrt{P}\) is the adjoint of \(\sqrt{PD_x}\) (that is, \(\pi(u)(D_x\sqrt{P})(u, v) = \pi(v)(\sqrt{P}D_x)(v, u)\) for any \(u, v \in V\)). Hence, we have 
\[
\|D_xPD_xf\|_{2,\pi} \leq \sqrt{\rho(D_xPD_x)\rho(D_xPD_x)}\|f\|_{2,\pi} \leq \left( 1 - \frac{1}{t_{\text{hit}}(P)} \right) \|f\|_{2,\pi}
\]
for any vector \(f\) (Lemmas A.1 and A.5). Combining this inequality and similar arguments in the proof of Lemma 1.11, we obtain 
\[
\Pr \left[ \bigwedge_{t=0}^{T} (X_t \neq w_t) \right] \leq \left\| \left( \prod_{i=1}^{T} (D_{w_{t-1}}P_{t}D_{w_{t}}) \right) \right\|_{2,\pi} \leq \prod_{i=1}^{T} \left( 1 - \frac{1}{t_{\text{hit}}(P)} \right).
\]

1.5 Related work

1.5.1 Random walk on a static graph

Consider a simple random walk on a connected graph \(G\) with \(n\) vertices and \(m\) edges. Aleliunas, Karp, Lipton, Lovász, and Rackoff [3] showed that the cover time is at most \(2m(n - 1)\). Kahn, Linial, Nisan, and Saks [30] showed that the cover time is at most \(16mn/d_{\text{min}}\). While the hitting time is at least \((1/2)m/d_{\text{min}}\) (see corollary 3.3 of Lovász [35]). Brightwell and Winkler [9] proved that the lollipop graph has a \((4/27 + o(1))n^3\) hitting time, while Feige [23] gave proved that the cover time is upper bounded by \((4/27)n^3 + O(n^{5/2})\) for any graph. In addition to the trivial relation of \(t_{\text{hit}} \leq t_{\text{cov}}\), it is known that \(t_{\text{cov}} \leq t_{\text{hit}}\log n\) holds for any \(G\) (see Matthews [36]).

The cover time \(t_{\text{cov}}(k)\) of \(k\) independent simple random walks has been studied in [4, 10, 21, 22, 42]. If \(k\) walkers start from the stationary distribution, Broder, Karlin, Raghavan, and Upfal [10] showed that the cover time is at most \(O \left( \left( \frac{m}{k} \right)^2 \log^2 n \right)\). More recently, Rivera, Sauerwald, and Sylvester [42] proved an improved bound of \(O \left( \left( \frac{m}{kd_{\text{min}}} \right)^2 \log^2 n \right)\). From the worst starting positions of \(k\) walkers, Elsässer and Sauerwald [22] showed that 
\[
t_{\text{cov}}(k) = O \left( t_{\text{mix}} + \frac{t_{\text{hit}} \log n}{k} \right)
\]
for \(k \leq n\).

It is well known that local degree information gives surprising power to random walks. For example, the \(\beta\)-random walk proposed by Ikeda, Kubo, Okumoto, and Yamashita [28, 29] and the Metropolis walk proposed by Nonaka, Ono, Sadakane, and Yamashita ([38], the definition is (3)) have
the \(O(n^2)\) hitting time and \(O(n^2 \log n)\) cover time for any \(G\). These bounds improve the worst-case \(\Omega(n^3)\) hitting time of the simple random walk (on the lollipop graph). Recently, David and Feige [18] showed the \(O(n^2)\) cover time for the minimum-degree random walk proposed by Abdullah, Cooper, and Draief [1]. This is best possible since any random walk on the path has \(\Omega(n^2)\) cover time [29]. It is easy to see that the \(\beta\)-random walk and the minimum-degree random walk have exponential hitting times on the Sisyphus Wheel.

The meeting and coalescing times have been well studied in the context of distributed computation such as leader election and consensus protocols [26]. Consider the simple random walk on a connected and non-bipartite \(G\). Tetali and Winkler showed that the meeting time is at most \((16/27 + o(1))n^3\). Hassin and Peleg [26] showed \(t_{coal} \leq t_{meet} \log n\), while \(t_{meet} \leq t_{coal}\) is trivial. Recent works on the meeting and coalescing times consider the lazy simple random walk on a connected graph \(G\) [8, 17, 31, 40]. For example, Kanade, Mallmann-Trenn, and Sauerwald [31] showed \(t_{coal} = O(t_{meet}(1 + \sqrt{t_{mix} \log n}))\). Oliveira and Peres [40] proved \(t_{coal} = O(t_{hit})\).

1.5.2 Random walk on dynamic graphs
Avin, Koucký and Lotker [5] presented the Sisyphus Wheel on which the hitting time is \(2^{\Omega(n)}\) for the lazy simple random walk (Figure 1). To avoid the exponential hitting time, Avin et al. [5] considered the \(d_{max}\)-lazy random walk on \(G\) and showed that the cover time of this random walk on any sequence of connected graphs is \(O(n^5 \log^2 n)\). They also showed that the mixing time of the walk is \(O(n^3 \log n)\). Denisyuk and Rodrigues [19] improved the cover time bound to \(O(n^3 \log n)\). Sauerwald and Zanetti [46] considered a lazy simple random walk on a sequence \((G_t)_{t \geq 1}\) of graphs such that all \(G_t\) have a common degree distribution. They showed that the \(O(n/\pi_{min})\) mixing time and \(O(n \log n)/\pi_{min}\) hitting time. Moreover, if all \(G_t\) are \(d\)-regular, then the hitting time is \(O(n^2)\). This bound matches that of static regular graphs [30].

The study of time-inhomogeneous Markov chains has applications in a wide range of fields, including consensus algorithms [41] and cryptography [37]. In an early paper, Griffeath [25] studied the ergodic theorem of time-inhomogeneous Markov chains. Saloff-Conste and Zúñiga [44, 45] considered the merging time for time-inhomogeneous Markov chains in terms of the stability of the stationary distributions. In another paper [43], they obtained an upper bound on the mixing time for time-inhomogeneous Markov chains under the assumption that the chains are irreducible and have a common stationary distribution.

Cai, Sauerwald, and Zaneti [11] considered the lazy simple random walk on a sequence of edge-Markovian random graphs. They introduced the notion of mixing time on this sequence (note that the stationary distribution changes over time) and obtained several mixing time bounds. Lampropoulou, Martin, and Spirakis [33] studied the cover time of the simple random walk on a variant of edge-Markovian random graphs.

1.5.3 Pull voting
The pull voting according to \(P_{LS}(G)\) for a static graph \(G\) has been intensively studied [17, 26] in the literature of distributed computing and stochastic process. It is widely known that the consensus time and the coalescing time are equal. Therefore, bounds for coalescing time yields bounds for consensus time. In particular, the consensus time on any (static) connected nonbipartite graphs is \(O(n^3 \log n)\) from Hassin and Peleg [26].

Berenbrink, Giakkoupis, Kermarrec, and Mallmann-trenn [8] studied the pull voting according to \((P_{LS}(G_t))_{t \geq 1}\) for a sequence \((G_t)_{t \geq 1}\) of graphs constructed by an adaptive adversary. That is, for every \(t\),
the graph $G_t$ can depend on the history of opinion configurations. Under the assumption that all $G_t$ must have the same degree distribution, they obtained an upper bound on $\tau_{\text{cons}}$ in terms of the conductance of $G_t$ for the binary opinion setting (i.e., $|\Sigma| = 2$).

\section{Notations and Definitions}

This section defines the mixing, hitting, cover, meeting, and coalescing times formally. For $b \geq a \geq 1$ and a sequence $(P_t)_{t \geq 1}$ of transition matrices, let $P_{[a,b]} := P_a P_{a+1} \cdots P_b$. For $\pi \in [0, 1]^V$, let $\pi_{\min} := \min_{v \in V} \pi(v)$.

For $p \geq 1$ and probability vectors $\mu \in [0, 1]^V$ and $\pi \in (0, 1]^V$, let

$$d^{(p,\pi)}(\mu) := \begin{cases} \left\| \frac{\mu}{\pi} - 1 \right\|_{p,\pi} = \left( \sum_{v \in V} \pi(v) \left| \frac{\mu(v)}{\pi(v)} - 1 \right|^p \right)^{1/p} & \text{if } 1 \leq p < \infty, \\
\max_{v \in V} \left| \frac{\mu(v)}{\pi(v)} - 1 \right| & \text{if } p = \infty \end{cases}$$

be the $\ell^p$-distance between $\mu/\pi$ and 1. It is known that $d^{(p,\pi)}(\mu) \leq d^{(p+1,\pi)}(\mu)$ holds for any $p \geq 1$ (see, e.g., section 4.7 in [34]). For example, $\sum_{v \in V} |\mu(v) - \pi(v)| = d^{(1,\pi)}(\mu) \leq d^{(2,\pi)}(\mu) \leq d^{(\infty,\pi)}(\mu)$. For $P = (P_t)_{t \geq 1}$, a probability vector $\pi \in (0, 1]^V$, and $\epsilon > 0$, we define the $\ell^p$-mixing time as

$$t^{(p,\pi)}_{\text{mix}}(P, \epsilon) := \min \left\{ t \geq 0 : \max_{s,t \geq 1, v \in V} d^{(p,\pi)}(P_{[s+1,t]}(v, \cdot)) \leq \epsilon \right\}.$$

Write $t^{(p,\pi)}_{\text{mix}}(P) := t^{(p,\pi)}_{\text{mix}}(P, 1/2)$.

Consider $k$ independent random walks $(X_t(1))_{t \geq 0}, \ldots, (X_t(k))_{t \geq 0}$, where each walker moves according to $P = (P_t)_{t \geq 1}$. Let $\tau^{(k)}_{\text{hit}}(P, w)$ (for $w \in V$) and $\tau^{(k)}_{\text{cov}}(P)$ be the random variables denoting hitting and cover times of the $k$ random walks, respectively. Formally,

$$\tau^{(k)}_{\text{hit}}(P, w) = \inf \left\{ t \geq 0 : \bigcup_{i \in [k], 0 \leq s \leq t} \{X_t(i)\} \ni w \right\}, \quad (4)$$

$$\tau^{(k)}_{\text{cov}}(P) = \inf \left\{ t \geq 0 : \bigcup_{i \in [k], 0 \leq s \leq t} \{X_t(i)\} = V \right\}. \quad (5)$$

Let $t^{(1)}_{\text{hit}}(P) := \max_{x \in V, w \in V} E \left[ \tau^{(1)}_{\text{hit}}(P, w) \big| X_0 = x \right]$ for the expected hitting time of $k$ random walks. Here, $X_t = (X_t(1), \ldots, X_t(k)) \in V^k$ is a vector-valued random variable. Similarly, the expected cover time of $k$ random walks is defined by $t^{(k)}_{\text{cov}}(P) := \max_{x \in V} E \left[ \tau^{(k)}_{\text{cov}}(P) \big| X_0 = x \right]$. In particular, let $t^{(k)}_{\text{hit}}(P) := t^{(1)}_{\text{hit}}(P)$ and $t^{(k)}_{\text{cov}}(P) := t^{(1)}_{\text{cov}}(P)$.

Let $(X_t(1))_{t \geq 2}$ and $(X_t(2))_{t \geq 2}$ be two independent random walks, where each walker moves according to $P = (P_t)_{t \geq 1}$. Write $X_t = (X_t(1), X_t(2)) \in V^2$. Then, let $\tau^{(k)}_{\text{meet}}(P) := \min \{ t \geq 0 : X_t(1) = X_t(2) \}$ and define the meeting time of $P$ as $\tau^{(k)}_{\text{meet}}(P) := \max_{x \in V} E \left[ \tau^{(k)}_{\text{meet}}(P) \big| X_0 = x \right]$. In the coalescing random walks, once two or more walkers meet at the same vertex, they merge into one walker. Formally, from a given initial state $C_0 = (C_0(1), \ldots, C_0(n)) \in V^n$, we inductively determine $C_t(a)$ for each $t \geq 1$ and $a \in [n]$, as follows. Suppose that $C_{t-1} = (C_{t-1}(1), \ldots, C_{t-1}(n))$
and \(C_t(1), \ldots, C_t(a-1)\) are determined. If there is some \(b < a\) such that \(C_{t-1}(a) = C_{t-1}(b)\), let \(C_t(a) := C_t(b)\). Otherwise, \(C_t(a)\) is determined by the random walk according to \(P_t\), that is, \(\Pr[C_t(a) = v|C_{t-1}(a) = u] = P_t(u,v)\) for \(u, v \in V\). For \(x = (x_1, x_2, \ldots, x_n) \in V^n\), let \(S(x) := \bigcup_{i=1}^n \{x_i\}\) (e.g., \(S(x) = \{a, b\}\) for \(x = (a, a, b)\)). Then, let \(\tau_{\text{coal}}(P) = \min\{t \geq 0 : |S(C_t)| = 1\}\) and define the coalescing time of \(P\) as \(t_{\text{coal}}(P) := \max_{x \in V} \mathbb{E}[\tau_{\text{coal}}(P)|C_0 = x]\).

### 3 | MIXING TIME

In this section, we show the following lemma, which immediately yields Theorem 1.2.

**Lemma 3.1.** Let \(P = (P_t)_{t \geq 1}\) be a sequence of irreducible, reversible, and lazy transition matrices. Suppose that all \(P_t\) have the same stationary distribution \(\pi\). Then, for any \(u, v \in V\) and any \(0 < \epsilon < 1\), \(\left| \frac{P_{\lfloor t \epsilon^2 \min(\langle \mu, \pi \rangle \rangle)} - 1}{\pi(v)} \right| \leq \epsilon^2\) holds if \(T \geq \frac{\epsilon^2}{e-1} \text{HIT}(P) + 2 \log(4 \text{HIT}(P)) + 1 + t_{\text{REL}}(P) \log e^{-1}\).

We need some more terminology. Let \(\pi \in (0, 1)^V\) be a positive probability distribution and \(f \in \mathbb{R}^V\) be a vector. Let \(E_\pi(f) := \sum_{v \in V} \pi(v)f(v) = \langle f, \mathbb{1}_\pi \rangle\) and \(\text{Var}_\pi(f) := \sum_{v \in V} \pi(v)(f(v) - E_\pi(f))^2 = \langle f, f \rangle_\pi - \langle f, 1 \rangle_\pi^2\). Note that, for any probability vector \(\mu \in [0, 1]^V\), we have

\[
d^{(2, \pi)}(\mu) = \sum_{v \in V} \left( \frac{\mu(v)}{\pi(v)} - 1 \right)^2 = \sum_{v \in V} \pi(v) \left( \frac{\mu(v)}{\pi(v)} \right)^2 - 1 = \text{Var}_\pi \left( \frac{\mu}{\pi} \right).
\]

For a transition matrix \(P\) such that \(\pi(u)P(u,v) = \pi(v)P(v,u)\) holds for all \(u, v \in V\), let \(E_{P,\pi}(f) := \langle f, f \rangle_\pi - \langle f, Pf \rangle_\pi = \frac{1}{2} \sum_{u,v \in V} \pi(u)P(u,v)(f(u) - f(v))^2\) be the Dirichlet form.

#### 3.1 | Key lemma

The following key lemma connects the Dirichlet form, \(\ell^2\)-distance, and hitting time.

**Lemma 3.2.** Let \(P = (P_t)_{t \geq 1}\) be a sequence of irreducible, reversible, and lazy transition matrices. Suppose that all \(P_t\) have the same stationary distribution \(\pi\). Then, for any probability vector \(\mu \in [0, 1]^V\),

\[
E_{P,\pi} \left( \frac{\mu}{\pi} \right) \geq \frac{\text{Var}_\pi \left( \frac{\mu}{\pi} \right)^2}{t_{\text{hit}}(P)}.
\]

**Proof.** Write \(f \in \mathbb{R}^V\) and let \(f_{\text{max}} := \max_{v \in V} f(v)\) and \(g := f_{\text{max}} \mathbb{1} - f\). Since \(f(u) - f(v) = g(v) - g(u)\) for any \(u, v \in V\), we have

\[
E_{P,\pi}(f) = \frac{1}{2} \sum_{u,v \in V} \pi(u)P(u,v)(f(u) - f(v))^2 = E_{P,\pi}(g) = \langle g, g \rangle_\pi - \langle Pg, g \rangle_\pi.
\]

Let \(v_{\text{max}}\) denote a vertex satisfying \(f(v_{\text{max}}) = f_{\text{max}}\). Recall that \(D_v \in \{0, 1\}^{V \times V}\) is a diagonal matrix where \(D(v,v) := \mathbb{1}_{v \neq v}\) for all \(v \in V\). From Lemmas A.1 and A.2, we have

\[
\langle Pg, g \rangle_\pi = (D_{v_{\text{max}}}PD_{v_{\text{max}}}g, g)_\pi \leq \rho(D_{v_{\text{max}}}PD_{v_{\text{max}}})(g, g)_\pi \leq \left(1 - \frac{1}{t_{\text{hit}}(P)}\right) \langle g, g \rangle_\pi.
\]
Note that we have \( g(u)P(u, v)g(v) = g(u)(D_{\max}P D_{\max})(u, v)g(v) \) holds for any \( u, v \in V \). Furthermore,

\[
\langle g, g \rangle_\pi = \sum_{v \in V} \pi(v)(f_{\max} - f(v))^2 = \sum_{v \in V} \pi(v)(f_{\max} - 1 + 1 - f(v))^2 = (f_{\max} - 1)^2 + \sum_{v \in V} \pi(v)(1 - f(v))^2 \geq \text{Var}(f)^2 + \text{Var}(f).
\]

The last inequality follows for \( 0 \leq \text{Var}_\pi(f) = \sum_{v \in V} \pi(v)\left(\frac{\mu(v)}{\pi(v)}\right)^2 - 1 \leq f_{\max} - 1 \). Therefore, we obtain

\[
\mathcal{E}_{p, \pi}(f) = \langle g, g \rangle_\pi - \langle Pg, g \rangle_\pi \geq \frac{\langle g, g \rangle_\pi}{\text{hit}(P)} - \frac{\text{Var}(f)}{\text{hit}(P)}.
\]

\[\blacksquare\]

### 3.2 Upper bound of mixing time

To show Lemma 3.1, we introduce the following bound shown in [43].

\textbf{Lemma 3.3} ([43]). Let \( \mathcal{P} = (P_t)_{t \geq 1} \) be a sequence of irreducible, aperiodic, and reversible transition matrices. Suppose that all \( P_t \) have the same stationary distribution \( \pi \). Then, for any probability vector \( \mu \in [0, 1]^V \),

\[
d^{(2, \pi)}(\mu P_{(1, T)}) \leq d^{(2, \pi)}(\mu) \prod_{t=1}^{T} \lambda_*(P_t) \leq d^{(2, \pi)}(\mu) \exp\left(-\frac{T}{t_{\text{rel}}(\mathcal{P})}\right).
\]

\textbf{Proof.} Applying Lemma A.3 repeatedly to \( d^{(2, \pi)}(\mu P_{(1, T)}) = \left\| \frac{\mu P_{(1, T)}}{\pi} - 1 \right\|_{2, \pi} \), we have

\[
\left\| \frac{\mu P_{(1, T)}}{\pi} - 1 \right\|_{2, \pi} \leq \left\| \frac{\mu P_{(1, T-1)}}{\pi} - 1 \right\|_{2, \pi} \lambda_*(P_T) \leq \cdots \leq \left\| \frac{\mu}{\pi} - 1 \right\|_{2, \pi} \prod_{t=1}^{T} \lambda_*(P_t).
\]

Then, since

\[
\prod_{t=1}^{T} \lambda_*(P_t) = \prod_{i=1}^{T} \left(1 - \frac{1}{t_{\text{rel}}(P_i)}\right) \leq \exp\left(-\sum_{t=1}^{T} \frac{1}{t_{\text{rel}}(P_i)}\right) \leq \exp\left(-\frac{T}{t_{\text{rel}}(\mathcal{P})}\right),
\]

holds, we obtain the claim. \[\blacksquare\]

Combining Lemmas 3.2 and 3.3, we obtain the following bounds of \( \ell^2 \)-distance.

\textbf{Lemma 3.4.} Let \( \mathcal{P} = (P_t)_{t \geq 1} \) be a sequence of irreducible, reversible, and lazy transition matrices. Suppose that all \( P_t \) have the same stationary distribution \( \pi \). Then, for any probability distribution \( \mu \in [0, 1]^V \) and any \( 0 < \epsilon < 1 \), \( d^{(2, \pi)}(\mu P_{(1, T)}) \leq \epsilon \) holds if \( T \geq \epsilon^{-2}t_{\text{HIT}}(\mathcal{P}) + 2 \log(4t_{\text{HIT}}(\mathcal{P})) + 1 + t_{\text{rel}}(\mathcal{P}) \log \epsilon^{-1} \).

\textbf{Proof.} Since \( P \) is reversible with respect to \( \pi \),

\[
\left(\frac{\mu P}{\pi}\right)(v) = \sum_{u \in V} \frac{\mu(u)P(u, v)}{\pi(v)} = \sum_{u \in V} P(v, u)\frac{\mu(u)}{\pi(u)} = \left(\frac{P}{\pi}\right)(v),
\]

(6)
holds for any \( v \in V \), that is, \( \mu_P = P \left( \frac{\mu}{\pi} \right) \) holds. Combining (6) and Lemmas 3.2 and A.4, we have

\[
\text{Var}_\pi \left( \frac{\mu_P}{\pi} \right) = \text{Var}_\pi \left( P \left( \frac{\mu}{\pi} \right) \right) \leq \text{Var}_\pi \left( \frac{\mu}{\pi} \right) - \mathcal{E}_{P,\pi} \left( \frac{\mu}{\pi} \right) \leq \text{Var}_\pi \left( \frac{\mu}{\pi} \right) \left( 1 - \frac{\text{Var}_\pi \left( \frac{\mu}{\pi} \right)}{t_{\text{hit}}(P)} \right). \tag{7} \]

Write \( x(i) = \text{Var}_\pi(\mu_P[1,i]/\pi) = d^{(2,\pi)}(\mu_P[1,i])^2 \). From (7), \( x \) is nonincreasing and \( x(t + 1) \leq x(t) \left( 1 - \frac{x(t)}{t_{\text{hit}}(P)} \right) \) holds. Hence, applying Lemma C.4 to \( x \), \( x(L) \leq 1 \) holds if \( L \geq \frac{e^2}{e-1} t_{\text{hit}}(P) + 2 \log \pi_{\text{min}}^{-1} + 1 = \left( \frac{e^2}{e-1} + o_\pi(1) \right) t_{\text{hit}}(P) \). Note that \( t_{\text{hit}}(P) \geq \pi_{\text{min}}^{-1}(1 - \pi_{\text{min}})^2 \geq \pi_{\text{min}}^{-1}/4 \) holds (see e.g., [2]). From Lemma 3.3,

\[
d^{(2,\pi)}(\mu_P[1,T]) \leq d^{(2,\pi)}(\mu_P[1,L]) \exp \left( - \frac{T - L}{t_{\text{REL}}(P)} \right) \leq \exp \left( - \frac{t_{\text{REL}}(P) \log e^{-1}}{t_{\text{REL}}(P)} \right) \leq e. \]

\[ \square \]

**Proof of Lemma 3.1.** Write \( P_{[2T,T+1]} = P_{2T}P_{2T-1} \ldots P_{T+1} \) for convenience. From the reversibility, it is easy to see that \( \pi(u)P_{[T+1,2T]}(u,v) = \pi(v)P_{[2T,T+1]}(v,u) \) holds. Hence, we have

\[
P_{[1,2T]}(u,v) = \sum_{w \in V} P_{[1,T]}(u,w)P_{[T+1,2T]}(w,v) = \sum_{w \in V} \pi(w) P_{[1,T]}(u,w) P_{[2T,T+1]}(v,w) \pi(w) = \sum_{w \in V} \pi(w) \left( \frac{P_{[1,T]}(u,w)}{\pi(w)} - 1 \right) \left( \frac{P_{[2T,T+1]}(v,w)}{\pi(w)} - 1 \right) + 1.
\]

Combining the above and the Cauchy–Schwarz inequality,

\[
\left| \frac{P_{[1,2T]}(u,v)}{\pi(v)} - 1 \right| \leq \sqrt{\sum_{w \in V} \pi(w) \left( \frac{P_{[1,T]}(u,w)}{\pi(w)} - 1 \right)^2} \sqrt{\sum_{w' \in V} \pi(w') \left( \frac{P_{[2T,T+1]}(v,w')}{\pi(w')} - 1 \right)^2} = d^{(2,\pi)}(P_{[1,T]}(u, \cdot)) \ d^{(2,\pi)}(P_{[2T,T+1]}(v, \cdot)), \tag{8}
\]

holds. Hence, from Lemma 3.4, we obtain the claim. \[ \square \]

For completeness, we prove Proposition 1.1.

**Proof of Proposition 1.1.** We have \( d^{(2,\pi)}(\mu) = \sqrt{\sum_{v \in V} \pi(v) \left( \frac{\mu(v)}{\pi(v)} - 1 \right)^2} \leq \frac{1}{\pi_{\text{min}}} \) for any probability distribution \( \mu \). Let \( T \geq t_{\text{REL}}(P) \log(\pi_{\text{min}}^{-1} e^{-1}) \). Applying Lemma 3.3, we have \( d^{(2,\pi)}(\mu P_{[1,T]}) \leq \epsilon. \) Hence, from (8), it holds for any \( u,v \in V \) that

\[
\left| \frac{P_{[1,2T]}(u,v)}{\pi(v)} - 1 \right| \leq \epsilon^2.
\]

\[ \square \]

## 4 Hitting and Cover Times

In this section, we consider \( k \) independent random walks \( (X_i(1))_{i \geq 0}, \ldots, (X_i(k))_{i \geq 0} \) according to \( P = (P_t)_{t \geq 0} \). Let \( (X_i)_{i \geq 0} \) be a random variable defined as \( X_i = (X_i(1), \ldots, X_i(k)) \in V^k \). Let \( \tau_{\text{hit}}^k(P, w) \)
(for \( w \in V \)) and \( t_{\text{hit}}^{(k)}(P) \) be the random variables denoting hitting and cover times of the \( k \) random walks, defined by (4) and (5). We bound the expected hitting and cover times: \( t_{\text{hit}}^{(k)}(P) \) and \( t_{\text{cov}}^{(k)}(P) \) (see Section 2 for the definitions). We sometimes abbreviate \( P \) and write \( \tau_{\text{hit}}^{(k)}(w) \) and \( \tau_{\text{cov}}^{(k)}(w) \) if \( P \) is clear from the context. This section is devoted to the proof of the following results.

**Theorem 4.1** (Hitting time bound of Theorem 1.4). Let \( P = (P_t)_{t \geq 1} \) be a sequence of irreducible and reversible transition matrices. Suppose that all \( P_t \) have the same stationary distribution \( \pi \). Then, for any \( k \),

\[
t_{\text{hit}}^{(k)}(P) \leq 20t_{\text{sep}}(P) + \frac{400t_{\text{HIT}}(P)}{k}.
\]

**Theorem 4.2** (Cover time bound of Theorem 1.4). Let \( P = (P_t)_{t \geq 1} \) be a sequence of irreducible and reversible transition matrices. Suppose that all \( P_t \) have the same stationary distribution \( \pi \). Then, for any \( k \),

\[
t_{\text{cov}}^{(k)}(P) \leq 20t_{\text{sep}}(P) + \frac{400t_{\text{HIT}}(P) \log n}{k}.
\]

The constant factors 20 and 400 in Theorems 4.1 and 4.2 may be improved by tuning parameters but we do not focus on it.

### 4.1 Key lemma

We prove Lemma 1.11.

**Proof of Lemma 1.11.** Recall \( D_w \in \{0, 1\}^{V \times V} \) is a diagonal matrix defined by \( D_w(x, x) := 1_{x \neq w} \). From the definition of \( D_w \), it is easy to see that

\[
\Pr \left[ \bigwedge_{t=0}^{T} \{X_t \neq w\}, X_T = y \mid X_0 = x \right] = (D_wP_1D_wP_2D_w \cdots D_wP_TD_w)(x, y),
\]

holds for any \( x, y \in V \). Hence, from the assumption on \( X_0 \) and the Cauchy–Schwarz inequality, we have

\[
\Pr \left[ \bigwedge_{t=0}^{T} \{X_t \neq w\} \right] \leq \sum_{x \in V} \pi(x) \sum_{y \in V} \left( \prod_{t=1}^{T} (D_wP_tD_w) \right)(x, y) \leq \left\| \prod_{t=1}^{T} (D_wP_tD_w) \right\|_{2, \pi}.
\]  

From the reversibility of \( P_t \), we have \( \pi(u)(D_wP_tD_w)(u, v) = \pi(v)(D_wP_tD_w)(v, u) \) for any \( u, v \in V \) and \( t \geq 1 \). Hence, we can apply Lemma A.2 repeatedly to (9) and obtain

\[
\left\| \prod_{t=1}^{T} (D_wP_tD_w) \right\|_{2, \pi} \leq \rho(D_wP_tD_w) \left\| \prod_{t=2}^{T} (D_wP_tD_w) \right\|_{2, \pi} \leq \cdots \leq \prod_{t=1}^{T} \rho(D_wP_tD_w).
\]

Then, using Lemma A.1, \( \rho(D_wP_tD_w) \leq 1 - \frac{1}{t_{\text{hit}}(P_t)} \) holds for all \( t \). Thus, we obtain the claim.

Note that all \( P_t \) are irreducible by assumption.
4.2 Upper bound of hitting time

**Lemma 4.3.** Let \( P = (P_t)_{t \geq 1} \) be a sequence of irreducible and reversible transition matrices. Suppose that all \( P_t \) have the same stationary distribution \( \pi \). Then, for any \( k, x \in V^k \) and \( t \geq 0 \), it holds that

\[
\text{Pr} \left[ \tau_{\text{hit}}^{(k)}(w) \geq t_{\text{sep}}(P) + \frac{20t_{\text{HIT}}(P)}{k} \right] + \text{Pr}[X_t = x] < 0.9.
\]

**Proof.** Let \( T := t_{\text{sep}}(P) = t_{\text{sep}}(P, 1/2) \), where \( t_{\text{sep}} \) is the separation time defined by (2). Then, the matrix \( P_{[1,T]} = \prod_{t=1}^T P_t \) can be written as

\[
P_{[1,T]}(u, v) = \frac{1}{2} \pi(v) + \frac{1}{2} Q(u, v),
\]

for some transition matrix \( Q \in [0, 1]^{V \times V} \). Thus, the position of a walker after a transition according to \( P_{[1,T]} \) has the same distribution as the following transition: The walker flips a fair coin. If it is head, the walker moves according to the stationary distribution \( \pi \). Otherwise, the next position of the walker at vertex \( u \in V \) is determined by the distribution \( Q(u, \cdot) \).

Suppose \( k \) independent walkers flip their own coins and then move to the position \( X_T \in V^k \) according to the transition probability \( P_{[1,T]} \). Let \( S \subseteq [k] \) be the random subset of indices of walkers with a head coin. Then, the distribution of \( X_T(i) \) conditioned on \( i \in S \) is \( \pi \). Let \( w \in V \) be a target vertex. From Lemma 1.11 and the independency of the walkers, for any \( T' \geq 0 \) and \( U \subseteq [k] \), we obtain

\[
\text{Pr} \left[ \bigwedge_{i \in U} \bigwedge_{T \leq t \leq T + T'} \{X_T(i) \neq w\} \bigg| S = U \right] \leq \prod_{i \in U} \prod_{T \leq t \leq T + T'} \left( 1 - \frac{1}{t_{\text{hit}}(P_t)} \right) \leq \exp \left( -\frac{|U|T'}{t_{\text{HIT}}(P)} \right). \tag{10}
\]

From the Chernoff inequality (Lemma C.2), we have \( \text{Pr}[|S| < k/4] \leq \exp(-k/16) \) (note that \( E[|S|] = k/2 \)). For any events \( A \) and \( B \), \( \text{Pr}[A \cap B] = \text{Pr}[A \wedge B] + \text{Pr}[A \wedge \overline{B}] \leq \text{Pr}[A \mid B] + \text{Pr}[\overline{B}] \) holds. Therefore, setting \( T' = 20t_{\text{HIT}} / k \), we obtain

\[
\text{Pr}[\tau_{\text{hit}}^{(k)}(P) \geq T + T'] \leq \text{Pr} \left[ \bigwedge_{i \in [k]} \bigwedge_{T \leq t \leq T + T'} \{X_T(i) \neq y\} \right] \leq \text{Pr} \left[ \bigwedge_{i \in S} \bigwedge_{T \leq t \leq T + T'} \{X_T(i) \neq y\} \bigg| |S| \geq k/4 \right] + \text{Pr}[|S| < k/4] \leq \exp \left( -\frac{kT'}{4t_{\text{HIT}}} \right) + \exp \left( -\frac{k}{16} \right) = \exp(-5) + \exp\left(-\frac{1}{16}\right) < 0.95,
\]

for any \( k \geq 1 \). Since this inequality holds regardless of the initial position \( X_0 \in V^k \), we obtain the claim. \( \blacksquare \)

**Proof of Theorem 4.1.** Theorem 4.1 follows from Corollary B.2 and Lemma 4.3. \( \blacksquare \)
4.3 | Upper bound of cover time

**Lemma 4.4.** Let \( P = (P_t)_{t \geq 1} \) be a sequence of irreducible and reversible transition matrices. Suppose that all \( P_t \) have the same stationary distribution \( \pi \). Then, for any \( k, x \in V^k \), \( t \geq 0 \) and every sufficiently large \( n \), it holds that

\[
\Pr \left[ \tau_{\text{cov}}^{(k)}(w) \geq t_{\text{sep}}^{(x)}(P) + \frac{20t_{\text{HIT}}(P) \log n}{k} + t \mid X_t = x \right] < 0.95.
\]

**Proof.** From (10) and the union bound over the target vertex \( w \in V \), we have

\[
\Pr \left[ \bigvee_{w \in V} \bigwedge_{i \in U} \bigwedge_{T \leq t \leq T + T'} \{ X_t(i) \neq w \} \bigg| |S| = U \right] \leq n \prod_{i \in U} \prod_{T \leq t \leq T + T'} \left( 1 - \frac{1}{t_{\text{hit}}(P_t)} \right) \leq n \exp \left( -\frac{|U| T'}{t_{\text{HIT}}(P)} \right).
\]

Setting \( T' = 20t_{\text{HIT}} \log n/k \), we obtain

\[
\Pr[\tau_{\text{cov}}^{(k)}(P) > T + T'] \leq \Pr \left[ \bigvee_{w \in V} \bigwedge_{i \in [k]} \bigwedge_{T \leq t \leq T + T'} \{ X_t(i) \neq y \} \right] \\
\leq \Pr \left[ \bigvee_{w \in V} \bigwedge_{i \in S} \bigwedge_{T \leq t \leq T + T'} \{ X_t(i) \neq y \} \bigg| |S| \geq k/4 \right] + \Pr[|S| < k/4] \\
\leq n \exp \left( -\frac{kT'}{4t_{\text{HIT}}} \right) + \exp \left( -\frac{k}{16} \right)
\]

\[
= n \exp(-5 \log n) + \exp \left( -\frac{k}{16} \right)
\]

\[
< 0.94 + O(n^{-4}) < 0.95,
\]

for any \( k \geq 1 \) and every sufficiently large \( n \). Since this inequality holds regardless of the initial position \( X_0 \in V^k \), we obtain the claim.  

**Proof of Theorem 4.2.** Theorem 4.2 follows from Corollary B.2 and Lemma 4.4.

5 | MEETING AND COALESCING TIMES

We prove Proposition 1.5 and Theorem 1.6.

5.1 | Meeting time lemma on dynamic graph

**Proof of Lemma 1.12.** Recall that \( D_w \in \{0, 1\}^{V \times V} \) is a diagonal matrix defined by \( D_w(x, x) = 1_{x \neq \text{w}} \). In the same way as the proof of Lemma 1.11 (see Section 4.1),

\[
\Pr \left[ \bigwedge_{t=0}^T \{ X_t \neq w \}, X_T = y \right| X_0 = x ] = (D_{w_0}P_1D_{w_1}P_2D_{w_2} \cdots D_{w_{T-1}}P_T D_{w_T})(x, y),
\]
holds for any \( x, y \in V \), and hence we have
\[
\Pr \left[ \bigwedge_{t=0}^{T} \{ X_t \neq w_i \} \right] = \sum_{x \in V} \pi(x) \sum_{y} \left( \prod_{t=1}^{T} (D_{w_{i-1}} P_{D_{w_i}}) \right) (x, y) \leq \left\| \prod_{t=1}^{T} (D_{w_{i-1}} P_{D_{w_i}}) \right\| _{2, \pi}.
\]

Here, we used the Caushy–Schwarz inequality. Since \( P_t \) is reversible and lazy, applying Lemma A.5 repeatedly yields
\[
\left\| \left( \prod_{t=1}^{T} (D_{w_{i-1}} P_{D_{w_i}}) \right) \right\| _{2, \pi} \leq \sqrt{\rho(D_{w_{i}} P_{D_{w_i}}) \rho(D_{w_{i-1}} P_{D_{w_i}}) \left\| \prod_{t=2}^{T} (D_{w_{i-1}} P_{D_{w_i}}) \right\| _{2, \pi}} \leq \cdots \leq \prod_{t=1}^{T} \rho(D_{w_{i-1}} P_{D_{w_i}}) \rho(D_{w_{i}} P_{D_{w_i}}).
\]

Finally, from Lemma A.1, we have \( \rho(D_{w} P_{D_{w}}) \leq 1 - \frac{1}{t_{iw}(P_{w})} \) for any \( t \). Thus, we obtain the claim. Note that \( P_t \) is irreducible for any \( t \).

### 5.2 Upper bound of coalescing time

Consider the coalescing random walks \((C_t(1))_{t \geq 0}, (C_t(2))_{t \geq 0}, \ldots, (C_t(n))_{t \geq 0}\) according to \( \mathcal{P} = (P_t)_{t \geq 1} \) (see Section 2 for the definition). This section is devoted to the proof of Theorem 1.6. To prove Theorem 1.6, we introduce the following notation. Let \( \ell_x := \lceil \log_2(\pi_x) \rceil \) and \( \ell_n := \ell_{\pi} \). Let \( K \) be a suitable constant that will be determined later. Define \( L_{\ell}, L_{\ell-1}, \ldots, L_0 \) recursively by \( L_{\ell} = t_{\text{sep}}(\mathcal{P}) \) and
\[
L_i = L_{i+1} + \left\lceil \frac{K_{\text{hit}}(P_{\pi})}{2^\ell} \right\rceil.
\]
In other words, \( L_i = t_{\text{sep}}(\mathcal{P}) + \sum_{j=1}^{\ell-i} \left\lceil \frac{K_{\text{hit}}(P_{\pi})}{2^j} \right\rceil \) and \( L_0 = t_{\text{sep}}(\mathcal{P}) + \sum_{j=0}^{\ell} \left\lceil \frac{K_{\text{hit}}(P_{\pi})}{2^j} \right\rceil \).

The following result states a relation between the coalescing random walk and the independent random walks. The proof is essentially based on the argument of [39].

**Lemma 5.1.** Let \((X_t)_{t \geq 0} = ((X_t(1), \ldots, X_t(n)))_{t \geq 0}\) be \( n \) independent random walks, where each walker moves according to \( \mathcal{P} = (P_t)_{t \geq 1} \). Suppose that \( C_0 = x \) and \( X_0 = x \) for an arbitrary initial position of walkers \( x = (x_1, x_2, \ldots, x_n) \in V^n \). Then, we have
\[
\Pr \left[ \tau_{\text{coal}}(\mathcal{P}) > L_0 \right] \leq \Pr \left[ \bigvee_{a=2}^{n} \bigwedge_{b \in [2^a - 1]} \bigwedge_{i=1}^{L_{\ell+a-1}} \{ X_i(b) \neq X_i(a) \} \right].
\]

**Proof.** First, we define the random walks with killings \((Y_t)_{t \geq 0} = ((Y_t(1), \ldots, Y_t(n)))_{t \geq 0}\), where a walker is killed when it meets another walker with a smaller index. Formally, \((Y_t)_{t \geq 0}\) is a Markov chain on a state space \((V \cup \{ \partial \})^n\), where \( \partial \not\in V \) is a coffin state. Set \( Y_t(1) = X_t(1) \) for all \( t \geq 0 \). For \( t \geq 0 \) and \( a \geq 1 \), define \( Y_t(a) \) inductively as follows. Suppose that \((Y_{\ell-1})_{t \geq 0}\) and \( Y_t(1), \ldots, Y_t(a-1) \) are determined. Then, define
\[
\tau(a) := \min\{ s \geq 0 : X_s(a) = Y_s(b) \text{ for some } b < a \}, \quad \text{ and}
\]
\[
Y_t(a) := \begin{cases} X_t(a) & \text{if } t < \tau(a), \\ \partial & \text{otherwise.} \end{cases}
\]

Next, we define the random walks with a list of allowed killings \((Z_t)_{t \geq 0} = ((Z_t(1), \ldots, Z_t(n)))_{t \geq 0}\), where a walker \( a \geq 2^{\ell+1} \) is killed when it meets a “killer” walker.
\[2^i \leq b < 2^{i+1}\] during the time period \(L_i \leq t \leq L_{i+1}\). Figure 2 is an overall picture. Formally, define the list of allowed killings \(A = \{A_t\}_{t \geq 0}\) as \(A_t := \emptyset\) for \(t \leq L_\ell\) and \(A_t := \{(b,a) : 2^i \leq b < 2^{i+1}, a \geq 2^{i+1}\}\) for \(L_{i+1} < t \leq L_i\) with \(0 \leq i < \ell\). Let \(Z_i(1) = X_i(1)\) for all \(t \geq 0\). For \(t \geq 0\) and \(a \geq 1\), \(Z_t(a)\) is inductively defined as follows. Suppose that \((Z_s)_{s=0}^0\) and \(Z_s(1), \ldots, Z_s(a-1)\) are determined. Then, let

\[
\tau^A(a) := \min\{s \geq 0 : X_s(a) = Z_s(b)\text{ for some } (b,a) \in A_s\}
\]

and

\[
Z_t(a) := \begin{cases} X_t(a) & \text{if } t < \tau^A(a), \\ \partial & \text{otherwise.} \end{cases}
\]

For a vector \(x \in (V \cup \{\partial\})^n\), let \(S(x) = \{i \in [n] : x(i) \neq \partial\}\). Obviously, \(\Pr[|S(C_t)| \geq z|C_0 = x] = \Pr[|S(Y_t)| \geq z|X_0 = x]\) holds for any \(t \geq 0, z \geq 0\) and \(x \in V^n\). We claim the following.

**Lemma 5.2.** For any \(t \geq 0\) and \(z \geq 0\), \(\Pr[|S(Y_t)| \geq z] \leq \Pr[|S(Z_t)| \geq z]\).

This lemma is already given in [39, proposition 3.4]. For completeness, we prove it in Appendix D.

We bound the second term of (11). From the definition of the random walk with allowed killings, \(Z_t(i) = X_t(i)\) until it meets a killer. Hence, we have

\[
\Pr[\tau_{\text{coal}}(\mathcal{P}) > L_0] \leq \Pr[|S(Z_{L_b})| \geq 2] = \Pr\left[\bigvee_{a=2}^n \{\text{Walker } a \text{ is not killed}\}\right] = \Pr\left[\bigvee_{a=2}^n \bigwedge_{b=1}^{L_{a-1}} \{Z_t(a) \neq Z_t(b)\}\right] = \Pr\left[\bigvee_{a=2}^n \bigwedge_{b=1}^{L_{a-1}} \{X_t(b) \neq X_t(a)\}\right].
\]

Combining Lemmas 1.12 and 5.1, we obtain the following lemma.

**Lemma 5.3.** Let \(\mathcal{P} = (P_i)_{i \geq 1}\) be a sequence of irreducible, reversible, and lazy transition matrices. Suppose that all \(P_i\) have the same stationary distribution \(\pi\). Then, for any \(x \in V^n\),

\[
\Pr[\tau_{\text{coal}}(\mathcal{P}) > T | C_0 = x] \leq 1 - 10^{-5}\] holds if \(T \geq t_{\text{coal}}(\mathcal{P}) + 80f_{\text{hit}}(\mathcal{P}) + \log_2(n)\).
Proof. From the definition (2) of the separation time \( t^{(x)}_{\text{sep}}(P) = L_\ell \), there is a transition matrix \( Q \in [0,1]^{V \times V} \) such that

\[
P_{[1,L_\ell]}(x,u) = \frac{1}{2} \pi(u) + \frac{1}{2} Q(x,u),
\]

holds for all \( x,u \in V \). Hence, the distribution of \( n \) walkers \( X_{L_\ell}(1), \ldots, X_{L_\ell}(n) \) can be simulated as follows: each walker \( i \in [n] \) flips its own fair coin. If it is head, the walker’s position \( X_{L_\ell}(i) \) is sampled according to \( \pi \). Otherwise, it is sampled according to the distribution \( Q(X_{0}(i)) \). Let \( I \subseteq [n] \) denote a random subset of indices with a head coin. Let \( \mathcal{W} := \{ W \subseteq [n] : |\{ 2^i, \ldots, 2^{i+1} - 1 \} \cap W | \geq 2^i/4 \} \) holds for all \( 0 \leq i < \ell \) be a set of subsets of \([n]\). Then, from Lemma 5.1,

\[
\Pr[\tau_{\text{coal}}(P) > L_0] 
\leq \sum_{W \in [n]} \Pr \left[ \bigvee_{a=2}^{n} b \in [2^{a-1}] \bigcap \bigwedge_{i \in [L_{\ell_1+1}, L_{\ell_2}]} \{ X_{i}(b) \neq X_{i}(a) \} \right] \Pr[I = W] 
\leq \Pr[I \notin \mathcal{W}] + \max_{W \in \mathcal{W}} \left[ \bigvee_{i=2}^{n} \bigwedge_{a=2} \bigcap_{b \in [2^{a-1}] \cap W} \bigwedge_{i \in [L_{\ell_1+1}, L_{\ell_2}]} \{ X_{i}(b) \neq X_{i}(a) \} \right] \bigwedge_{i \in W} \{ X_{L_{\ell_i}}(i) \sim \pi \}. \tag{11}
\]

Let \( I_j = \mathbb{1}_{j \in I} \in \{0,1\} \) denote the binary indicator for the random subset \( I \). For the first term of (11), applying the Chernoff inequality (Lemma C.2) yields

\[
\Pr[I \in \mathcal{W}] = \prod_{i=0}^{\ell-1} \left[ \sum_{j=2^i}^{2^{i+1}-1} \Pr[I_j \geq 2^i/4] \right] 
\geq \prod_{i=0}^{5} \left( 1 - \exp \left( -\frac{2^i}{16} \right) \right) \prod_{i=6}^{\ell-1} \left( 1 - \exp \left( -\frac{2^i}{16} \right) \right) 
\geq 0.00033 \cdot \prod_{i=2}^{\ell-5} (1 - \exp(-2^i)) \geq 0.00033 \cdot \prod_{i=2}^{\infty} (1 - 2^{-i}) 
\geq 0.00033 \cdot \left( 1 - \sum_{i=2}^{\infty} 2^{-i} \right) \geq 0.00016. \tag{12}
\]

Note that \( \mathbb{E} \left[ \sum_{j=2^i}^{2^{i+1}-1} I_j \right] = 2^i/2 \). Next, we bound the second term of (11). For any \( W \in \mathcal{W} \),

\[
\Pr \left[ \bigvee_{a=2}^{n} b \in [2^{a-1}] \cap W \bigcap \bigwedge_{i \in [L_{\ell_1+1}, L_{\ell_2}]} \{ X_{i}(b) \neq X_{i}(a) \} \right] \bigwedge_{i \in W} \{ X_{L_{\ell_i}}(i) \sim \pi \} 
\leq \sum_{a=2}^{n} \prod_{b \in [2^{a-1}] \cap W} \Pr \left[ \bigwedge_{i \in [L_{\ell_1+1}, L_{\ell_2}]} \{ X_{i}(b) \neq X_{i}(a) \} \right] \bigwedge_{i \in W} \{ X_{L_{\ell_i}}(i) \sim \pi \},
\]
holds. Applying the meeting time lemma (Lemma 1.12), for any \( b \in W \), we have

\[
\Pr \left[ \bigwedge_{i \in [L_{C_b+1}, L_{C_b}]} \{ X_t(b) \neq X_t(a) \} \bigwedge_{i \in W} \{ X_L(i) \sim \pi \} \right] \leq \exp \left( -\frac{L_{C_b} - L_{C_b+1}}{t_{\text{HIT}}(P)} \right) \leq \exp \left( -\frac{K}{2C_b} \right).
\]

Hence, for any \( W \in \mathcal{W} \),

\[
\Pr \left[ \bigwedge_{a=2}^{n} \bigwedge_{i=1}^{L_{C_b+1} - 1} \prod_{j=1}^{\ell - 2^{a+1} - 1} \prod_{b \in [2^{a+1} - 1]} \{ X_t(b) \neq X_t(a) \} \bigwedge_{i \in W} \{ X_L(i) \sim \pi \} \right] \leq \sum_{i=1}^{\ell - 2^{a+1} - 1} \sum_{j=1}^{\ell - 2^{a+1} - 1} \prod_{b \in [2^{a+1} - 1]} \exp \left( -\frac{K}{2} \right) \leq \sum_{i=1}^{\ell - 2^{a+1} - 1} \sum_{j=1}^{\ell - 2^{a+1} - 1} \prod_{b \in [2^{a+1} - 1]} \exp \left( -\frac{K}{4} \right)
\]

\[
\leq \sum_{i=1}^{\ell - 2^{a+1} - 1} \prod_{j=1}^{\ell - 2^{a+1} - 1} \exp \left( -\frac{K}{4} \right) \leq \frac{2}{e^{K/4} - 2}.
\]

Combining (11) to (13) with \( K = 40 \), \( \Pr \left[ \tau_{\text{coal}}(P) > T \big| C_0 = x \right] \leq 1 - 0.00016 + \frac{2}{e^{10} - 2} \leq 1 - 10^{-5} \) holds.

\textit{Proof of Theorem 1.6.} Let \( T := [(85 + o(1))t_{\text{HIT}}(P)] \). Lemma 5.3 implies that, for any \( t \geq 0 \) and \( x \in \mathcal{V}^m \), \( \Pr[\tau_{\text{coal}}(P) > t + T \big| C_t = x] \leq 1 - 10^{-5} \) holds. Thus we obtain the claim from Corollary B.2.

\section{Lower bound of meeting time}

In this section, we prove Proposition 1.5.

\textit{Proof of Proposition 1.5.} We consider the lazy simple random walk on the graph sequence given in proposition 12 of [41]. For completeness, we present the sequence formally. For a graph \( H \) and a permutation \( \eta \) on \( V(H) \), let \( \eta(H) \) be the graph given by \( V(\eta(H)) = V(H) \) and \( E(\eta(H)) = \{ \{ \eta(u), \eta(v) \} : \{ u, v \} \in E(H) \} \).

For any integer \( m \in \mathbb{N} \), let \( U = \{ u_0, \ldots, u_{m-1} \} \) and \( W = \{ w_0, \ldots, w_{m-1} \} \). Define the graph \( G \) by \( V(G) := U \cup W \) and

\[
E(G) := \{ u_0, w_0 \} \cup \bigcup_{i=1}^{m-1} \{ u_i, u_0 \} \cup \bigcup_{j=1}^{m-1} \{ w_j, w_0 \}.
\]

(14)

Let \( \eta \) be the permutation on \( V \) defined by \( \eta(u_i) = u_{(i+1) \mod m} \) and \( \eta(w_j) = w_{(j+1) \mod m} \).

We claim that the lazy simple random walk on the sequence \( (G_t)_{t \geq 1} \) given by \( G_1 = G \) and \( G_{t+1} = \eta(G_t) \) \((t \geq 1)\) has the desired property (see Figure 3). Consider two independent lazy simple random walks \( (X_t(1))_{t \geq 0} \) and \( (X_t(2))_{t \geq 0} \) with initial positions \( (X_0(1), X_0(2)) = (u_{m-1}, w_{m-1}) \). Suppose \( \tau_{\text{meet}} \leq t \). Then, there exists \( t' \leq t \) such that either \( X(1) \) or \( X(2) \) moves along the edge \( \{ u_j, w_j \} \) for \( j = t' \mod m \). Focus on the walk \( X(1) \) and suppose \( X_{t'-1}(1) \in U \) and \( X_{t'}(1) \in W \). To reach \( u_j \), the walker \( X(1) \) must choose the self loop for \( m - 1 \) consecutive times, which occurs with probability \( 2^{-m+1} \). Therefore, by the union bound over \( X(0) \) and \( X(1) \), we have \( \Pr[\tau_{\text{meet}} \leq t] \leq 2 \cdot 2^{-m+1} \) and we have \( t_{\text{meet}} = 2^{\Omega(m)} \).
6 | PULL VOTING

In this section, we prove Theorem 1.8. Our proof of bounding $E[\tau_{\text{cons}}]$ is inspired by the idea of well-known duality between the pull voting and coalescing random walk [26].

**Proposition 6.1** (Duality in static setting [26]). Let $P \in [0, 1]^{V \times V}$ be an irreducible transition matrix. Let $\tau_{\text{cons}}(P)$ be the consensus time of the pull voting according to $P$ where all vertices initially hold distinct opinions. Let $\tau_{\text{coal}}(P)$ be the coalescing time of the coalescing random walk according to $P$. Then, for every $j \geq 0$, $\Pr[\tau_{\text{coal}}(P) \leq j] = \Pr[\tau_{\text{cons}}(P) \leq j]$ holds.

From Proposition 6.1, we can obtain bounds of $\tau_{\text{cons}}$ by studying $\tau_{\text{coal}}$. Indeed, the proof of previous results bounding $\tau_{\text{cons}}$ on a static graph relies on the duality. In this paper, we obtain the following consensus–coalescing relation that is analogous to Proposition 6.1 in the time inhomogeneous setting.

**Proposition 6.2** (Consensus–Coalescing relation on dynamic graphs). Let $P$ be a sequence of transition matrices (not necessarily has a time-homogeneous stationary distribution). Consider the pull voting according to $P$ such that initially all vertices have $n$ distinct opinions. Then, there is a sequence $(Q(i))_{i \geq 0}$ where each $Q(i) = (Q(i)_t)_{t \geq 1}$ is a transition matrix sequence such that, for every $j \geq 0$,

$$\Pr[\tau_{\text{cons}}(P) \leq j] = \Pr[\tau_{\text{coal}}(Q(i)) \leq j],$$

holds. Moreover, if $P$ is reversible and has the time-invariant stationary distribution $\pi$, then so do the sequences $Q(i)$ for all $i$.

Indeed, if $P = (P_t)_{t \geq 1}$ is a sequence of a static transition matrix, then $Q(i) = P$ for all $i \geq 0$, implying Proposition 6.1.

The proof of the duality theorem in the static setting (Proposition 6.1) is obtained by constructing a coupling of the pull voting and the coalescing random walk with equal consensus and coalescing times. Our proof of Proposition 6.2 is based on essentially the same argument. In Section 6.4, we present a sequence of graphs on which the pull voting according to $PLS$ on it has an exponential consensus time as follows:

**Proposition 6.3.** There is a sequence $(G_i)_{i \geq 1}$ of graphs on which the pull voting according to $P = (PLS(G_i))_{i \geq 1}$ over opinion set $\Sigma = \{0, 1\}$ satisfies $t_{\text{cons}} = 2^{\Omega(n)}$.

6.1 | Consensus–coalescing relation

We prove Proposition 6.2. The proof is essentially based on the notion of linear voter model of [16].
\begin{proof}[Proof of Proposition 6.2] Let $P \in [0, 1]^{V \times V}$ be a transition matrix and $S$ be the set of all binary $V \times V$ matrices such that each row contains exactly one 1. For each matrix $S \in S$, we define a probability distribution $\mu_P$ over $S$ by

$$\mu_P(S) = \prod_{(i,j) \in V \times V : S_{ij} = 1} P_{ij},$$

for each $S \in S$. We interpret $S \in S$ as the list of selections at a specific round of the pull voting: Specifically, $s_{ij} = 1$ if and only if $i$ selects $j$ at the pull voting. Then, $\mu_P(\cdot)$ can be seen as the probability distribution over the set of all possible selection lists during the pull voting according to $P$.

Given a sequence $S_1, \ldots, S_l \in S$ of $l$ matrices, we can simulate the pull voting for $l$ rounds as follows. Let $y_0 \in \Sigma^V$ denote the initial opinion configuration where each vertex has a distinct opinion. Then, $y_i = S_i y_{i-1} = S_i S_{i-1} \ldots S_1 y_0 = \prod_{t=1}^l S_{i-t+1} y_0$. We say that an opinion vector $y \in \Sigma^V$ is in consensus if $y = \sigma 1$ holds for some $\sigma \in \Sigma$. For fixed $i \in \mathbb{N}$, let

$$S_{\text{cons}}^{(i)} = \{ (S_1, \ldots, S_l) \in S^l : y_i \text{ is in consensus} \}.$$

Here, note that, if $y_{i-1}$ is in consensus, then so does $y_i = S_i y_{i-1}$. Then, we have

$$\Pr[\tau_{\text{cons}}(P) \leq i] = \sum_{(S_1, \ldots, S_l) \in S_{\text{cons}}^{(i)}} \prod_{l=1}^i \mu_P(S_l). \quad (15)$$

We show that, for every $i$, there is a sequence $Q = Q^{(i)} = (Q^{(i)}_t)_{t \geq 0}$ of transition matrices such that the right-hand side of (15) is equal to $\Pr[\tau_{\text{coal}}(Q^{(i)}) \leq i]$. Consider the sequence $Q^{(i)}$ defined by

$$Q^{(i)}_t = \begin{cases} P_{t-1} + 1 & \text{if } t \leq i, \\ P_0 & \text{if } t > i. \end{cases}$$

Note that, if $P$ is reversible and has a time-invariant stationary distribution $\pi$, then so does $Q^{(i)}$ for every $i$.

Fix $i$ and consider the coalescing random walk according to $Q^{(i)}$. We call a vector $c \in \mathbb{Z}^V_{\geq 0}$ satisfying $\sum_{v \in V} c_v = |V|$ a walker configuration vector. A walker configuration vector can be interpreted as the vector denoting the number of walkers on vertices, that is, $c_v$ is the number of walkers on $v$. Given $S_1', \ldots, S_l' \in S$, we can simulate the coalescing random walk for $i$ rounds as follows: Let $c_0 = 1^T$ (initially, each vertices has exactly one walker). For $1 \leq j \leq i$, let $c_j = c_{j-1} S_j' = c_0 \prod_{k=1}^j S_k'$. Intuitively speaking, the matrix $S_j'$ denotes the transition result: $(S_j')_{u,v} = 1$ if and only if $u$ sends all walkers on it to $v$ at the $j$th round. A vertex $v \in V$ at round $i$ has a walker if and only if $(c_i)_v > 0$.

We say that a walker configuration vector $c$ is in coalesce if $c = |V| e_v$ for some $v \in V$, where $e_v \in \{0, 1\}^V$ is the binary indicator vector for a vertex $v$ (i.e., $(e_v)_u = 1$ if and only if $v = u$). Note that, if $c_{j-1}$ is in coalesce, then so does $c_j = c_{j-1} S_j'$. For a coalescing random walk according to a transition matrix $Q$, a transition result $S' \in S$ occurs with probability $\mu_Q(S')$. Let

$$S_{\text{coal}}^{(i)} = \{ (S_1', \ldots, S_l') \in S^l : c_i \text{ is coalescing} \}.$$


Then, we have

$$\Pr[\tau_{\text{coal}}(Q^{(i)}) \leq i] = \sum_{(S'_1, \ldots, S'_n) \in S^{(i)}_{\text{coal}}} \prod_{i=1}^{n} \mu_{Q^{(i)}}(S'_i) = \sum_{(S'_1, \ldots, S'_n) \in S^{(i)}_{\text{coal}}} \prod_{i=1}^{n} \mu_{P_i}(S'_{i-t+1}).$$

(16)

We compare (15) and (16). Indeed, it holds that \((S_1, \ldots, S_t) \in S_{\text{cons}}^{(i)}\) if and only if \((S_i, \ldots, S_i) \in S^{(i)}_{\text{coal}}\). To see this, suppose \((S_1, \ldots, S_t) \in S_{\text{cons}}^{(i)}\). Then, the opinion configuration vector \(y_i\) at the \(i\)th round satisfies \(y_i = S_i \ldots S_1 v_0 = \sigma_w 1\) for some \(w \in V\), where \(\sigma_w \in \Sigma\) is the opinion that \(w \in V\) initially holds. Since this relation holds regardless of the labels \(\Sigma\) of opinions, we have \(S_i \ldots S_1 e_w = 1\) and \(S_i \ldots S_1 e_v = 0\) for \(v \neq w\), where \(0\) denotes the all-zero vector. Therefore,

$$1^T S_i S_{i-1} \ldots S_1 e_v = \begin{cases} |V| & \text{if } v = w, \\ 0 & \text{otherwise,} \end{cases}$$

and thus \(1^T \prod_{i=1}^{t} S_i = |V| e_v^T\). In other words, \((S'_1, \ldots, S'_n) := (S_i, \ldots, S_i) \in S^{(i)}_{\text{coal}}\). The converse direction (i.e., \((S'_1, \ldots, S'_n) \in S^{(i)}_{\text{coal}}\) implies \((S'_1, \ldots, S'_n) \in S^{(i)}_{\text{cons}}\)) can be checked similarly: If \((S'_1, \ldots, S'_n) \in S^{(i)}_{\text{coal}}\), then \(1^T S'_1 \ldots S'_t = |V| e_w^T\) for some \(w \in V\). Then, for \((S_1, \ldots, S_t) := (S'_1, \ldots, S'_n)\), we have \(1^T S_i \ldots S_1 e_w = 1^T S'_i \ldots S'_1 e_w = |V|\). Since \(0 \leq S_i \ldots S_1 e_w \leq S_i \ldots S_1 1 = 1\) (here, we write \(x_1, \ldots, x_n\) \((y_1, \ldots, y_n)\) if \(x_1 \leq y_1\) for all \(i \in [n]\)), we have \(S_i \ldots S_1 e_w = 1\). This implies \((S_1, \ldots, S_t) \in S^{(i)}_{\text{cons}}\).

The mapping \(\phi : (S_1, \ldots, S_t) \mapsto (S'_1, \ldots, S'_n)\) is a bijection between \(S^{(i)}_{\text{cons}}\) and \(S^{(i)}_{\text{coal}}\) preserving the product measure \(\prod_{i=1}^{n} \mu_{P_i}(S_i)\). This implies that (15) and (16) are equal, completing the proof of Proposition 6.2.

6.2 | Consensus time

**Proof of Theorem 1.8.** If \(P\) is irreducible, lazy, and reversible with respect to \(\pi \in (0, 1]^V\), so does \(Q^{(i)}\) for all \(i \geq 0\). Therefore, from Theorem 1.6, \(E[\tau_{\text{coal}}(Q^{(i)})] \leq T\) for all \(i\), where \(T = C \cdot t_{\text{HIT}}(P)\) for some absolute constant \(C > 0\). Then, from Proposition 6.2, we have

$$\Pr[\tau_{\text{cons}}(P) \geq 2T] = \Pr[\tau_{\text{coal}}(Q^{(2T)}) \geq 2T] \leq \frac{E[\tau_{\text{coal}}(Q^{(2T)})]}{2T} \leq \frac{1}{2}.$$\n
In the initial opinion configuration, all vertices have \(n\) distinct opinions. Therefore, for any fixed \(t \geq 0\), \(\Pr[\tau_{\text{cons}}(P_i+1)_{t \geq 0}] \leq 1/2\) holds for any initial opinion configuration. From Corollary B.2 we have \(E[\tau_{\text{cons}}(P_t)] \leq 4T = 4t_{\text{HIT}}(P) = O(t_{\text{HIT}}(P))\).

6.3 | Winning probability

We prove Proposition 1.10. Our proof is based on the voting martingale argument that was used to obtain the winning probability result for the static graph setting (cf. [16, 26]). We just verify that the argument works for our dynamic graph setting. For completeness, we write the proof in this subsection.

**Proof of Proposition 1.10.** We first consider the special case of \(\Sigma = \{0, 1\}\) and then go on to the general case \(\Sigma \subseteq \{0, \ldots, n-1\}\).
The case of $\Sigma = \{0, 1\}$. Let $(Y_t)_{t \geq 0}$ be the pull voting according to $\mathcal{P} = (P_t)_{t \geq 1}$ with a time-homogeneous stationary distribution $\pi$. Note that $Y_t \in \{0, 1\}^V$. In this proof, we promise that $Y_t$ is an $n \times 1$ vector and $\pi \in [0, 1]^V$ is a $1 \times n$ vector. Let $\pi(Y_t) = \sum_{u \in V} \pi(u)Y_t(u)$. We first claim that $(\pi(Y_t))_{t \geq 0}$ is a martingale with respect to $Y_t$. To see this, observe

$$
E[\pi(Y_{t+1})|Y_t] = \sum_{u \in V} \pi(u)\Pr[Y_{t+1}(u) = 1|Y_t] = \sum_{u \in V} \pi(u) \sum_{w \in V} (P_t)(u, w)Y_t(w) = \sum_{w \in V} (\pi P_t)(w)Y_t(w) = \pi(Y_t).
$$

Since $\pi(Y_t)$ are bounded, we can apply the Optimal Stopping Theorem and obtain $E[\pi(Y_{\text{cons}})] = \pi(Y_0)$. Note that, since $Y_{\text{cons}}$ is either all-zero or all-one, we have

$$
E[\pi(Y_{\text{cons}})] = \pi(1)\Pr[Y_{\text{cons}} = 1] = \Pr[Y_{\text{cons}} = 1].
$$

In other words, the probability that the opinion 1 wins is equal to $\pi(Y_0) = \sum_{v \in V: Y_0(v) = 1} \pi(v)$.

General case. We reduce the general case to the binary opinion case by regarding $\sigma = 1$ for a fixed opinion $\sigma \in \Sigma$ and all other opinions are zero. Then we obtain the claim from the argument for the binary opinion case.

6.4 Exponential consensus time

In this subsection, we prove Proposition 6.3. Consider the graph $G = (U \cup W, E)$ and permutation $\eta : V(G) \to V(G)$ defined in Section 5.3. Suppose vertices in $U$ have opinion 0 and that in $W$ have opinion 1 initially. We claim that the sequence $(G_t)_{t \geq 1}$ given by $G_1 = G$ and $G_t = \eta(G_{t-1})$ with the initial opinion configuration above has the exponential consensus time.

By the monotonicity of the pull voting, we consider the following setting. Suppose only vertices in $U = \{u_0, \ldots, u_{m-1}\}$ perform the pull voting and vertices in $W$ always have opinion 1. Let $\tau$ be the time to reach the opinion configuration where all vertices in $U$ have opinion 1. It suffices to prove $E[\tau] = 2^{\Omega(m)}$, where $|U| = m$.

Since the graph dynamics is given by the iteration of applying the permutation $\eta$, it is convenient to consider the following equivalent process: At the $r$th round, vertices perform the one-round pull voting and then opinions are shuffled according to $\eta$. That is, if $Y_{r-1} \in \{0, 1\}^U$ denote the opinion configuration at the beginning of the $r$th round, we first update $Y_{r-1}$ by the pull voting to obtain $Z_r \in \{0, 1\}^U$ and then set $Y_r(\eta(u)) = Z_r(u)$ for every $u \in U$. Note that the permutation $\eta$ defined in Section 5.3 satisfies $\eta(U) = U$. We consider the sequence $(Y_r)_{r \geq 0}$ described above, where $Y_0$ is the all-zero vector. Note that the process agrees with the opinion 1 at the last of the $r$th round if and only if $Y_r(u) = 1$ for all $u \in U$.

Let $r' = \inf\{i \geq 0 : Y_i(0) = 1\}$. Then, for any $i = 1, \ldots, m-1$, the vertex $u_i$ must choose the self-loop in the pull voting procedure at the $(r' - i)$th round to keep $u_i$'s opinion 1 (otherwise, $u_i$ selects $u_0$, who has opinion 0). This happens with probability $2^{-m+1}$ and therefore we have $E[\tau] \geq E[\tau'] = 2^{\Omega(m)}$. This completes the proof of Proposition 6.3.

7 METROPOLIS WALK ON EDGE-MARKOVIAN GRAPH

In this section, we prove Theorem 1.7. Let $V$ be a vertex set with $|V| = n$ and $p, q \in [0, 1]$ be two parameters. Let $((Y_t(e))_{t \geq 0})_{e \in \binom{V}{2}}$ be $\binom{n}{2}$ independent Markov chains, where each $(Y_t(e))_{t \geq 0}$ is the
We show the following lemma in this section. Markov chain with the state space $\{0, 1\}$ and the transition matrix $M = M_{p,q} := \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix}$. Henceforth, write $Y_t = (Y_t(e))_{e \in \binom{V}{2}}$ for convenience. The edge-Markovian graph $\mathcal{G}(n, p, q)$ is a sequence of random graphs $(G_t)_{t \geq 0} = ((V,E_t))_{t \geq 0}$, where $E_t := \left\{ e \in \binom{V}{2} : Y_t(e) = 1 \right\}$ for all $t \geq 0$. We show the following lemma in this section.

**Lemma 7.1.** Suppose that $0 < p+q \leq 1$ and $\frac{p}{p+q} \geq 32(c+1) \frac{\log n}{n}$ hold for an arbitrary $c > 0$. Let $\mathcal{G}(n, p, q) = (G_t)_{t \geq 0}$ be the edge-Markovian graph. Let $I = I(p, q) := \left\lfloor \frac{\max\{1, \log(q/p)\}}{p+q} \right\rfloor$ and $J \leq n'/2$. For any $\ell' \geq 0$ and $0 \leq i \leq \ell'$, let $S(\ell', i) := \left( \frac{\ell' - 1}{2} + i \right) (I + J) = \left( \sum_{j=1}^{\ell'-1} j + i \right) (I + J)$. Let $C = 8192$. Then, for any $Y_0 = b \in \{0, 1\}^{\binom{V}{2}}$ (Figure 4),

$$\Pr \left[ \bigwedge_{i=1}^{\ell} \bigvee_{t=1}^{\ell} \left\{ t_{\text{REL}} \left( (P_{LM}(G_t))_{s \in S(\ell', i-1)+l} \right) \leq C \right\} \right] \geq 1 - \frac{1}{n}.$$

**Proof.** For $S \subseteq V$, let $Q_t(S) = \sum_{u \in S} \sum_{v \notin S} \pi(u)p_t(u,v)$ and $\pi(S) = \sum_{v \in S} \pi(v)$. Let $C_t(S) = \sum_{u \in S} \sum_{v \in S} Y_t(\{u,v\})$. From (3), we have

$$\frac{Q_t(S)}{\pi(S)} = \sum_{u \in S} \sum_{v \notin S} \frac{1}{n} \sum_{u \in S} \sum_{v \notin S} \frac{Y_t(\{u,v\})}{\max \{ \deg(G_t, u), \deg(G_t, v) \}} \geq \frac{C_t(S)}{2 |S| d_{\text{max}}(G_t)}. \quad (17)$$

Now, it is easy to check that

$$M' = \frac{1}{p+q} \begin{pmatrix} q & p \\ q & p \end{pmatrix} + (1-p-q)' \begin{pmatrix} p & -p \\ -q & q \end{pmatrix},$$

holds for any $t \geq 1$. Hence, for any $e \in \binom{V}{2}$, $s \geq 0$, $\ell' \geq I = \left\lfloor \frac{\max\{1, \log(q/p)\}}{p+q} \right\rfloor$ and $b \in \{0, 1\}$, $\Pr[Y_{s+t'}(e) = 1|Y_s(e) = b] = M'_t(b, 1)$ and

$$\frac{p}{2(p+q)} \leq \frac{p}{p+q} \left( 1 - (1-p-q)' \right) \leq M'_t(b, 1) \leq \frac{p}{p+q} \left( 1 + (1-p-q)' \frac{q}{p} \right) \leq \frac{2p}{p+q}. \quad (18)$$

\(^3\)Formally, $\Pr[Y_t(e) = b_t|Y_0(e) = b_0, \ldots, Y_{t-1}(e) = b_{t-1}] = \Pr[Y_t(e) = b_t|Y_{t-1}(e) = b_{t-1}] = M(b_{t-1}, b_t)$ holds for all $t \geq 1$, $(b_0, \ldots, b_t) \in \{0, 1\}^{t+1}$ and $e \in \binom{V}{2}$.\end{document}
Using (18) and the Chernoff inequality (Lemma C.2), for any \( s \geq 0, I' \geq I \) and \( b \in \{0, 1\}^{|Y_v|} \), we have

\[
\Pr \left[ \bigvee_{S \subseteq V: |S| \leq n/2} \left\{ C_{s+I'}(S) \leq \frac{|S|np}{2(p+q)} \right\} \bigg| Y_s = b \right] \\
\leq \sum_{S \subseteq V: |S| \leq n/2} \exp \left( -\frac{|S|np}{32(p+q)} \right) \leq \sum_{S \subseteq V: |S| \leq n/2} \exp(-c + 1)|S| \log n \\
= \sum_{1 \leq s \leq n/2} \left( \frac{n}{x} \right)^{n/n-1} \leq \frac{1}{n^c}. \tag{19}
\]

Note that \( \mathbb{E} \left[ C_{s+I'}(S) \big| Y_s = b \right] \geq |S|(n - |S|) \frac{p}{2(p+q)} \geq \frac{|S|np}{4(p+q)} \) for any \( 1 \leq |S| \leq n/2 \). Furthermore, since \( \mathbb{E} \left[ \deg(G_{s+I'}, v) \big| Y_s = b \right] \leq n \frac{2p}{p+q} \), we have

\[
\Pr \left[ \bigvee_{v \in V} \left\{ \deg(G_v, v) \geq 2n \frac{2p}{p+q} \right\} \bigg| Y_s = b \right] \leq \sum_{v \in V} \exp \left( -\frac{2np}{3(p+q)} \right) \leq \frac{1}{n^{64n(n+1)/3-1}}. \tag{20}
\]

Combining (17), (19), and (20), it holds with probability at least 1 - \( \frac{2}{n^c} \) that

\[
\min_{S \subseteq V: 0 < \pi(S) \leq 1/2} \frac{Q_{s+I}(S)}{\pi(S)} \geq \frac{|S|np}{8(p+q)} \geq \frac{|S|np}{4(p+q)} = \frac{1}{64}.
\]

Hence, applying Lemma C.3 yields the following: for any \( s \geq 0, I' \geq I \) and \( b \in \{0, 1\}^{|Y_v|} \),

\[
\Pr \left[ t_{REL}(P_{LM}(G_{s+I})) \leq C \big| Y_s = b \right] \geq 1 - \frac{2}{n^c}. \tag{21}
\]

Here, \( C = 2 \cdot 64^2 = 8192 \). From (21) and the union bound,

\[
\Pr \left[ t_{REL} \left( \left( P_{LM}(G_i) \right)_{i=s+I}^{s+I'} \right) > C \big| Y_s = b \right] = \Pr \left[ \bigvee_{i=s+I}^{s+I'} \left\{ t_{REL}(P_{LM}(G_i)) > C \big| Y_s = b \right\} \right] \leq \frac{2I}{n^c} \leq \frac{1}{n}, \tag{22}
\]

holds for any \( s \geq 0 \). Let \( E_i^{(e)} \) be the event that \( t_{REL} \left( \left( P_{LM}(G_i) \right)_{i=S(e,i-1)+1}^{S(e,i)} \right) > C \). Fix \( Y_0 = b \).

From (22), we have

\[
\Pr \left[ \bigwedge_{i=1}^{\ell} E_i^{(e)} \right] = \sum_{b' \in [0,1]^{|Y_v|}} \Pr \left[ E_i^{(e)} \big| Y_{S(e,\ell-1)} = b' \right] \Pr \left[ \bigwedge_{i=1}^{\ell-1} E_i^{(e)}, Y_{S(e,\ell-1)} = b' \right] \\
\leq \frac{1}{n} \Pr \left[ \bigwedge_{i=1}^{\ell-1} E_i^{(e)} \right] \leq \cdots \leq \frac{1}{n^c}.
\]

Hence, we obtain

\[
\Pr \left[ \bigwedge_{\ell=1}^{\infty} \bigvee_{i=1}^{\ell} E_i^{(e)} \right] = 1 - \Pr \left[ \bigvee_{\ell=1}^{\infty} \bigwedge_{i=1}^{\ell} E_i^{(e)} \right] \geq 1 - \sum_{\ell=1}^{\infty} \Pr \left[ \bigwedge_{i=1}^{\ell} E_i^{(e)} \right] \geq 1 - \sum_{\ell=1}^{\infty} \frac{1}{n^c}. \tag*{■}
\]
Proof of the hitting time bound in Theorem 1.7. Let $J = \lceil 8192(2 \log n + \log 2) + 20 \cdot 8192n/k \rceil$. From Lemmas 4.3 and 7.1, it holds for any $\ell \geq 1$ that $\Pr[\tau_{\text{hit}}^{(k)}(P) > \ell(1, 0)]$ holds for any $\ell \geq 1$ that $\Pr[\tau_{\text{hit}}^{(k)}(P) > S(\ell, 0)] \leq 0.9$. Applying Lemma B.3 yields $E[\tau_{\text{hit}}^{(k)}(P)] \leq \sum_{\ell=0}^{\infty} \ell(I + J)(0.9)^\ell \leq 100(I + J)$.

Proof of the cover time bound in Theorem 1.7. Let $J = \lceil 8192(2 \log n + \log 2) + 20 \cdot 8192n \log n/k \rceil$. From Lemmas 4.4 and 7.1, it holds for any $\ell \geq 1$ that $\Pr[\tau_{\text{cov}}^{(k)}(P) > \ell(1, 0)]$ holds for any $\ell \geq 1$ that $\Pr[\tau_{\text{cov}}^{(k)}(P) > S(\ell, 0)] \leq 0.9$. Applying Lemma B.3, $E[\tau_{\text{cov}}^{(k)}(P)] \leq \sum_{\ell=0}^{\infty} \ell(I + J)(0.9)^\ell \leq 100(I + J)$.

Proof of the coalescing time bound in Theorem 1.7. Let $J = \lceil 8192Cn \rceil$. From Lemmas 5.3 and 7.1, it holds for any $\ell \geq 1$ that $\Pr[\tau_{\text{coal}}^{(k)}(P) > S(\ell, 0)]$ holds for any $\ell \geq 1$ that $\Pr[\tau_{\text{coal}}^{(k)}(P) > S(\ell, 0)] \leq 1 - 10^{-5}$ for any $\ell$. Applying Lemma B.3, $E[\tau_{\text{coal}}^{(k)}(P)] \leq \sum_{\ell=0}^{\infty} \ell(I + J)(1 - 10^{-5})^\ell \leq C'(I + J)$ holds for some absolute constant $C' > 0$.

8 CONCLUSION

We obtain new bounds on the mixing, hitting, and cover times of the random walk according to the sequence of irreducible, reversible, and lazy transition matrices that have the same stationary distribution. These bounds generalize previous work for a lazy simple random walk or a $d_{\text{max}}$-lazy walk and improve them in several cases. We also obtain the first bounds on the hitting and cover times of multiple random walks and the coalescing time on dynamic graphs. In addition, we bound the consensus time of the pull-voting on dynamic graphs. Our results reinforce the observation that time-inhomogeneous Markov chains with an invariant stationary distribution behave almost identically to a time-homogeneous chain. Specifically, we prove that if all $P_i$ have the same stationary distribution, then $\tau_{\text{hit}}(P_{i \geq 1}) \leq O(\max_{\ell \geq 1} \tau_{\text{hit}}(P_i))$ holds (Theorem 1.4). It is natural to ask for the same relation for other parameters. For example, does $\tau_{\text{cov}}(P) \leq O(\max_{\ell \geq 1} \tau_{\text{cov}}(P_i))$ hold?

Most previous works on time-inhomogeneous random walks have translated techniques from time-homogeneous chains into time-inhomogeneous ones. In particular, several known upper bounds (including ours) are based on spectral arguments, which essentially requires the time-homogeneity of the stationary distribution. On the other hand, known lower bounds such as the Sisyphus wheel are based on some combinatorial arguments. To understand time-inhomogeneous random walks with time-varying stationary distributions, it might be important to interpolate the spectral and combinatorial arguments. For example, it is known that the simple random walk on any static connected graph has an $O(n^3)$ cover time, which was shown by the combinatorial trick using a spanning tree combined with a bound on the edge commute time. Can we translate this strategy to time-inhomogeneous chains to obtain an $O(n^3)$ cover time? This research question might be a possible future direction of the research of time-inhomogeneous chains.

REFERENCES

1. M. Abdullah, C. Cooper, and M. Draief. Speeding up cover time of sparse graphs using local knowledge (Proc. 26th Int. Workshop on Comb. Algorithms (IWOCA)), Verona, Italy, 2015, pp. 1–12.
2. D. J. Aldous and J. A. Fill. Reversible Markov chains and random walks on graphs. https://www.stat.berkeley.edu/users/aldous/RWG/book.html, 2002.
3. R. Aleliunas, R. M. Karp, R. J. Lipton, L. Lovász, and C. Rackoff. Random walks, universal traversal sequences, and the complexity of maze problems (Proc. 20th Symp. Found. Comput. Sci. (SFCS)), San Juan, Puerto Rico, 1979, pp. 218–223.
How to cite this article: N. Shimizu and T. Shiraga, Reversible random walks on dynamic graphs, Random Struct. Alg. 63 (2023), 1100–1136. https://doi.org/10.1002/rsa.21164

APPENDIX A: TOOLS FOR KEY LEMMAS

In this section, we introduce technical tools for Lemmas 1.11, 1.12, 3.2, and 3.3. The first one is concerned with the spectral radius $\rho(D_w P D_w)$ of the substochastic matrix $D_w P D_w$ (see Section 1.4 for the definition of $D_w$). It is known that $\rho(D_w P D_w) \leq 1 - 1/t_{hit}(P)$ for any irreducible $D_w P D_w$ (section 3.6.5 of [2]). For completeness, we show this under the assumption of irreducibility and reversibility of $P$ as follows.

Lemma A.1. Let $P \in [0, 1]^{V \times V}$ be an irreducible and reversible transition matrix. Then, for any $w \in V$,

$$\rho(D_w P D_w) \leq 1 - \frac{1}{t_{hit}(P)}.$$ 

Proof. Define $P_w \in [0, 1]^{V \setminus \{w\} \times V \setminus \{w\}}$ by $P_w(u, v) = P(u, v)$ for any $u, v \in V \setminus \{w\}$. The Perron–Frobenius theorem implies that $\lambda = \rho(P_w)$ is an eigenvalue of $P_w$ and there is a nonnegative nonzero eigenvector $g \in \mathbb{R}^{V \setminus \{w\}}$ satisfying $P_w g = \lambda g$ (see, e.g., theorem 8.3.1 in [27]). Write $Q_w = D_w P D_w$ for convenience. Define $h \in \mathbb{R}^V$ by $h(v) = \frac{g(v)\pi(v)}{Z}$ for any $v \in V \setminus \{w\}$, where $Z = \sum_{v \in V \setminus \{w\}} g(v)\pi(v)$. Then, $h$ is a probability vector. Furthermore,

$$(hQ_w)(v) = \sum_{u \in V} h(u)Q_w(u, v) = \sum_{u \in V \setminus \{w\}} g(u)\pi(u)Z^{-1} P_w(u, v) = \sum_{u \in V \setminus \{w\}} g(u)\pi(u)Z^{-1} P_w(v, u) = \pi(v)Z^{-1} (P_w g)(v) = \frac{\pi(v)}{Z} \lambda g(v) = \lambda h(v),$$

where $D_w$ is a substochastic matrix.

Let $\Pi_{s_i}$ be the $s_i$-th row of $P$. Then, there is a probability vector $\pi_{s_i}$ such that $P \Pi_{s_i} = \pi_{s_i}$ and $\pi_{s_i} = \frac{1}{s_i} \Pi_{s_i}$. We define $h_{s_i} = \sum_{u \in V} h(u)\Pi_{s_i}(u)$ for any $u \in V$. We take $h_{s_i}$ to be a probability vector. Furthermore, by the Perron–Frobenius theorem, $\lambda_{s_i} = \rho(P_{s_i})$ is an eigenvalue of $P_{s_i}$ and there is a nonnegative nonzero eigenvector $g_{s_i} \in \mathbb{R}^{V \setminus \{w\}}$ satisfying $P_{s_i} g_{s_i} = \lambda_{s_i} g_{s_i}$ (see, e.g., theorem 8.3.1 in [27]). Write $Q_{s_i} = D_{s_i} P_{s_i} D_{s_i}$ for convenience. Then, $Q_{s_i}$ is a substochastic matrix. Define $h_{s_i} \in \mathbb{R}^V$ by $h_{s_i}(v) = \frac{g_{s_i}(v)\pi_{s_i}(v)}{Z}$ for any $v \in V \setminus \{w\}$, where $Z = \sum_{v \in V \setminus \{w\}} g_{s_i}(v)\pi_{s_i}(v)$. Then, $h_{s_i}$ is a probability vector. Furthermore,

$$(h_{s_i}Q_{s_i})(v) = \sum_{u \in V} h_{s_i}(u)Q_{s_i}(u, v) = \sum_{u \in V \setminus \{w\}} g_{s_i}(u)\pi_{s_i}(u)Z^{-1} P_{s_i}(u, v) = \sum_{u \in V \setminus \{w\}} g_{s_i}(u)\pi_{s_i}(u)Z^{-1} P_{s_i}(v, u) = \pi_{s_i}(v)Z^{-1} (P_{s_i} g_{s_i})(v) = \frac{\pi_{s_i}(v)}{Z} \lambda_{s_i} g_{s_i}(v) = \lambda_{s_i} h_{s_i}(v),$$

where $D_{s_i}$ is a substochastic matrix.

Proof. Define $P_{a_i} \in [0, 1]^{V \times V}$ by $P_{a_i}(u, v) = P(u, v)$ for any $u, v \in V \setminus \{w\}$. The Perron–Frobenius theorem implies that $\lambda = \rho(P_{a_i})$ is an eigenvalue of $P_{a_i}$ and there is a nonnegative nonzero eigenvector $g_{a_i} \in \mathbb{R}^{V \setminus \{w\}}$ satisfying $P_{a_i} g_{a_i} = \lambda_{a_i} g_{a_i}$ (see, e.g., theorem 8.3.1 in [27]). Write $Q_{a_i} = D_{a_i} P_{a_i} D_{a_i}$ for convenience. Then, $Q_{a_i}$ is a substochastic matrix. Define $h_{a_i} \in \mathbb{R}^V$ by $h_{a_i}(v) = \frac{g_{a_i}(v)\pi_{a_i}(v)}{Z}$ for any $v \in V \setminus \{w\}$, where $Z = \sum_{v \in V \setminus \{w\}} g_{a_i}(v)\pi_{a_i}(v)$. Then, $h_{a_i}$ is a probability vector. Furthermore,

$$(h_{a_i}Q_{a_i})(v) = \sum_{u \in V} h_{a_i}(u)Q_{a_i}(u, v) = \sum_{u \in V \setminus \{w\}} g_{a_i}(u)\pi_{a_i}(u)Z^{-1} P_{a_i}(u, v) = \sum_{u \in V \setminus \{w\}} g_{a_i}(u)\pi_{a_i}(u)Z^{-1} P_{a_i}(v, u) = \pi_{a_i}(v)Z^{-1} (P_{a_i} g_{a_i})(v) = \frac{\pi_{a_i}(v)}{Z} \lambda_{a_i} g_{a_i}(v) = \lambda_{a_i} h_{a_i}(v),$$

where $D_{a_i}$ is a substochastic matrix.
holds for any \( v \in V \setminus \{w\} \). Since \( (hQ_w)(w) = 0 = \lambda h(w) \), we have \( hQ_w = \lambda h \). Hence, \( hQ_w = \lambda' h \) holds for any \( t \geq 1 \). This implies that

\[
Pr_h [\tau_w > t] = Pr_h \left[ \bigwedge_{i=0}^{t-1} \{X_i \neq w\} \right] = \sum_{v \in V} h(v) \sum_{u \in W} Q_u(v, u) = \lambda' \sum_{u \in V} h(u) = \lambda',
\]

holds for any \( t \geq 1 \). Since \( P \) is irreducible, there is a \( t^* \geq 1 \) such that \( Pr_h [\tau_w > t^*] < 1 \). Hence, \( \lambda < 1 \) and we have

\[
E_h [\tau_w] = \sum_{t=0}^{\infty} Pr_h [\tau_w > t] = \frac{1}{1 - \lambda}.
\]

Note that \( Pr_h [\tau_w > 0] = 1 \) holds from \( h(w) = 0 \). Thus, \( \rho(D_wPD_w) = \lambda = 1 - \frac{1}{E_h [\tau_w]} \leq 1 - \frac{1}{t_{\text{hit}}(P)} \) holds and we obtain the claim. Note that we have \( E_h [\tau_w] = \sum_{v \in V} h(v)E_v [\tau_w] \leq \sum_{v \in V} h(v)t_{\text{hit}}(P) = t_{\text{hit}}(P) \).

The following lemmas are already known in the literature. We put the proofs of them for completeness.

**Lemma A.2.** Let \( M \in \mathbb{R}^{V \times V} \) be a matrix and \( v \in \mathbb{R}^V \) be a positive vector. Suppose that \( v(u)M(u, v) = v(v)M(v, u) \) holds for all \( u, v \in V \). Then,

\[
\langle Mf, f \rangle_v \leq \rho(M)\langle f, f \rangle_v \quad \text{and} \quad ||Mf||_{2,v} \leq \rho(M)||f||_{2,v},
\]

hold for any \( f \in \mathbb{R}^V \). Furthermore, if \( M \) is a transition matrix,

\[
\langle Mf, f \rangle_v \leq \lambda_*(M)\langle f, f \rangle_v \quad \text{and} \quad ||Mf||_{2,v} \leq \lambda_*(M)||f||_{2,v},
\]

hold for any \( f \in \mathbb{R}^V \) satisfying \( \langle f, 1 \rangle_v = 0 \).

**Proof.** From the assumption, \( \langle Mf, g \rangle_v = \langle f, Mg \rangle_v \) holds for any \( f, g \in \mathbb{R}^V \). Hence, from the spectral theorem, the inner product space \( (\mathbb{R}^V, \langle \cdot, \cdot \rangle_v) \) has an orthonormal basis of real-valued eigenvectors \( \{\psi_i\}_{i=1}^{|V|} \) corresponding to real eigenvalues \( \{\lambda_i(M)\}_{i=1}^{|V|} \) (see, e.g., lemma 12.1 in [34]). In other words, for any \( i, j \) and \( f \in \mathbb{R}^V \), we have \( M\psi_i = \lambda_i(M)\psi_i \), \( \langle \psi_i, \psi_j \rangle_v = 1 \) if \( i = j \), and \( f = \sum_{i=1}^{|V|} \langle f, \psi_i \rangle_{v, i} \psi_i \). Without loss of generality, assume \( \lambda_1(M) \geq \lambda_2(M) \geq \cdots \lambda_{|V|}(M) \). For any \( f \in \mathbb{R}^V \), we have

\[
||f||_{2,v}^2 = \langle f, f \rangle_v = \sum_{i=1}^{|V|} \langle f, \psi_i \rangle_{v, i}^2 = \sum_{i=1}^{|V|} \langle f, \psi_i \rangle_v^2, \quad (A1)
\]

\[
||Mf||_{2,v}^2 = \sum_{i=1}^{|V|} \langle Mf, \psi_i \rangle_v^2 = \sum_{i=1}^{|V|} \langle f, M\psi_i \rangle_v^2 = \sum_{i=1}^{|V|} \lambda_i(M)^2 \langle f, \psi_i \rangle_v^2, \quad (A2)
\]

\[
\langle Mf, f \rangle_v = \sum_{i=1}^{|V|} \langle Mf, \psi_i \rangle_v \langle \psi_i, f \rangle_v = \sum_{i=1}^{|V|} \lambda_i(M) \langle f, \psi_i \rangle_v^2. \quad (A3)
\]

Combining (A1) and (A3), \( \langle Mf, f \rangle_v \leq \rho(M)\langle f, f \rangle_v \) holds. Combining (A1) and (A2), \( ||Mf||_{2,v} \leq \rho(M)^2||f||_{2,v} \) holds. If \( M \) is a transition matrix, we have \( \lambda_1(M) = 1 \) and
\(\psi_1 = 1\). Furthermore, \(|\lambda_i(M)| \leq 1\) holds for all \(1 \leq i \leq |V|\). Combining (A1) and (A3),

\[
\langle Mf, f \rangle_v \leq \lambda_1(M)\langle f, \psi_1 \rangle_v + \lambda_\ast(M)\langle f, f \rangle_v = \lambda_\ast(M)\langle f, f \rangle_v
\]

holds. Combining (A1) and (A2),

\[
\|Mf\|_{2,\pi}^2 \leq \lambda_1(M)\|f, \psi_1\|_v^2 + \lambda_\ast(M)\|f, f\|_v = \lambda_\ast(M)\|f\|_{2,\pi}^2
\]

holds. \hfill \blacksquare

**Lemma A.3** (see, e.g., (12.8) of [34]). Let \(P \in [0, 1]^{V \times V}\) be a transition matrix. Suppose that \(\pi(u)P(u, v) = \pi(v)P(v, u)\) holds for any \(u, v \in V\) and some probability distribution \(\pi \in (0, 1)^V\). Then, for any probability vector \(\mu \in [0, 1]^V\),

\[
\left\| \frac{\mu P}{\pi} - 1 \right\|_{2,\pi}^2 \leq \lambda_\ast(P)^2 \left\| \frac{\mu}{\pi} - 1 \right\|_{2,\pi}^2.
\]

**Proof.** Combining (6) and Lemma A.2, we have

\[
\left\| \frac{\mu P}{\pi} - 1 \right\|_{2,\pi}^2 = \left\| P \left( \frac{\mu}{\pi} - 1 \right) + P1 \right\|_{2,\pi}^2 = \left\| P \left( \frac{\mu}{\pi} - 1 \right) \right\|_{2,\pi}^2 \leq \lambda_\ast(P)^2 \left\| \frac{\mu}{\pi} - 1 \right\|_{2,\pi}^2.
\]

Note that we have \(\left\langle \frac{\mu}{\pi} - 1, 1 \right\rangle_\pi = \sum_{i \in V} \pi(i) \left( \frac{\mu(i)}{\pi(i)} - 1 \right) = 0\). \hfill \blacksquare

**Lemma A.4** (see, e.g., proposition 2.5 in [24]). Let \(P \in [0, 1]^{V \times V}\) be a lazy transition matrix. Suppose that \(\pi(u)P(u, v) = \pi(v)P(v, u)\) holds for any \(u, v \in V\) and some probability distribution \(\pi \in (0, 1)^V\). Then for any \(f \in \mathbb{R}^V\),

\[
\text{Var}_\pi(Pf) \leq \text{Var}_\pi(f) - \mathcal{E}_{\pi}(f, f).
\]

**Proof.** It is straightforward to see that

\[
\text{Var}_\pi(Pf) = \langle Pf, Pf \rangle_\pi - \langle Pf, f \rangle_\pi^2 = \langle P^2f, f \rangle_\pi - \langle f, Pf \rangle_\pi^2 = \langle f, f \rangle_\pi - \langle f, Pf \rangle_\pi + \langle P^2f, f \rangle_\pi - \langle Pf, f \rangle_\pi^2 = \text{Var}_\pi(f) - \mathcal{E}_{\pi}(f, f)
\]

holds. From (A1) and (A3), \(\mathcal{E}_{\pi}(f, f) = \langle f, f \rangle_\pi - \langle Pf, f \rangle_\pi = \lambda_\ast(P)^2 \sum_{i=1}^{|V|} \psi_i(f)\psi_i(f)^2\). Hence, we have \(\mathcal{E}_{\pi}(f, f) = \lambda_\ast(P)^2 \sum_{i=1}^{|V|} (1 - \lambda_i(P)) (f, \psi_i)^2 \geq \sum_{i=1}^{|V|} (1 - \lambda_i(P)) (f, \psi_i)^2 = \mathcal{E}_{\pi}(f, f)\). Thus, we obtain the claim. Note that all eigenvalues of \(P\) are non-negative since \(P\) is lazy.

**Lemma A.5** (see, e.g., theorem 4.1 in [40]). Let \(P \in [0, 1]^{V \times V}\) be a lazy transition matrix. Suppose that \(\pi(u)P(u, v) = \pi(v)P(v, u)\) holds for any \(u, v \in V\) and some probability distribution \(\pi \in (0, 1)^V\). Then for any \(x, y \in V\) and any \(f \in \mathbb{R}^V\),

\[
\left\| D_xPD_yf \right\|_{2,\pi}^2 \leq \rho(D_xPD_x) \rho(D_yPD_y) \left\| f \right\|_{2,\pi}^2.
\]

**Proof.** From assumption, the inner product space \((\mathbb{R}^V, \langle \cdot, \cdot \rangle_\pi)\) has an orthonormal basis of real-valued eigenvectors \(\{\psi_i\}_{i=1}^{|V|}\) corresponding to real eigenvalues \(\{\lambda_i(P)\}_{i=1}^{|V|}\). This implies that, for all \(u, v \in V\), \(P(v, u) = \pi(u)\sum_{i=1}^{|V|} \sqrt{\lambda_i(P)}\psi_i(v)\psi_i(u)\). Let \(\sqrt{P} \in [0, 1]^{V \times V}\) be the positive semidefinite square root of \(P\), that is, \(\sqrt{P}(v, u) = \pi(u)\sum_{i=1}^{|V|} \sqrt{\lambda_i(P)}\psi_i(v)\psi_i(u)\). Note that all eigenvalues are nonnegative since \(P\) is lazy. It is easy to see that \((\sqrt{P})^2 = P\) and \(\pi(v)\sqrt{P}(v, u) = \pi(u)\sqrt{P}(u, v)\) holds for any \(u, v \in V\). Hence, we have

\[
\pi(v)\sqrt{P}(v, u) = \pi(v)D_n(v, v)\sqrt{P}(v, u) = \pi(u)D_n(v, u)\sqrt{P}(u, v) = \pi(u)(\sqrt{P}D_n)(u, v),
\]
that is, \( (D_w \sqrt{P} f, g)_{\mathcal{F}} = (f, \sqrt{P} D_w g)_{\mathcal{F}} \) holds for any \( f, g \). This implies that both \( \|D_w \sqrt{P} f\|_{2,\mathcal{F}} \) and \( \|\sqrt{P} D_w f\|_{2,\mathcal{F}} \) are upper bounded by \( \sqrt{\rho(D_w PD_w)} \|f\|_{2,\mathcal{F}} \). Consequently,

\[
\|D_w PD_w f\|_{2,\mathcal{F}}^2 = \left\| D_w \sqrt{P} \sqrt{PD_w f} \right\|_{2,\mathcal{F}}^2 \leq \rho(D_w PD_w) \left\| \sqrt{PD_w f} \right\|_{2,\mathcal{F}}^2 = \rho(D_w PD_w) \rho(D_w PD_w) \|f\|_{2,\mathcal{F}}^2,
\]

holds, and we obtain the claim. \( \square \)

**APPENDIX B: TOOLS FOR EXPECTED STOPPING TIME**

Our upper bounds of the hitting, cover, and coalescing times rely on the following observation.

**Lemma B.1.** Let \((Z_t)_{t \geq 0}\) be a sequence of random variables where \(Z_t \in S\) for a finite state space \(S\). For an event \(E \subseteq S\), let \(\tau(E) = \inf\{t \geq 0 : Z_t \in E\}\) be the stopping time. Suppose there exist \(T > 0\) and \(c > 0\) such that, for any \(t \geq 0\),

\[
\Pr[\tau(E) \geq T + t | \tau(E) \geq t] \leq 1 - c,
\]

holds. Then, \(E[\tau(E)] \leq \frac{T}{c}\).

**Proof.** From the assumption, for any \(k \geq 0\), we have

\[
\Pr[\tau(E) \geq kT + t] = \Pr[\tau(E) \geq kT + t | \tau(E) \geq (k - 1)T + t] \Pr[\tau(E) \geq (k - 1)T + t] \\
\leq (1 - c) \Pr[\tau(E) \geq (k - 1)T + t] \\
\vdots \\
\leq (1 - c)^k.
\]

Therefore, we have

\[
E[\tau(E)] = \sum_{k=0}^{\infty} \sum_{t=0}^{T-1} \Pr[\tau(E) \geq kT + t] \leq T \sum_{k=0}^{\infty} (1 - c)^k = T \frac{c}{c}.
\]

\( \square \)

**Corollary B.2.** Let \((Z_t)_{t \geq 0}\) be a sequence of random variables where \(Z_t \in S\) for a finite state space \(S\). For an event \(E \subseteq S\), let \(\tau(E) = \inf\{t \geq 0 : Z_t \in E\}\) be the stopping time. Suppose there exist \(T > 0\) and \(c > 0\) such that, for any \(t \geq 0\) and \(z \in S\),

\[
\Pr[\tau(E) \geq T + t | Z_t = z] \leq 1 - c,
\]

holds. Then, \(E[\tau(E)] \leq \frac{T}{c}\).

**Proof.** Note that

\[
\Pr[\tau(E) \geq T + t | \tau(E) \geq t] \leq \Pr[\tau(E) \geq T + t | Z_t] \leq 1 - c.
\]

To obtain upper bounds for hitting, cover, and coalescing times, it suffices to prove that the corresponding stopping time satisfies the condition of Corollary B.2.
Finally, we introduce the following lemma, which we use in the proof of the edge-Markovian graph (Section 7).

**Lemma B.3.** Let $\tau$ be a stopping time of a sequence of random variables $(Z_i)_{i \geq 0}$. Let $0 = T_0 \leq T_1 \leq \cdots$ be a nondecreasing sequence and $\varepsilon$ be a positive constant. Suppose that $\Pr[\tau > T_\varepsilon | \tau > T_{\varepsilon-1}] \leq 1 - \varepsilon$ holds for all $\varepsilon \geq 1$. Then, $\mathbb{E}[\tau] \leq \sum_{\ell=0}^{\infty}(T_{\ell+1} - T_\ell)(1 - \varepsilon)^\ell$.

**Proof.** From the assumption,

$$
\Pr[\tau > T_\varepsilon] = \Pr[\tau > T_\varepsilon, \tau > T_{\varepsilon-1}] = \Pr[\tau > T_\varepsilon | \tau > T_{\varepsilon-1}] \Pr[\tau > T_{\varepsilon-1}]
$$

holds for any $\varepsilon \geq 1$. The first equality follows since $\tau$ is a stopping time. Hence, we obtain

$$
\mathbb{E}[\tau] = \sum_{\ell=0}^{\infty} \sum_{t=I_\ell}^{T_{\ell+1}-1} \Pr[\tau > t] \leq \sum_{\ell=0}^{\infty} \sum_{t=I_\ell}^{T_{\ell+1}-1} \Pr[\tau > T_\varepsilon] \leq \sum_{\ell=0}^{\infty}(T_{\ell+1} - T_\ell)(1 - \varepsilon)^\ell.
$$

\[\blacksquare\]

**APPENDIX C: OTHER TOOLS**

**Lemma C.1** (lemmas 4.24 and 4.25 in [2]). Suppose that $P$ is irreducible and reversible. Then, $\frac{1}{1 - \lambda_2(P)} \leq \text{hit}(P) \leq \frac{2}{\pi_{\min}(1 - \lambda_2(P))}$ holds.

**Lemma C.2** (The Chernoff inequality, see, e.g., theorem 1.10.21 in [20]). Let $X_1, X_2, \ldots, X_n$ be $n$ independent random variables taking values in $[0, 1]$. Let $X = \sum_{i=1}^{n} X_i$. Let $\mu^- \leq \mathbb{E}[X] \leq \mu^+$. Then,

$$
\Pr[X \geq (1 + \varepsilon)\mu^+] \leq \exp \left( -\frac{\min\{\varepsilon, \varepsilon^2\} \mu^+}{3} \right) \quad \text{for any } \varepsilon \geq 0,
$$

$$
\Pr[X \leq (1 - \varepsilon)\mu^-] \leq \exp \left( -\frac{\varepsilon^2 \mu^-}{2} \right) \quad \text{for any } 0 \leq \varepsilon \leq 1.
$$

The following is well known as the Cheeger inequality for reversible Markov chains.

**Lemma C.3** (see, e.g., theorem 13.10 in [34]). Let $P \in [0, 1]^{V \times V}$ be an irreducible and reversible transition matrix. For $S \subseteq V$, let $\pi(S) := \sum_{v \in S} \pi(v)$ and $Q(S) := \sum_{u \in S} \sum_{v \in S} \pi(u)P(u, v)$. Let $\Phi_* := \min_{S: 0 < \pi(S) \leq 1/2} \frac{Q(S)}{\pi(S)}$. Then, $\frac{\Phi_*^2}{2} \leq 1 - \lambda_*(P) \leq 2\Phi_*$.

**Lemma C.4** ([46]). Let $x : \mathbb{N} \to \mathbb{R}_{>0}$ be a positive nonincreasing function. Suppose that $x(t+1) \leq x(t) \left(1 - \frac{x(t)}{K}\right)$ holds for any $t \geq 0$. Then, $x(t) \leq 1$ holds for any $t \geq \frac{e^t}{e-1}K + \log x(0) + 1$.

**Proof.** First, we show that $x(a + \lfloor eK/x(a) \rfloor) \leq x(a)/e$ holds for any $a \geq 0$ by contradiction: suppose that $x(a + \ell) > x(a)/e$ holds for $\ell = \lfloor eK/x(a) \rfloor$. Since $x$ is nonincreasing and $x(t+1) \leq x(t) \left(1 - \frac{x(t)}{K}\right)$ holds, we have

$$
x(a + \ell) \leq x(a + \ell - 1) \left(1 - \frac{x(a + \ell - 1)}{K}\right) \leq x(a + \ell - 1) \left(1 - \frac{x(a + \ell - 1) - 1}{K}\right) \leq x(a + \ell - 1) \left(1 - \frac{x(a + \ell - 1)}{K}\right)
$$

\[\blacksquare\]
\[ \leq \cdots \leq x(a) \left( 1 - \frac{x(a + \epsilon)}{K} \right)^\epsilon \leq x(a) \exp \left( -\frac{\epsilon x(a + \epsilon)}{K} \right) \leq x(a) \exp \left( -\frac{\epsilon x(a)}{x(a)} \right) < x(a)/e. \]

This contradicts the assumption and we obtain the claim: \( x(a + \epsilon) \leq x(a)/e \) holds. Now, let \( \epsilon'(1) = \left[ \frac{eK}{x(a)} \right] \) and \( \epsilon'(i) = \left[ \frac{eK}{x(L(i-1))} \right] \) for \( i \geq 1 \), where \( L(i) := \sum_{j=1}^{i} \epsilon'(j) \) for \( i \geq 1 \) and \( L(0) := 0 \). From the above argument, we have \( x(L(i+1)) \leq x(L(i))/e \) for any \( i \geq 0 \). Let \( H \) be the first number with \( x(L(H)) \leq 1 \). Since \( 1 < x(L(H-1)) \leq x(L(H-2))/e \leq \cdots \leq x(0)/e^{H-1} \) holds, we have \( H - 1 < \log x(0) \). Hence, we have

\[
L(H) = \sum_{i=1}^{H} \epsilon'(i) = \sum_{i=0}^{H-1} \left[ \frac{eK}{x(L(i))} \right] \leq \sum_{i=0}^{H-1} \left[ \frac{eK}{e} \times x(L(H-1)) \right] \leq \sum_{i=0}^{H-1} \left[ \frac{eK}{e} \right] \leq H + \frac{e^2}{e-1} \leq \log x(0) + 1 + \frac{e^2}{e-1}.K.
\]

**APPENDIX D: DEFERRED PROOFS**

**Proof of Lemma 5.2.** Suppose that a walker \( a \) and \( b \) (with \( a < b \)) meet for the first time. In the \( Y \)-process, \( b \) must be in the coffin state. However, in the \( Z \)-process, \( b \) can be alive depending on the list of allowed killings \( A \). Such \( b \) can contribute +1 to \( |S(Z_j)| - |S(Y_j)| \). On the other hand, if such \( b \) kills two or more walkers in the \( Z \)-process and these victims are in \( S(Y_j) \), then \( b \) and the victims \( c_1, c_2, \ldots, c_r \) contribute −1 or less in \( |S(Z_j)| - |S(Y_j)| \). Our idea is to modify the \( Y \)-process such that \( c_1, \ldots, c_r \) follow the orbit of \( b \) (Figure D1). To state it more formally, for the list of allowed killings \( A = (A_t)_{t \geq 0} \), we define the \( A \)-coalescing random walks \( (C^A_t)_{t \geq 0} = ((C^A_t(1), \ldots, C^A_t(n))_{t \geq 0} \) as follows: Let \( (X_t(1), \ldots, X_t(n)) \) be \( n \) independent random walks. Then, define \( (C^A_t(1), \ldots, C^A_t(n))_{t \geq 0} \) as

\[
C^A_t(a) = \begin{cases} 
X_t(a) & \text{if } t < \tau^A(a), \\
X_t(b) & \text{otherwise,}
\end{cases}
\]

where \( b \) is the smallest index \( b \) such that \((b, a) \in A_{\tau^A(a)} \) and \( X_t(\omega)(a) = X_t(\omega)(b) \). In other words, \( C^A_t \) is the process such that a walker \( a \) follows another walker \( b \) if \( b \) kills \( a \). Observe that, if we simulate the \( Y \)-process using \( C^A_t \) as a seed instead of \( X_t \), the resulting process has the same distribution as \( Y \). Let \( Z \) be the \( Z \)-process constructed using \( X_t \) as a seed and \( Y' \) be the \( Y \)-process constructed using \( C^A_t \) as a seed.

We claim that, with probability 1 (over the choice of \( X \)), \( |S(Y'_t)| \leq |S(Z_t)| \) for all \( t \geq 0 \). Write \( S(Z_t) = \{b_1, b_2, \ldots, b_k\} \). Let \( K(b_1) = \{b_1, c_1, \ldots, c_r\} \) be the set of walkers that follow \( b_1 \) in the \( C^A \)-process. Note that \( (K(b_1), K(b_2), \ldots, K(b_k)) \) is a partition of \( [n] \). Since \( b_1, c_1, \ldots, c_r \) coalesce in the \( C^A \)-process, \( S(Y'_t) \cap K(b_1) = \{b_1\} \) if \( b_1 \in S(Y'_t) \) and \( S(Y'_t) \cap K(b_1) = \{\min\{c_1, \ldots, c_r\}\} \) if \( b_1 \notin S(Y'_t) \). Note that \( S(Y'_t) \cap K(b_i) = \emptyset \) if \( b_i \notin S(Y'_t) \) and \( \epsilon' = 0 \). Therefore, \( |S(Y'_t) \cap K(b_i)| \leq 1 \) holds for any \( b_i \in S(Z_t) \). Thus, \( |S(Y'_t)| = \sum_{i=1}^{k} |S(Y'_t) \cap K(b_i)| \leq k = |S(Z_t)| \) and we obtain the claim.
In the $Y$-process, the walker $b$ is killed by $a$ in the early stage. Such $b$ can kill two or more walkers in the $Z$-process. In this case, the $Y$-process has more survivors than the $Z$-process. To avoid this issue, we consider another process $Y'$ obtained by modifying $Y$ such that the victims $c_1, c_2$ follow the orbit of $b$. In the example above, the walker $c_1$ kills $c_2$ instead of $b$ and thus the number of survivors is not less than that of $Z$. 

![Diagram showing walkers in $Y$-process, $Z$-process, and $Y'$-process.](image)