CAMBRIDGE
UNIVERSITY PRESS

ARTICLE

Multiple random walks on graphs: mixing few to cover many†

Nicolás Rivera[1] ⓘ, Thomas Sauerwald[2]* ⓘ and John Sylvester[3] ⓘ

[1]Universidad de Valparaíso, Valparaíso, Chile, [2]University of Cambridge, Cambridge, UK, and [3]University of Liverpool, Liverpool, UK
*Corresponding author. Email: thomas.sauerwald@cl.cam.ac.uk

Abstract

Random walks on graphs are an essential primitive for many randomised algorithms and stochastic processes. It is natural to ask how much can be gained by running k multiple random walks independently and in parallel. Although the cover time of multiple walks has been investigated for many natural networks, the problem of finding a general characterisation of multiple cover times for *worst-case* start vertices (posed by Alon, Avin, Koucký, Kozma, Lotker and Tuttle in 2008) remains an open problem. First, we improve and tighten various bounds on the *stationary* cover time when k random walks start from vertices sampled from the stationary distribution. For example, we prove an unconditional lower bound of $\Omega((n/k) \log n)$ on the stationary cover time, holding for any n-vertex graph G and any $1 \le k = o(n \log n)$. Secondly, we establish the *stationary* cover times of multiple walks on several fundamental networks up to constant factors. Thirdly, we present a framework characterising *worst-case* cover times in terms of *stationary* cover times and a novel, relaxed notion of mixing time for multiple walks called the *partial mixing time*. Roughly speaking, the partial mixing time only requires a specific portion of all random walks to be mixed. Using these new concepts, we can establish (or recover) the *worst-case* cover times for many networks including expanders, preferential attachment graphs, grids, binary trees and hypercubes.

Keywords: multiple random walks; mixing time; cover time; Markov chains

2020 MSC codes: Primary: 05C81, 60J10, 60J20, Secondary: 68R10

1. Introduction

A random walk on a graph is a stochastic process that at each time step chooses a neighbour of the current vertex as its next state. The fact that a random walk visits every vertex of a connected, undirected graph in polynomial time was first used to solve the undirected $s - t$ connectivity problem in logarithmic space [4]. Since then random walks have become a fundamental primitive in the design of randomised algorithms which feature in approximation algorithms and sampling [44, 57], load balancing [35, 59], searching [25, 45], resource location [36], property testing [15, 38, 39], graph parameter estimation [7, 14] and biological applications [8, 10, 27].

The fact that random walks are local and memoryless (Markov property) ensures they require very little space and are relatively unaffected by changes in the environment, for example, dynamically evolving graphs or graphs with edge failures. These properties make random walks a natural candidate for parallelisation, where running parallel walks has the potential of lower

---

time overheads. One early instance of this idea are space-time trade-offs for the undirected $s-t$ connectivity problem [9, 22]. Other applications involving multiple random walks are sublinear algorithms [16], local clustering [6, 60] or epidemic processes on networks [41, 55].

Given the potential applications of multiple random walks in algorithms, it is important to understand fundamental properties of multiple random walks. The *speed-up*, first introduced in [5], is the ratio of the worst-case cover time by a single random walk to the cover time of k parallel walks. Following [5] and subsequent works [19, 20, 31, 37, 53, 58] our understanding of when and why a speed-up is present has improved. In particular, various results in [5, 19, 20] establish that as long as the lengths of the walks are not smaller than the mixing, the speed-up is linear in k. However, there are still many challenging open problems, for example, understanding the effect of different start vertices or characterising the magnitude of speed-up in terms of graph properties, a problem already stated in [5]: '... *which leads us to wonder whether there is some other property of a graph that characterises the speed-up achieved by multiple random walks more crisply than hitting and mixing times*'. Addressing the previous questions, we introduce new quantities and couplings for multiple random walks, that allow us to improve the state-of-the-art by refining, strengthening or extending results from previous works.

While there is an extensive body of research on the foundations of (single) random walks (and Markov chains), it seems surprisingly hard to transfer these results and develop a systematic theory of *multiple* random walks. One of the reasons is that processes involving multiple random walks often lead to questions about *short* random walks, for example, shorter than the mixing time. Such short walks may also arise in applications including generating random walk samples in massively parallel systems [40, 57], or in applications where random walk steps are expensive or subject to delays (e.g. when crawling social networks like Twitter [14]). The challenge of analysing *short* random walks (shorter than mixing or hitting time) has been mentioned not only in the area of multiple cover times (e.g. [19, Sec. 6]), but also in the context of concentration inequalities for random walks [43, p. 863] and property testing [16].

1.1. Our contribution

Our first set of results provide several tight bounds on $t_{\mathsf{cov}}^{(k)}(\pi)$ in general (connected) graphs, where $t_{\mathsf{cov}}^{(k)}(\pi)$ is the expected time for each vertex to be visited by at least one of k independent walks each started from a vertex independently sampled from the stationary distribution π.

The main findings of Section 3 include:

- Proving general bounds of $\mathcal{O}\left(\left(\frac{|E|}{kd_{\min}}\right)^2 \log^2 n\right)$, $\mathcal{O}\left(\frac{\max_{v \in V} \mathbb{E}_\pi[\tau_v]}{k} \log n\right)$ and $\mathcal{O}\left(\frac{|E| \log n}{kd_{\min}\sqrt{1-\lambda_2}}\right)$ on $t_{\mathsf{cov}}^{(k)}(\pi)$, where $d_{\min}$ is the minimum degree, $\mathbb{E}_\pi[\tau_v]$ is the hitting time of $v \in V$ from a stationary start vertex and λ_2 is the second largest eigenvalue of the transition matrix of the walk. All three bounds are tight for certain graphs. The first bound improves over [9], the second result is a Matthew's type bound for multiple random walks, and the third yields tight bounds for non-regular expanders such as preferential attachment graphs.

- We prove that for any graph G and $1 \leq k = o(n \log n)$, $t_{\mathsf{cov}}^{(k)}(\pi) = \Omega((n/k) \log n)$. Weaker versions of this bound were obtained in [20], holding only for certain values of k or under additional assumptions on the mixing time. Our result matches the fundamental $\Omega(n \log n)$ lower bound for single random walks ($k = 1$) [3, 21], and generalises it in the sense that total amount of work by all k stationary walks together for covering is always $\Omega(n \log n)$. We establish the $\Omega((n/k) \log n)$ bound by reducing the multiple walk process to a single, reversible Markov chain and applying a general lower bound on stationary cover times [3].

- A technical tool that provides a bound on the lower tail of the cover time by k walks from stationary for graphs with a large and (relatively) symmetric set of hard to hit vertices (Lemma 3.9). When applied to the 2d torus and binary trees this yields a tight lower bound.

In Section 4, we introduce a novel quantity for multiple walks we call *partial mixing*. Intuitively, instead of mixing all (or at least half) of the k walks, we only need to mix a specified number $\tilde{k}$ of them. We put this idea on a more formal footing and prove min-max theorems which relate worst-case cover times $t_{\text{cov}}^{(k)}$ to partial mixing times $t_{\text{mix}}^{(\tilde{k},k)}$ and stationary cover times:

- For any graph G and any $1 \leq k \leq n$, we prove that:

$$t_{\text{cov}}^{(k)} \leq 12 \cdot \min_{1 \leq \tilde{k} < k} \max\left(t_{\text{mix}}^{(\tilde{k},k)}, t_{\text{cov}}^{(\tilde{k})}(\pi)\right).$$

For now, we omit details such as the definition of the partial mixing time $t_{\text{mix}}^{(\tilde{k},k)}$ as well as some min-max characterisations that serve as lower bounds (these can be found in Section 4). Intuitively these characterisations suggest that for any number of walks k, there is an 'optimal' choice of $\tilde{k}$ so that one first waits until $\tilde{k}$ out of the k walks are mixed, and then considers only these $\tilde{k}$ stationary walks when covering the remainder of the graph.

This argument involving mixing only some walks extends and generalises prior results that involve mixing all (or at least a constant portion) of the k walks [5, 19, 20]. Previous approaches only imply a linear speed-up as long as the lengths of the walks are not shorter than the mixing time of a *single* random walk. In contrast, our characterisation may still yield tight bounds on the cover time for random walks that are much shorter than the mixing time.

In Section 5, we demonstrate how our insights can be used on several well-known graph classes. As a first step, we determine their stationary cover times; this is based on our bounds from Section 3. Secondly, we derive lower and upper bounds on the partial mixing times. Finally, with the stationary cover times and partial mixing times at hand, we can apply the characterisations from Section 4 to infer lower and upper bounds on the worst-case stationary times. For some of those graphs the worst-case cover times were already known before, while for, e.g., binary trees and preferential attachment graphs, our bounds are new.

- For the graph families of binary trees, cycles, d-dim. tori ($d = 2$ and $d \geq 3$), hypercubes, cliques and (possibly non-regular) expanders we determine the cover time up to constants, for both worst-case and stationary start vertices (see Table 1 for the quantitative results).

We believe that this new methodology constitutes some progress towards the open question of Alon, Avin, Koucký, Kozma, Lotker, and Tuttle [5] concerning a characterisation of worst-case cover times.

1.2. Novelty of our techniques

While a lot of the proof techniques in previous work [5, 19, 20, 58] are based on direct arguments such as mixing time (or relaxation time), our work introduces a number of new methods which, to the best of our knowledge, have not been used in the analysis of cover time of multiple walks before. In particular, one important novel concept is the introduction of the so-called *partial mixing time*. The idea is that instead of waiting for all (or a constant portion of) k walks to mix, we can just mix some $\tilde{k} \leq k$ walks to reap the benefits of coupling these $\tilde{k}$ walks to stationary walks. This then presents a delicate balancing act where one must find an optimal $\tilde{k}$ minimising the overall bound on the cover time, for example in expanders the optimal $\tilde{k}$ is linear in k whereas in binary trees it is approximately $\sqrt{k}$, and for the cycle it is roughly $\log k$. This turning point reveals

Table 1. All results above are $\Theta(\,\cdot\,)$, that is bounded above and below by a multiplicative constant, apart from the mixing time of expanders which is only bounded from above.

Graph family	Cover t_{cov}	Hitting t_{hit}	Mixing t_{mix}	k-Cover Time, where $2 \leq k \leq n$	
				worst-case $t_{\mathrm{cov}}^{(k)}$	From π^k, $t_{\mathrm{cov}}^{(k)}(\pi)$
Cycles	n^2	n^2	n^2	$\dfrac{n^2}{\log k}$	$\left(\dfrac{n}{k}\right)^2 \log^2 k$
Binary trees	$n \log^2 n$	$n \log n$	n	$(n/k) \log^2 n$ **if** $k \leq \log^2 n$. $(n/\sqrt{k}) \log n$ **if** $k \geq \log^2 n$.	$\dfrac{n \log n}{k} \log\left(\dfrac{n \log n}{k}\right)$
2-Dim. Tori	$n \log^2 n$	$n \log n$	n	$(n/k) \log^2 n$ **if** $k \leq \log^2 n$. $\dfrac{n}{\log(k/\log^2 n)}$ **if** $k \geq \log^2 n$.	$\dfrac{n \log n}{k} \log\left(\dfrac{n \log n}{k}\right)$
d-Dim. Tori $d \geq 3$	$n \log n$	n	$n^{2/d}$	$(n/k) \log n$ **if** $k \leq n^{1-2/d} \log n$. $\dfrac{n^{2/d}}{\log(k/(n^{1-2/d} \log n))}$ **if** $k \geq n^{1-2/d} \log n$.	$\dfrac{n}{k} \log n$
Hypercubes	$n \log n$	n	$\log n \log \log n$	$(n/k) \log n$ **if** $k \leq n/\log \log n$. $\log n \log \log n$ **if** $k \geq n/\log \log n$.	$\dfrac{n}{k} \log n$
Expanders	$n \log n$	n	$\mathcal{O}(\log n)$	$\dfrac{n}{k} \log n$	$\dfrac{n}{k} \log n$
PA, $m \geq 2$	$n \log n$	n	$\mathcal{O}(\log n)$	$\dfrac{n}{k} \log n$	$\dfrac{n}{k} \log n$
Barbells	n^2	n^2	n^2	n^2/k	$\dfrac{2^{-k} n^2}{k} + \dfrac{n \log n}{k}$

Notes: PA above is the preferential attachment process where each vertex has m initial links, the results hold w.h.p., see [12, 48]. Cells shaded in Green are new results proved in this paper with the exception that for $k = \mathcal{O}(\log n)$ upper bounds on the stationary cover time for binary trees, expanders and preferential attachment graphs can be deduced from general bounds for the worst-case cover time in [5]. Cells shaded Gray in the second to last column are known results we re-prove in this paper using our partial mixing time results, for the 2-dim grid we only re-prove upper bounds. References for the second to last column are given in Section 5, except for the barbell, see [19], p. 2]. The barbell consists of two cliques on $n/2$ vertices connected by single edge; we include this in the table as an interesting example where the speed-up by stationary walks is exponential in k. All other results for single walks can be found in [2], for example.

something about the structure of the graph and our results relating partial mixing to hitting time of sets helps one find this. Another tool we frequently use is a reduction to random walks with geometric resets, similar to a PageRank process, which allows us to relate multiple walks from stationary to a single reversible Markov chain.

2. Notation and preliminaries

Throughout $G = (V, E)$ will be a finite undirected, connected graph with $n = |V|$ vertices and $m = |E|$ edges. For $v \in V(G)$ let $d(v) = |\{u \in V : uv \in E(G)\}|$ denote the degree of v and $d_{\min} = \min_{v \in V} d(v)$ and $d_{\max} = \max_{v \in V} d(v)$ denote the minimum and maximum degrees in G respectively. For any $k \geq 1$, let $X_t = \left(X_t^{(1)}, \ldots, X_t^{(k)} \right)$ be the *multiple random walk* process where each $X_t^{(i)}$ is an independent random walk on G. Let

$$\mathbb{E}_{u_1, \ldots, u_k}[\,\cdot\,] = \mathbb{E}[\,\cdot\, \mid X_0 = (u_1, \ldots, u_k)]$$

denote the conditional expectation where, for each $1 \leq i \leq k$, $X_0^{(i)} = u_i \in V$ is the start vertex of the ith walk. Unless mentioned otherwise, walks will be *lazy*, that is, at each step the walk stays at its current location with probability $1/2$, and otherwise moves to a neighbour chosen uniformly at random. We let the random variable $\tau_{\mathsf{cov}}^{(k)}(G) = \inf \left\{ t : \bigcup_{i=0}^{t} \{ X_i^{(1)}, \ldots, X_i^{(k)} \} = V \right\}$ be the first time every vertex of the graph has been visited by some walk $X_t^{(i)}$. For $u_1, \ldots, u_k \in V$ let

$$t_{\mathsf{cov}}^{(k)}((u_1, \ldots, u_k), G) = \mathbb{E}_{u_1, \ldots, u_k} \left[\tau_{\mathsf{cov}}^{(k)}(G) \right], \qquad t_{\mathsf{cov}}^{(k)}(G) = \max_{u_1, \ldots, u_k \in V} t_{\mathsf{cov}}^{(k)}((u_1, \ldots, u_k), G)$$

denote the cover time of k walks from $(u_1, \ldots, u_k)$ and the cover time of k walks from worst-case start positions respectively. For simplicity, we drop G from the notation if the underlying graph is clear from the context. We shall use π to denote the stationary distribution of a single random walk on a graph G. For $v \in V$ this is given by $\pi(v) = \frac{d(v)}{2m}$ which is the degree over twice the number of edges. Let $\pi_{\min} = \min_{v \in V} \pi(v)$ and $\pi_{\max} = \max_{v \in V} \pi(v)$. We use π^k, which is a distribution on V^k given by the product measure of π with itself, to denote the stationary distribution of a multiple random walk. For a probability distribution μ on V let $\mathbb{E}_{\mu^k}[\,\cdot\,]$ denote the expectation with respect to k walks where each start vertex is sampled independently from μ and

$$t_{\mathsf{cov}}^{(k)}(\mu, G) = \mathbb{E}_{\mu^k} \left[\tau_{\mathsf{cov}}^{(k)}(G) \right].$$

In particular, $t_{\mathsf{cov}}^{(k)}(\pi, G)$ denotes the expected cover time from k independent stationary start vertices. For a set $S \subseteq V$ (if $S = \{v\}$ is a singleton set we use $\tau_v^{(k)}$, dropping brackets), we define

$$\tau_S^{(k)} = \inf \left\{ t : \text{there exists } 1 \leq i \leq k \text{ such that } X_t^{(i)} \in S \right\}$$

as the first time at least one vertex in the set S is visited by any of the k independent random walks. Let

$$t_{\mathsf{hit}}^{(k)}(G) = \max_{u_1, \ldots, u_k \in V} \max_{v \in V} \mathbb{E}_{u_1, \ldots, u_k} \left[\tau_v^{(k)} \right]$$

be the worst-case vertex to vertex hitting time. When talking about a single random walk we drop the (1) index, that is, $t_{\mathsf{cov}}^{(1)}(G) = t_{\mathsf{cov}}(G)$; we also drop G from the notation when the graph is clear. If we wish the graph G to be clear we shall also use the notation $\mathbb{P}_{u,G}[\,\cdot\,]$ and $\mathbb{E}_{u,G}[\cdot]$. For $t \geq 0$ we let $N_v(t)$ denote the number of visits to $v \in V$ by a single random walk up-to time t.

For a single random walk X_t with stationary distribution π and $x \in V$, let $d_{\mathrm{TV}}(t)$ and $s_x(t)$ be the total variation and separation distances for X_t given by

$$d_{\mathrm{TV}}(t) = \max_{x \in V} ||P_{x,\cdot}^t - \pi||_{\mathrm{TV}}, \qquad \text{and} \qquad s_x(t) = \max_{y \in V} \left[1 - \frac{P_{x,y}^t}{\pi(y)} \right],$$

where $P_{x,\cdot}^t$ is the t-step probability distribution of a random walk starting from x and, for probability measures μ, ν, $||\mu - \nu||_{\mathrm{TV}} = \frac{1}{2} \sum_{x \in V} |\mu(x) - \nu(x)|$ is the total variation distance.

Let $s(t) = \max_{x \in V} s_x(t)$, then for $0 < \varepsilon \leq 1$ the mixing and separation times [42, (4.32)] are

$$t_{\mathsf{mix}}(\varepsilon) = \inf\{t : d_{\mathsf{TV}}(t) \leq \varepsilon\} \qquad \text{and} \qquad t_{\mathsf{sep}}(\varepsilon) = \inf\{t : s(t) \leq \varepsilon\},$$

and $t_{\mathsf{mix}} := t_{\mathsf{mix}}(1/4)$ and $t_{\mathsf{sep}} := t_{\mathsf{sep}}(1/e)$. A strong stationary time (SST) σ, see [42, Ch. 6] or [1], is a randomised stopping time for a Markov chain Y_t on V with stationary distribution π if

$$\mathbb{P}_u\big[Y_\sigma = v \mid \sigma = k\big] = \pi(v) \qquad \text{for any } u, v \in V \text{ and } k \geq 0.$$

Let $t_{\mathsf{rel}} = \frac{1}{1-\lambda_2}$ be the relaxation time of G, where λ_2 is the second largest eigenvalue of the transition matrix of the (lazy) random walk on G. A sequence of graphs (G_n) is a *sequence of expanders* (or simply an *expander*) if $\liminf_{n \to \infty} 1 - \lambda_2(G_n) > 0$, which implies that $t_{\mathsf{rel}} = \Theta(1)$.

For random variables Y, Z we say that Y dominates Z ($Y \succeq Z$) if $\mathbb{P}[Y \geq x] \geq \mathbb{P}[Z \geq x]$ for all real x. Finally, we shall use the following inequality [50, Proposition B.3]:

$$(1 + x/n)^n \geq e^x(1 - x^2/n) \quad \text{for } n \geq 1, |x| \leq n. \tag{1}$$

3. Multiple stationary cover times

We shall state our general upper and lower bound results for multiple walks from stationary in Sections 3.1 and 3.2 before proving these results in Sections 3.3 and 3.4, respectively.

3.1. Upper bounds

Broder, Karlin, Raghavan and Upfal [9] showed that for any graph G (with $m = |E|$) and $k \geq 1$,

$$t_{\mathsf{cov}}^{(k)}(\pi) = \mathcal{O}\left(\left(\frac{m}{k}\right)^2 \log^3 n\right).$$

We first prove a general bound which improves this bound by a multiplicative factor of $d_{\mathsf{min}}^2 \log n$ which may be $\Theta(n^2 \log n)$ for some graphs.

Theorem 3.1. *For any graph G and any $k \geq 1$,*

$$t_{\mathsf{cov}}^{(k)}(\pi) = \mathcal{O}\left(\left(\frac{m}{k d_{\mathsf{min}}}\right)^2 \log^2 n\right).$$

This bound is tight for the cycle if $k = n^{\Theta(1)}$, see Theorem 5.2. Theorem 3.1 is proved by relating the probability a vertex v is not hit up to a certain time t to the expected number of returns to v by a walk of length t from v and applying a bound by Oliveira and Peres [51].

The next bound is analogous to Matthew's bound [2, Theorem 2.26] for the cover time of single random walks from worst-case, however it is proved by a different method.

Theorem 3.2. *For any graph G and any $k \geq 1$, we have*

$$t_{\mathsf{cov}}^{(k)}(\pi) = \mathcal{O}\left(\frac{\max_{v \in V} \mathbb{E}_\pi[\tau_v] \log n}{k}\right).$$

This bound is tight for many graphs, see Table 1. Following this paper the stronger bound $t_{\mathsf{cov}}^{(k)}(\pi) = \mathcal{O}(t_{\mathsf{cov}}/k)$ was recently proved by Hermon and Sousi [30]. A version of Theorem 3.2 for

$t_{\text{cov}}^{(k)}$ was established in [5] provided $k = \mathcal{O}(\log n)$, the restriction on k is necessary (for worst-case) as witnessed by the cycle. Theorem 3.2 also gives the following explicit bound.

Corollary 3.3. *For any graph G and any $k \geq 1$, we have*

$$t_{\text{cov}}^{(k)}(\pi) = \mathcal{O}\left(\frac{m}{k d_{\min}} \sqrt{t_{\text{rel}}} \log n\right).$$

Proof. Use $\max_{v \in V} \mathbb{E}_\pi[\tau_v] \leq 20 m \sqrt{t_{\text{rel}} + 1}/d_{\min}$ from [51, Theorem 1] in Theorem 3.2. □

Notice that, for any $k \geq 1$, this bound is tight for any expander with $d_{\min} = \Omega(m/n)$, such as preferential attachment graphs with $m \geq 2$, see Theorems 5.14 and 5.13. We now establish some bounds for classes of graphs determined by the return probabilities of random walks.

Lemma 3.4. *Let G be any graph satisfying $\pi_{\min} = \Omega(1/n)$, $t_{\text{mix}} = \mathcal{O}(n)$ and $\sum_{i=0}^{t} P_{vv}^i = \mathcal{O}(1 + t\pi(v))$ for any $t \leq t_{\text{rel}}$ and $v \in V$. Then for any $1 \leq k \leq n$,*

$$t_{\text{cov}}^{(k)}(\pi) = \Theta\left(\frac{n}{k} \log n\right).$$

The bound above applies to a broad class of graphs including the hypercube and high dimensional grids. The following bound holds for graphs with sub-harmonic return times, this includes binary trees and 2d-grid/torus.

Lemma 3.5. *Let G be any graph satisfying $\pi_{\max} = \mathcal{O}(\pi_{\min})$, $\sum_{i=0}^{t} P_{v,v}^i = \mathcal{O}(1 + \log t)$ for any $t \leq t_{\text{rel}}$ and $v \in V$, and $t_{\text{mix}} = \mathcal{O}(n)$. Then for any $1 \leq k \leq (n \log n)/3$,*

$$t_{\text{cov}}^{(k)}(\pi) = \mathcal{O}\left(\frac{n \log n}{k} \log\left(\frac{n \log n}{k}\right)\right).$$

3.2. Lower bounds

Generally speaking, lower bounds for random walks seem to be somewhat more challenging to derive than upper bounds. In particular, the problem of obtaining a tight lower bound for the cover time of a simple random walk on an undirected graph was open for many years [2]. This was finally resolved by Feige [21] who proved $t_{\text{cov}} \geq (1 - o(1))n \log n$ (this bound was known up-to constants for stationary Markov chains by Aldous [3]). We prove a generalisation of this bound, up to constants, that holds for k random walks starting from stationary (thus also for worst-case).

Theorem 3.6. *There exists a constant $c > 0$ such that for any graph G and $1 \leq k \leq c \cdot n \log n$,*

$$t_{\text{cov}}^{(k)}(\pi) \geq c \cdot \frac{n}{k} \cdot \log n.$$

We remark that in this section all results hold (and are proven) for *non-lazy* random walks, which by stochastic domination implies that the same result also holds for *lazy* random walks. Theorem 3.6 is tight, uniformly for all $1 \leq k \leq n$, for the hypercube, expanders and high-dimensional tori, see Theorem 5.14. We note that [20] proved this bound for any start vertices under the additional assumption that $k \geq n^\epsilon$, for some constant $\varepsilon > 0$. One can track the constants in the proof of Theorem 3.6 and show that $c > 2 \cdot 10^{-11}$, we have not optimised this but note that $c \leq 1$ must hold in either condition of Theorem 3.6 due to the complete graph.

To prove this result, we introduce the *geometric reset graph*, which allows us to couple the multiple random walk to a single walk to which we can apply a lower bound by Aldous [3]. The geometric reset graph is a small modification to a graph G which gives an edge-weighted graph $\widehat{G}(x)$ such that the simple random walk on $\widehat{G}(x)$ emulates a random walk on G with $\text{Geo}(x)$ *resets* to stationarity, where $\text{Geo}(x)$ is a geometric random variable with expectation $1/x$. This is achieved

by adding a dominating vertex z to the graph and weighting edges from z so that after z is visited the walk moves to a vertex $v \in V(G)$ proportional with probability $\pi(v)$.

Definition 3.7. *(The Geometric Reset Graph $\widehat{G}(x)$).* For any graph G the undirected, edge-weighted graph $\widehat{G}(x)$, where $0 < x \leq 1$, consists of all vertices $V(G)$ and one extra vertex z. All edges from G are included with edge-weight $1 - x$. Further, z is connected to each vertex $u \in G$ by an edge with edge-weight $x \cdot d(u)$, where $d(u)$ is the degree of vertex u in G.

Given a graph with edge weights $\{w_e\}_{e \in E}$ the probability a *non-lazy* random walk moves from u to v is given by $w_{uv} / \sum_{w \in V} w_{uw}$. Thus the walk on $\widehat{G}(x)$ behaves as a random walk in G, apart from that in any step, it may move to the extra vertex z with probability $\frac{xd(u)}{xd(u)+(1-x)d(u)} = x$. Once the walk is at z it moves back to a vertex $u \in V \setminus \{z\}$ with probability

$$P_{z,u} = \frac{x \cdot d(u)}{\sum_{v \in V \setminus \{z\}} x \cdot d(v)} = \pi(u).$$

Hence, the stationary distribution $\widehat{\pi}$ of the random walk on $\widehat{G}(x)$ is proportional to π on $V(G)$, and for the extra vertex z we have

$$\widehat{\pi}(z) = \frac{\sum_{u \in V} xd(u)}{\sum_{u \in V}(1-x)d(u) + \sum_{u \in V} xd(u)} = \frac{x}{(1-x)+x} = x.$$

Using the next lemma, we can then obtain bounds on the multiple stationary cover time by simply bounding the cover time in the augmented graph $\widehat{G}(x)$ for some x.

Lemma 3.8. *Let G be any graph, $k \geq 1$ and $x = Ck/T$ where $C > 30$ and $T \geq 5Ck$. Then,*

$$\mathbb{P}_{\pi^k, G}\left[\tau_{\mathsf{cov}}^{(k)} > \frac{T}{10Ck}\right] > \mathbb{P}_{\widehat{\pi}, \widehat{G}(x)}[\tau_{\mathsf{cov}} > T] - \exp\left(-\frac{Ck}{50}\right).$$

The result in Lemma 3.8 will also be used later in this work to prove a lower bound for the stationary cover time of the binary tree and two-dimensional grid when k is small.

The next result we present utilises the second moment method to obtain a lower bound which works well for $k = n^{\Theta(1)}$ walks on symmetric (e.g. transitive) graphs. In particular, we apply this to get tight lower bounds for cycles, the 2-dim. torus and binary trees in Section 5.

Lemma 3.9. *For any graph G, subset $S \subseteq V$ and $t \geq 1$, let $p = \max_{u \in S} \mathbb{P}_\pi[\tau_u \leq t]$. Then, for any $k \geq 2$ satisfying $2p^2 k < 1$, we have*

$$\mathbb{P}_{\pi^k}\left[\tau_{\mathsf{cov}}^{(k)} \leq t\right] \leq \frac{4kp^2 e^{2kp}}{|S| \min_{v \in S} \pi(v)}.$$

3.3. Proofs of upper bounds

We begin by stating some basic facts.

Lemma 3.10. *Let $N_v(t)$ be the number of visits to $v \in V$ by a t-step walk, $t \geq 1$, then*

(i) $\mathbb{P}_{\pi^k}\left[\tau_{\mathsf{cov}}^{(k)} \geq t\right] \leq \sum_{v \in V} \exp(-k\mathbb{P}_\pi[N_v(t) \geq 1])$,

(ii) $\dfrac{(t+1)\pi(v)}{\sum_{i=0}^{t} P_{v,v}^i} \leq \mathbb{P}_\pi[N_v(t) \geq 1] = \dfrac{\mathbb{E}_\pi[N_v(t)]}{\mathbb{E}_\pi[N_v(t) \mid N_v(t) \geq 1]} \leq \dfrac{2(t+1)\pi(v)}{\sum_{i=0}^{t/2} P_{v,v}^i}$,

(iii) $\displaystyle\sum_{i=0}^{t} P_{u,u}^i \leq \frac{e}{e-1}\left(\sum_{i=0}^{\lceil t_{\mathsf{rel}}\rceil - 1} P_{u,u}^i - \lceil t_{\mathsf{rel}} \rceil \pi(u)\right) + (t+1)\pi(u).$

Proof. For Item (i), by independence of the walks and the union bound we have

$$\mathbb{P}_{\pi^k}\left[\tau_{\mathsf{cov}}^{(k)} \geq t\right] \leq \sum_{v \in V} \prod_{i=1}^{k} \left(1 - \mathbb{P}_{\pi}\left[N_v^{(i)}(t) \geq 1\right]\right) \leq \sum_{v \in V} \exp\left(-k\mathbb{P}_{\pi}[N_v(t) \geq 1]\right).$$

For Item (ii), since $N_v(t)$ is a non-negative integer, we have

$$\mathbb{E}_{\pi}[N_v(t)] = \mathbb{E}_{\pi}[N_v(t) \mid N_v(t) \geq 1] \cdot \mathbb{P}_{\pi}[N_v(t) \geq 1].$$

For the two inequalities first observe that $\mathbb{E}_{\pi}[N_v(t)] = (t+1)\pi(v)$. Now, conditional on a walk first hitting v at time s we have $N_v(t) = \sum_{i=0}^{t-s} P_{v,v}^t$. The first inequality in Item (ii) then follows since $\mathbb{E}_{\pi}[N_v(t) \mid N_v(t) \geq 1] \leq \sum_{i=0}^{t} P_{v,v}^i$. For the last inequality in Item (ii), observe that, by reversibility, for any $t \geq 1$

$$\mathbb{P}_{\pi}\left[N_v\left(\lceil \tfrac{t-1}{2} \rceil\right) \geq 1, \, N_v(t) \geq 1\right] = \sum_{v_0 v_1 \cdots v_t : v \in \bigcup_{i=0}^{\lceil \frac{t-1}{2} \rceil} \{v_i\}} \pi(v_0) P_{v_0,v_1} \cdots P_{v_{t-1},v_t}$$

$$= \sum_{v_0 v_1 \cdots v_t : v \in \bigcup_{i=0}^{\lceil \frac{t-1}{2} \rceil} \{v_i\}} \pi(v_t) P_{v_t,v_{t-1}} \cdots P_{v_1,v_0}$$

$$= \mathbb{P}_{\pi}\left[N_v(t) - N_v\left(\lfloor \tfrac{t+1}{2} \rfloor - 1\right) \geq 1, \, N_v(t) \geq 1\right].$$

Since $\{N_v(t) \geq 1\} = \{N_v\left(\lceil \tfrac{t-1}{2} \rceil\right) \geq 1\} \cup \{N_v(t) - N_v\left(\lfloor \tfrac{t+1}{2} \rfloor - 1\right) \geq 1\}$ we have

$$\mathbb{P}_{\pi}\left[N_v\left(\lceil \tfrac{t-1}{2} \rceil\right) \geq 1 \mid N_v(t) \geq 1\right] \geq 1/2.$$

Thus $\mathbb{E}_{\pi}[N_v(t) \mid N_v(t) \geq 1] \geq (1/2) \cdot \sum_{i=0}^{\lceil \frac{t-1}{2} \rceil} P_{v,v}^i$ as claimed.

Finally, for Item (iii), the proof of [51, Lemma 1] shows that

$$\sum_{i=0}^{t} P_{u,u}^i - (t+1)\pi(u) \leq \frac{e}{e-1} \left(\sum_{i=0}^{\lceil t_{\mathsf{rel}} \rceil - 1} P_{u,u}^i - \lceil t_{\mathsf{rel}} \rceil \pi(u)\right),$$

rearranging gives the result. $\square$

Recall $a \wedge b$ denotes $\min(a, b)$. We shall use the following result of Oliveira and Peres [51, Theorem 2] for lazy walks: for any $v \in V$ and $t \geq 0$ we have

$$P_{v,v}^t - \pi(v) \leq \frac{10d(v)}{d_{\min}} \left(\frac{1}{\sqrt{t+1}} \wedge \frac{\sqrt{t_{\mathsf{rel}} + 1}}{t+1}\right). \tag{2}$$

Note that we prove Theorem 3.1 for lazy walks however this also applies to non-lazy walks as the cover time of a lazy walk stochastically dominates that of a non-lazy walk.

Proof of Theorem 3.1. Recall that we aim to prove $t_{\mathsf{cov}}^{(k)}(\pi) = \mathcal{O}\left(\left(\frac{m}{kd_{\min}}\right)^2 \log^2 n\right)$. To begin observe that if $k \geq 10(m/d_{\min})\log n$ then the probability any vertex u is unoccupied at time 0 is $(1 - \pi(u))^k \leq e^{-\pi(u)k} \leq e^{-d_{\min}k/(2m)}$. For any graph $t_{\mathsf{cov}} \leq 16mn/d_{\min} \leq 16n^3$ by [34, Theorem 2], thus we have

$$\mathbb{E}_{\pi^k}\left[\tau_{\mathsf{cov}}^{(k)}\right] \leq 16n^3 \cdot ne^{-d_{\min}k/(2m)} = o\left(\left(m/(kd_{\min})\right)^2 \log^2 n\right).$$

It follows that we can assume $k \leq 10(m/d_{\min})\log n$ for the remainder of the proof.

We will apply Lemma 3.10 to bound $\mathbb{P}_{\pi^k}\left[\tau_{\mathsf{cov}}^{(k)} \geq t\right]$ for $t \geq 1$. To begin, by (2) we have

$$\sum_{i=0}^{t} P_{u,u}^i \leq \frac{10d(u)}{d_{\min}} \sum_{i=0}^{t} \frac{1}{\sqrt{i+1}} + (t+1)\pi(u) \leq \frac{10d(u)}{d_{\min}} \left(\int_1^t \frac{1}{\sqrt{x}}\,\mathrm{d}x + 1\right) + (t+1)\pi(u).$$

Now, since $d_{\min} \leq d(u)$, $\pi(u) \leq 1$ and $t \geq 1$, we have

$$\sum_{i=0}^{t} P_{u,u}^i \leq \frac{10d(u)}{d_{\min}} \left(\left[2\sqrt{t} - 2\sqrt{1}\right] + 1\right) + t\pi(u) + 1 \leq \frac{20d(u)}{d_{\min}}\sqrt{t} + t\pi(u).$$

Thus, by Lemma 3.10 (ii) and dividing each term by a factor of $\pi(u) = d(u)/2m$ we have

$$\mathbb{P}_{\pi}[N_v(t) \geq 1] \geq \frac{t}{(40m/d_{\min})\sqrt{t} + t}.$$

We now define $t^* = \left(\frac{300m\log n}{kd_{\min}}\right)^2$. Firstly if $k \geq 10\log n$, then for any $t \geq t^*$ we have

$$k\mathbb{P}_{\pi}[N_v(t) \geq 1] \geq \frac{k\left(\frac{300m\log n}{kd_{\min}}\right)^2}{(40m/d_{\min}) \cdot \frac{300m\log n}{kd_{\min}} + \left(\frac{300m\log n}{kd_{\min}}\right)^2} \geq \frac{300\log n}{40 + 300/10} > 4\log n.$$

Now, since $t_{\mathsf{cov}} \leq 16n^3$, Lemma 3.10(i) gives $\mathbb{E}_{\pi^k}\left[\tau_{\mathsf{cov}}^{(k)}\right] \leq t^* + o(n^{-3}) \cdot 16n^3 = \mathcal{O}\left(\left(\frac{m}{kd_{\min}}\right)^2\log^2 n\right)$.

Finally, if $k \leq 10\log n$ then $t^* \geq (300m/d_{\min})^2/100 = 900(m/d_{\min})^2$. However, $t_{\mathsf{cov}} \leq 16mn/d_{\min} \leq 32m^2/d_{\min}^2$ by [34, Theorem 2]. Thus, $\mathbb{E}_{\pi^k}\left[t_{\mathsf{cov}}^{(k)}\right] \leq t^*$ holds as claimed. □

Proof of Theorem 3.2. We consider first the case $k < 8\log_2 n$. To begin, for any pair u, v it holds that $\mathbb{E}_u[\tau_v] \leq 2\max_{w \in V} \mathbb{E}_{\pi}[\tau_w]$ by [42, Lemma 10.2]. Thus by Markov's inequality

$$\mathbb{P}_u\left[\tau_v \geq 4\left\lceil \max_{w \in V} \mathbb{E}_{\pi}[\tau_w]\right\rceil\right] \leq \frac{\mathbb{E}_u[\tau_v]}{4\left\lceil \max_{w \in V} \mathbb{E}_{\pi}[\tau_w]\right\rceil} \leq \frac{1}{2},$$

for any $v \in V$. Then, the Markov property yields

$$\mathbb{P}_u\left[\tau_v \geq 20\left\lceil \frac{\log_2 n}{k}\right\rceil \cdot \left\lceil \max_{w \in V} \mathbb{E}_{\pi}[\tau_w]\right\rceil\right] \leq \left(\frac{1}{2}\right)^{5(\log_2 n)/k} = \frac{1}{n^{5/k}}.$$

Thus, by independence of the k walks

$$\mathbb{P}_{\pi^k}\left[\tau_v^{(k)} \geq 20\left\lceil \frac{\log_2 n}{k}\right\rceil \cdot \left\lceil \max_{w \in V} \mathbb{E}_{\pi}[\tau_w]\right\rceil\right] \leq \frac{1}{n^5},$$

and finally by the union bound,

$$\mathbb{P}_{\pi^k}\left[t_{\mathsf{cov}}^{(k)} \geq 20\left\lceil \frac{\log_2 n}{k}\right\rceil \cdot \left\lceil \max_{w \in V} \mathbb{E}_{\pi}[\tau_w]\right\rceil\right] \leq \frac{1}{n^4}.$$

Therefore, since the cover time of a single walk satisfies $t_{\mathsf{cov}} = \mathcal{O}(n^3)$, we have

$$\mathbb{E}_{\pi^k}\left[t_{\mathsf{cov}}^{(k)}\right] \leq 20\left\lceil \frac{\log_2 n}{k}\right\rceil \cdot \left\lceil \max_{w \in V} \mathbb{E}_{\pi}[\tau_w]\right\rceil + o(1).$$

We now cover the case $8\log_2 n \le k \le 100\max_{v\in V}\mathbb{E}_\pi[\tau_v]\cdot\log n$, where we apply Items (i) & (ii) of Lemma 3.10 to bound $\mathbb{P}_{\pi^k}\big[\tau_{\mathsf{cov}}^{(k)}\ge t\big]$. Observe that

$$\sum_{i=0}^{t} P_{v,v}^i = \sum_{i=0}^{t}(P_{v,v}^i - \pi(v)) + (t+1)\pi(v).$$

Since the walk is lazy $(P_{v,v}^i - \pi(v))$ is non-negative and non-increasing in i [42, Exercise 12.5]. Thus

$$\sum_{i=0}^{t} P_{v,v}^i \le \sum_{i=0}^{\infty}(P_{v,v}^i - \pi(v)) + (t+1)\pi(v) = \pi(v)\mathbb{E}_\pi[\tau_v] + (t+1)\pi(v),$$

as $\pi(v)\mathbb{E}_\pi[\tau_v] = \sum_{i=0}^{\infty}(P_{v,v}^i - \pi(v))$ by [2, Lemma 2.11]. Now, by Lemma 3.10 (ii) we have

$$\mathbb{P}_\pi[N_v(t)\ge 1] \ge \frac{(t+1)\pi(v)}{\pi(v)\mathbb{E}_\pi[\tau_v] + (t+1)\pi(v)} = \frac{(t+1)}{\mathbb{E}_\pi[\tau_v] + (t+1)}.$$

Choosing $t = 8\left\lceil\frac{\log n}{k}\max_{v\in V}\mathbb{E}_\pi[\tau_v]\right\rceil$ yields

$$\mathbb{P}_\pi[N_v(t)\ge 1] \ge \frac{8\log n}{k + 8\log n} \ge \frac{4\log n}{k},$$

since $k \ge 8\log n$. Then an application of Lemma 3.10 (i) gives

$$\mathbb{P}_{\pi^k}\big[\tau_{\mathsf{cov}}^{(k)}\ge t\big] \le n\exp(-4\log n) = \frac{1}{n^3}.$$

Again, as $t_{\mathsf{cov}} = \mathcal{O}(n^3)$, we conclude $\mathbb{E}_{\pi^k}\big[t_{\mathsf{cov}}^{(k)}\big] \le 8\left\lceil\frac{\log n}{k}\max_{v\in V}\mathbb{E}_\pi[\tau_v]\right\rceil + \mathcal{O}(1)$.

Finally assume that $k \ge 100\max_{v\in V}\mathbb{E}_\pi[\tau_v]\cdot\log n$. Recall that $\max_v\mathbb{E}_\pi[\tau_v] \ge (1/2)\cdot\max_{x,y}\mathbb{E}_x[\tau_y]$ by [2, Lemma 3.15] and $\max_{x,y}\mathbb{E}_x[\tau_y] \ge \max_{x,y}mR(x,y)$ by the commute time identity [42, Proposition 10.6], where $R(x,y)$ is the effective resistance between x and y (see [42, Section 9.4]). Thus we have

$$\max_v\mathbb{E}_\pi[\tau_v] \ge \frac{1}{2}\cdot m\cdot\max_v\frac{1}{d(v)} \ge \frac{1}{2}\cdot\max_v\frac{1}{\pi(v)}, \tag{3}$$

as $R(x,y) \ge \max\{1/d(x), 1/d(y)\}$ by the definition of the effective resistance. The probability that any vertex u is unoccupied at time 0 is $(1-\pi(u))^k \le e^{-\pi(u)k}$. For any graph $t_{\mathsf{cov}} \le 16n^3$ by [34, Theorem 2] and thus, one can check that by (3), we have

$$\mathbb{E}_{\pi^k}\big[\tau_{\mathsf{cov}}^{(k)}\big] \le \max_{v\in V}16n^3\cdot ne^{-\pi(v)k} = o\Big(\frac{\max_{v\in V}\mathbb{E}_\pi[\tau_v]\cdot\log n}{k}\Big),$$

as claimed.   □

Proof of Lemma 3.4. The lower bound will follow from the general lower bound we prove later in Theorem 3.6. For the upper bound notice that, by Lemma 3.10 (iii) and our hypothesis that $\sum_{i=0}^{t} P_{vv}^i = \mathcal{O}(1 + t\pi(v))$ for any $t \le t_{\mathsf{rel}}$, we have that for any $T \ge 0$,

$$\sum_{t=0}^{T} P_{v,v}^t \le \frac{e}{e-1}\left(\sum_{i=0}^{\lceil t_{\mathsf{rel}}\rceil - 1} P_{v,v}^i - \lceil t_{\mathsf{rel}}\rceil\pi(v)\right) + (T+1)\pi(v) \le C(1 + T\pi(v)),$$

for some constant $C < \infty$. Now, by Lemma 3.10 (ii) and as $\pi_{\min} = \Omega(1/n)$ holds by hypothesis, for any v and $T = \mathcal{O}(n)$ there exists a constant $C' < \infty$ such that

$$\mathbb{P}_\pi[N_v(T)\ge 1] \ge \frac{T\pi(v)}{C(1 + T\pi(v))} \ge \frac{T}{C'n}. \tag{4}$$

First consider $k = \omega(\log n)$, and let $T = \lceil 4C'(n/k) \log n \rceil = \mathcal{O}(n)$. Lemma 3.10 (i) and (4) give

$$\mathbb{P}_{\pi^k}\left[\tau_{\text{cov}}^{(k)} \geq T\right] \leq n \exp\left(-k\mathbb{P}_\pi[N_v(T) \geq 1]\right) = n \exp\left(-4 \log n\right) = \frac{1}{n^3}.$$

Otherwise, if $k = \mathcal{O}(\log n)$, then we consider consecutive periods of length $5t_{\text{mix}} = \mathcal{O}(n)$. For any $u, v \in V$, (4) gives

$$\mathbb{P}_u[N_v(5t_{\text{mix}}) \geq 1] \geq \frac{1}{4}\mathbb{P}_\pi[N_v(t_{\text{mix}}) \geq 1] \geq \frac{t_{\text{mix}}}{4C'n}.$$

The probability that a vertex v is not hit in a period, starting from any vector in V^k is

$$(1 - \min_{u \in V} \mathbb{P}_u[N_v(5t_{\text{mix}}) \geq 1])^k \leq e^{-kt_{\text{mix}}/(4C'n)}$$

and thus after $20C'\lceil \frac{n}{kt_{\text{mix}}} \cdot \log n \rceil$ periods the probability v has not been hit is at most $e^{-5 \log n} = n^{-5}$. By the union bound, and since worst-case cover time of a single walk on any graph is $\mathcal{O}(n^3)$, the cover time is $\mathcal{O}\left(t_{\text{mix}} \cdot \frac{n}{kt_{\text{mix}}} \cdot \log n\right) = \mathcal{O}\left(\frac{n}{k} \cdot \log n\right)$. □

Proof of Lemma 3.5. From Lemma 3.10 (iii) we deduce that $\sum_{i=0}^t P_{v,v}^i = \mathcal{O}(1 + t/n + \log t)$ for any $t \leq n(\log n)^2$. Thus we can apply Lemma 3.10 (ii) which, for any $2 \leq t \leq n(\log n)^2$, gives

$$k\mathbb{P}_\pi[\tau_v \leq t] \geq \frac{ckt}{n \log t + t}, \tag{5}$$

for some fixed constant $c > 0$. Let $t^* = C\frac{n \log n}{k} \log\left(\frac{n \log n}{k}\right)$ and $C \log n \leq k \leq (n \log n)/3$ for some constant $1 < C := C(c) < \infty$ to be determined later then,

$$n \log t^* = n \log\left(C\frac{n \log n}{k} \log\left(\frac{n \log n}{k}\right)\right) \geq n \log\left(\frac{n \log n}{k}\right) \geq t^*. \tag{6}$$

In addition we have

$$\log t^* = \log\left(C\frac{n \log n}{k} \log\left(\frac{n \log n}{k}\right)\right) \leq (2 + \log C) \log\left(\frac{n \log n}{k}\right). \tag{7}$$

Thus inserting (6) into (5) then applying (7) yields the following for any $v \in V$,

$$k\mathbb{P}_\pi[\tau_v \leq t] \geq \frac{c}{2} \cdot \frac{kt^*}{n \log t^*} \geq \frac{c}{2} \cdot \frac{(Cn \log n) \log\left(\frac{n \log n}{k}\right)}{(2 + \log C)n \log\left(\frac{n \log n}{k}\right)} = \frac{Cc \log n}{4 + 2 \log C}.$$

We can assume w.l.o.g. that $c < 1$ and thus taking $C = 100/c^2$ yields $\frac{Cc}{4+2\log C} > \frac{100/c}{14+4\log(1/c)} > 5$. So by independence of the walks we have

$$\mathbb{P}_{\pi^k}\left[\tau_v^{(k)} > t^*\right] \leq \left(1 - \mathbb{P}_\pi[\tau_v \leq t^*]\right)^k \leq \exp\left(-k\mathbb{P}_\pi[\tau_v \leq t^*]\right) \leq n^{-5}.$$

Thus, by the union bound $\mathbb{P}_{\pi^k}\left[\tau_{\text{cov}}^{(k)} > t^*\right] \leq n^{-4}$. So, since the worst-case expected cover time by a single walk on any graph is $\mathcal{O}(n^3)$, we have $t_{\text{cov}}^{(k)}(\mathcal{T}_n, \pi) = \mathcal{O}(t^*)$, as claimed.

The case $k \leq (100 \log n)/c^2$ remains. For any $k = \mathcal{O}(\log n)$, consider periods of length $5t_{\text{mix}} = \mathcal{O}(n)$. Then by Lemma 3.10 (ii), for any pair of vertices v, w and some $C' < \infty$, we have

$$\mathbb{P}_w[N_v(5t_{\text{mix}}) \geq 1] \geq \frac{1}{4}\mathbb{P}_\pi[N_v(t_{\text{mix}}) \geq 1] \geq \frac{t_{\text{mix}}}{C'n \log n}.$$

Thus the probability $v \in V$ is not hit in a period, starting from any configuration in V^k is

$$(1 - \mathbb{P}_w[N_v(5t_{\mathrm{mix}}) \geq 1])^k \leq e^{-kt_{\mathrm{mix}}/(C'n\log n)}.$$

Thus after $5C'\lceil(n\log^2 n)/(kt_{\mathrm{mix}})\rceil$ periods the probability v has not been hit is at most $e^{-5\log n} = n^{-5}$, so (similarly) the cover time is $\mathcal{O}\left(t_{\mathrm{mix}} \cdot \frac{n\log^2 n}{kt_{\mathrm{mix}}}\right) = \mathcal{O}\left(\frac{n}{k} \cdot \log^2 n\right)$ as claimed. $\square$

3.4. Proofs of lower bounds

We begin by proving the lower bound obtained by the reset coupling.

Proof of Lemma 3.8. Let $x = Ck/T$, and recall that in each step that the random walk on $\widehat{G}(x)$ is not at the vertex z, it moves to z with probability x. For $i \geq 1$, we refer to the portion of the walk between the ith and $(i+1)^{\mathrm{th}}$ visits to z as the ith sub-walk. For $i \geq 0$ let $X_i \sim \mathrm{Geo}(x)$ be i.i.d. geometric random variables with mean $1/x$, then it follows that the ith sub-walk has length $X_i + 1$ and takes $X_i - 1$ steps inside G (the steps leaving and entering z are not in G) and it takes at most X_0 steps to first visit z.

If we let $X = \sum_{i=0}^{Ck/3} X_i$ then we see that $X + Ck/3$ stochastically dominates the time taken for the first $Ck/3$ sub-walks to occur within a walk on $\widehat{G}(x)$. Observe that $\mathbb{E}[X] = T/3$ and so a Chernoff bound for sums of geometric r.v.'s [32, Theorem 2.1] gives

$$\mathbb{P}[X > 2\mathbb{E}[X]] < e^{-x\mathbb{E}[X](2-1-\ln(2))} \leq e^{-\frac{Ck}{10}}.$$

Note that $2\mathbb{E}[X] + Ck/3 \leq T$. Thus, if $\mathcal{E}_1$ is the event that there are at least $Ck/3$ sub-walks in a walk of length T, then $\mathbb{P}[\mathcal{E}_1^c] \leq e^{-\frac{Ck}{10}}$.

For each $i \geq 0$, we have $\mathbb{P}[X_i \geq 1/(2x)] = (1-x)^{1/(2x)} \geq 1/2$ by Bernoulli's inequality. Let $\mathcal{E}_2$ be the event that $\{X_i \geq 1/(2x)\}$ holds for more than k values $1 \leq i \leq Ck/3$. Recall that $C > 30$ and, as the X_i's are independent, a Chernoff bound [49, Theorem 4.5] gives

$$\mathbb{P}[\mathcal{E}_2^c] \leq \mathbb{P}\left[\mathrm{Bin}\left(\frac{Ck}{3}, \frac{1}{2}\right) \leq k\right] \leq \exp\left(-\frac{1}{2} \cdot \left(1 - \frac{6}{C}\right)^2 \cdot \frac{Ck}{6}\right) \leq e^{-\frac{Ck}{20}}.$$

Observe that conditional on $\mathcal{E}_1 \cap \mathcal{E}_2$ there are at least k sub-walks which take at least s many steps within G where $s = \frac{1}{2x} - 1 = \frac{T}{2Ck} - 1 \geq \frac{T}{10Ck}$, since $T \geq 5Ck$. Recall that at each visit to z the walk on $\widehat{G}(x)$ moves to a vertex $v \in V(G)$ proportional to $\pi(v)$. It follows that, conditional on $\mathcal{E}_1 \cap \mathcal{E}_2$, we can couple k stationary random walks of length s on G to a single random walk of length T on $\widehat{G}$ such that if the latter walk has not covered $\widehat{G}$ then the former walks have not covered G. Thus

$$\mathbb{P}_{\pi^k, G}\left[\tau_{\mathrm{cov}}^{(k)} > \frac{T}{10Ck}\right] \geq \mathbb{P}_{\widehat{\pi}, \widehat{G}(x)}[\tau_{\mathrm{cov}} > T, \mathcal{E}_1 \cap \mathcal{E}_2] \geq \mathbb{P}_{\widehat{\pi}, \widehat{G}(x)}[\tau_{\mathrm{cov}} > T] - \mathbb{P}[\mathcal{E}_1^c \cup \mathcal{E}_2^c],$$

and the result follows since $\mathbb{P}[\mathcal{E}_1^c \cup \mathcal{E}_2^c] \leq e^{-\frac{Ck}{10}} + e^{-\frac{Ck}{20}} < e^{-\frac{Ck}{50}}$. $\square$

We shall prove the general lower bound with this coupling, however, we must first state a technical lemma.

Lemma 3.11. *For any reversible Markov chain and $0 < b < 1$ we have $t_{\mathrm{cov}} \leq 2t_{\mathrm{cov}}(\pi)$ and $\mathbb{P}_\pi[\tau_{\mathrm{cov}} > b \cdot t_{\mathrm{cov}}(\pi)] > \frac{1-b}{2}$.*

Proof of Lemma 3.11. We first show $t_{\mathrm{cov}} \leq 2t_{\mathrm{cov}}(\pi)$. The Random Target Lemma [42, Lemma 10.1] states that $\mathbb{E}_x[\tau_\pi] = \sum_{v \in V} \mathbb{E}_x[\tau_v]\pi(v)$ does not depend on $x \in V$. Thus, for any $u \in V$,

$$\mathbb{E}_u[\tau_\pi] = \mathbb{E}_\pi[\tau_\pi] \leq \max_{v \in V} \mathbb{E}_\pi[\tau_v].$$

Then, since choosing a vertex v independently from π and waiting for the random walk to hit v is a strong stationary time for the random walk, we have

$$\mathbb{E}_u[\tau_{\mathsf{cov}}] \leq \mathbb{E}_u[\tau_\pi] + \mathbb{E}_\pi[\tau_{\mathsf{cov}}] \leq \max_{v \in V} \mathbb{E}_\pi[\tau_v] + \mathbb{E}_\pi[\tau_{\mathsf{cov}}] \leq 2\mathbb{E}_\pi[\tau_{\mathsf{cov}}].$$

Observe that, by the Markov Property, for any $0 < b < 1$ we have

$$t_{\mathsf{cov}}(\pi) \leq bt_{\mathsf{cov}}(\pi) \cdot \mathbb{P}_\pi\big[\tau_{\mathsf{cov}} \leq bt_{\mathsf{cov}}(\pi)\big] + (bt_{\mathsf{cov}}(\pi) + t_{\mathsf{cov}}) \cdot \mathbb{P}_\pi\big[\tau_{\mathsf{cov}} > bt_{\mathsf{cov}}(\pi)\big]$$

$$= bt_{\mathsf{cov}}(\pi) + t_{\mathsf{cov}} \cdot \mathbb{P}_\pi\big[\tau_{\mathsf{cov}} > bt_{\mathsf{cov}}(\pi)\big]$$

$$\leq bt_{\mathsf{cov}}(\pi) + 2t_{\mathsf{cov}}(\pi) \cdot \mathbb{P}_\pi\big[\tau_{\mathsf{cov}} > bt_{\mathsf{cov}}(\pi)\big],$$

as $t_{\mathsf{cov}} \leq 2t_{\mathsf{cov}}(\pi)$. Rearranging gives $\mathbb{P}_\pi\big[\tau_{\mathsf{cov}} > b \cdot t_{\mathsf{cov}}(\pi)\big] \geq \frac{1-b}{2}$. $\qquad\square$

Before proving Theorem 3.6 we must recall a result by Aldous [3].

Theorem 3.12. ([3, Theorem 1]). *Let $(X_t)_{t \geq 0}$ be a stationary Markov chain on a state space I with irreducible transition matrix $P_{i,j}$ and stationary distribution π. Let τ_{cov} be the cover time and define t^* to be the solution of*

$$\sum_{i \in I} \exp\big(-t^* \cdot \pi(i)\big) = 1.$$

Let $0 < \theta < 1$ and suppose the following hypotheses are satisfied:

(a) $P_{i,i} = 0$ for all $i \in I$.

(b) $\sum_{j \in J} \exp(-t^* \cdot \pi(j)) \geq \theta$ where $J \subseteq I$ is the set of states j such that $\max_{i \in I} P_{i,j} \leq 1 - \theta$.

(c) *The chain is reversible; that is $\pi(i)P_{i,j} = \pi(j)P_{j,i}$ for all $i, j \in I$.*

Then $\mathbb{E}[\tau_{\mathsf{cov}}] \geq c_0 t^*$, where $c_0 > \theta$ depends only on θ.

We now have all the pieces we need to prove Theorem 3.6.

Proof of Theorem 3.6. We wish to apply Theorem 3.12 to the geometric reset graph $\widehat{G}(x)$ for some suitable $x := x(n, k)$. The following claim is based on [3, Proposition 2].

Claim 3.13. *For $n \geq 3$ and $0 < x \leq 1/2$ the walk on $\widehat{G}(x)$ satisfies Theorem 3.12 with $\theta = 1/2$.*

Proof of Claim. Let π and $\widehat{\pi}$ denote the stationary distributions of G and $\widehat{G}$, respectively. Note that for any $x < 1$ the walk on the weighted graph $\widehat{G}(x)$ satisfies $P_{i,i} = 0$ and is reversible. These are items (a) and (c) from Theorem 3.12 and we now prove item (b) also holds for $\theta = 1/2$.

Partition $V(\widehat{G}) = V(G) \cup \{z\}$ into sets I_1, I_2, I_3 as follows. Let I_1 be the set of leaves of G, that is vertices of degree 1 in G (and degree 2 in $\widehat{G}$ as all vertices are connected to z). Let I_2 be the set of vertices of $V(G)$ which are adjacent to at least one leaf of G. Let I_3 be the remaining vertices (in particular $z \in I_3$). Let $J = I_1 \cup I_3$ and fix $j \in J$. For the case $j = z$, by the definition of $\widehat{G}$ and by hypothesis, we have $P_{v,z} = x \leq 1/2$ for any $v \in V$. Otherwise, if i is a neighbour of $j \neq z$ then i cannot be a leaf, so $P_{i,j} = \frac{1-x}{d(i)(1-x)+xd(i)} = \frac{1-x}{d(i)} \leq 1/2$. Thus it suffices to verify

$$\sum_{j \in J} \exp\big(-t^* \cdot \widehat{\pi}(j)\big) \geq 1/2. \tag{8}$$

Consider an edge $ik \in E(\widehat{G})$ where $i \in I_1$, then we must have $k \in I_2$. Since $\widehat{\pi}(i) < \widehat{\pi}(k)$ we have $\exp(-t^* \cdot \pi(i)) \geq \exp(-t^* \cdot \pi(k))$ and also $|I_1| \geq |I_2|$. Thus, summing over all leaves i gives

$$\sum_{i \in I_1} \exp\big(-t^* \cdot \widehat{\pi}(i)\big) \geq \sum_{k \in I_2} \exp\big(-t^* \cdot \widehat{\pi}(k)\big). \tag{9}$$

However by the definition of t^*, the sum over all $i \in I$ is equal to 1, thus so (9) implies that the sum over I_2 is at most $1/2$, and so (8) follows. □

Now since $t^* \geq n \log n$ by [3, (1.2)], it follows from Claim 3.13. and Theorem 3.12 that there exists some universal constant $0 < c < \infty$ such that for any $0 < x \leq 1/2$ we have

$$\mathbb{E}_{\widehat{\pi}, \widehat{G}(x)} [\tau_{\text{cov}}] \geq cn \log n.$$

Thus by Lemma 3.11 we have $\mathbb{P}_{\widehat{\pi}, \widehat{G}(x)}[\tau_{\text{cov}} > \frac{c}{2} \cdot n \log n] \geq \frac{1}{4}$. We seek to apply Lemma 3.8 with $C = 100$ and $T = \frac{c}{2} \cdot n \log n$. Firstly, we see that the condition $T \geq 5kC$ forces the restriction $k \leq c' \cdot n \log n$ where $c' = \frac{c}{2 \cdot 5 \cdot C} = \frac{c}{1000}$ for $0 < c < \infty$ universal. Secondly the condition $1/2 \geq x = Ck/T$ forces the condition $k \leq c'' n \log n$ where $c'' = \frac{c}{2 \cdot C} \geq c'$.

Thus, provided $1 \leq k \leq c' \cdot n \log n$ holds, all other assumptions are satisfied and we have

$$\mathbb{P}_{\pi^k, G}\left[\tau_{\text{cov}}^{(k)} > \frac{(c/2) \cdot n \log n}{10 \cdot 100 \cdot k}\right] > \frac{1}{4} - \exp\left(-\frac{100k}{50}\right) \geq \frac{1}{4} - e^{-2} > \frac{1}{10}$$

by Lemma 3.8. The result follows by taking the constant in the statement to be $c/2000 \leq c'$, since $c > 0$ is universal. □

Proof of Lemma 3.9. Let $X_v = \mathbf{1}(\tau_v > t)$ be the indicator that $v \in V$ has *not* been visited up-to time t and $X = \sum_{v \in S} X_v$. Let $p_v = \mathbb{P}_\pi[\tau_v \leq t]$, $p = \max_{v \in S} p_v$, and observe that $\mathbb{E}_{\pi^k}[X_v] = (1 - p_v)^k$. Thus, we have

$$\mathbb{E}_{\pi^k}[X] = \sum_{v \in S} (1 - p_v)^k \geq \sum_{v \in S} e^{-kp_v}(1 - p_v^2 k) \geq \sum_{v \in S} e^{-kp_v}(1 - p^2 k) > 0, \qquad (10)$$

where the first inequality is by (1) since $0 \leq p \leq 1$ and the last is by the assumption $p^2 k \leq 2p^2 k < 1$. Let $R(t)$ be the number of vertices in S that are visited by a single t-step random walk. Observe that $\mathbb{E}_\pi[R(t)] = \sum_{v \in S} p_v$. Let $r(v, w) = \mathbb{P}_\pi[\tau_v \leq t, \tau_w \leq t]$, then for any $v, w \in V$,

$$\mathbb{E}_{\pi^k}[X_v X_w] = (1 - \mathbb{P}_\pi[\{\tau_v \leq t\} \cup \{\tau_w \leq t\}])^k = (1 - p_v - p_w + r(v, w))^k.$$

Recall the identity $a^k - b^k = (a - b) \sum_{i=0}^{k-1} a^{k-1-i} b^i$ for $k \geq 2$. Thus,

$$\mathbb{E}_{\pi^k}[X^2] = \sum_{v \in S} \sum_{w \in S} (1 - p_v - p_w + r(v, w))^k$$

$$= \sum_{v \in S} \sum_{w \in S} \left((1 - p_v - p_w)^k + ((1 - p_v - p_w + r(v, w))^k - (1 - p_v - p_w)^k)\right)$$

$$= \sum_{v \in S} \sum_{w \in S} \left((1 - p_v - p_w)^k + r(v, w) \sum_{i=0}^{k-1} (1 - p_v - p_w)^i (1 - p_v - p_w + r(v, w))^{k-1-i}\right).$$

Since $2p^2 k < 1$ and $k \geq 2$, we have $1 - p_v - p_w \geq 1 - 2p > 0$. Additionally we have $r(v, w) \leq \min(p_v, p_w)$ and so $1 - p_v - p_w + r(v, w) \leq 1 - \max\{p_v, p_w\} \leq 1$. Thus

$$\mathbb{E}_{\pi^k}[X^2] \leq \sum_{v \in S} \sum_{w \in S} \left((1 - p_v - p_w)^k + r(v, w) \cdot k\right)$$

$$\leq \sum_{v \in S} \sum_{w \in S} e^{-kp_v - kp_w} + \left(\sum_{v \in S} \sum_{w \in S} r(v, w)\right) \cdot k$$

$$= \left(\sum_{v \in S} e^{-kp_v}\right)^2 + \mathbb{E}_\pi[R(t)^2] \cdot k. \qquad (11)$$

Recall the Paley-Zygmund inequality: for any non-negative random variable X, $\mathbb{P}_{\pi^k}[X > 0] \geq \mathbb{E}_{\pi^k}[X]^2/\mathbb{E}_{\pi^k}[X^2]$. Inserting (10) and (11) into the (inverted) fraction $\frac{\mathbb{E}_{\pi^k}[X^2]}{\mathbb{E}_{\pi^k}[X]^2}$ gives

$$\frac{\mathbb{E}_{\pi^k}\left[X^2\right]}{\mathbb{E}_{\pi^k}[X]^2} \leq \frac{1}{(1-p^2k)^2} + \frac{\mathbb{E}_{\pi}[R(t)^2]k}{|S|^2 e^{-2pk}(1-p^2k)^2} = \frac{1}{(1-p^2k)^2} + \frac{ke^{2pk}\mathbb{E}_{\pi}[R(t)^2]}{|S|^2(1-p^2k)^2}.$$

Finally, we claim that $\mathbb{E}_{\pi}[R(t)^2] \leq 2p^2|S|(\min_{v \in S} \pi(v))^{-1}$, therefore

$$\frac{\mathbb{E}_{\pi^k}\left[X^2\right]}{\mathbb{E}_{\pi^k}[X]^2} \leq \frac{1}{(1-p^2k)^2} + \frac{2kp^2 e^{2kp}|S|/(\min_{v \in S} \pi(v))}{|S|^2(1-p^2k)^2} \leq \frac{1 + \frac{2kp^2 e^{2kp}}{|S| \min_{v \in S} \pi(v)}}{(1-p^2k)^2}$$

Now since $\mathbb{P}_{\pi^k}\left[\tau_{\mathrm{cov}}^{(k)} \leq t\right] = 1 - \mathbb{P}_{\pi^k}\left[\tau_{\mathrm{cov}}^{(k)} > t\right] \leq 1 - \mathbb{E}_{\pi^k}[X]^2/\mathbb{E}_{\pi^k}[X^2]$ we have

$$\mathbb{P}_{\pi^k}\left[\tau_{\mathrm{cov}}^{(k)} \leq t\right] \leq 1 - \frac{(1-p^2k)^2}{1 + \frac{2kp^2 e^{2kp}}{|S| \min_{v \in S} \pi(v)}} \leq 1 - (1-2p^2k)\left(1 - \frac{2kp^2 e^{2kp}}{|S| \min_{v \in S} \pi(v)}\right),$$

since $\frac{1}{1+x} \geq 1 - x$ for any $x \geq 0$. Now, as $|S| \min_{v \in S} \pi(v) \leq 1$ we have

$$\mathbb{P}_{\pi^k}\left[\tau_{\mathrm{cov}}^{(k)} \leq t\right] \leq 2kp^2 + \frac{2kp^2 e^{2kp}}{|S| \min_{v \in S} \pi(v)} \leq \frac{4kp^2 e^{2kp}}{|S| \min_{v \in S} \pi(v)},$$

which concludes the main proof. To prove the claim, first observe that

$$\mathbb{E}_{\pi}\left[R(t)^2\right] = \mathbb{E}_{\pi}\left[\sum_{v \in S}\sum_{w \in S} \mathbf{1}(\tau_v \leq t)\mathbf{1}(\tau_w \leq t)\right] = \sum_{v \in S}\sum_{w \in S} \mathbb{P}_{\pi}[\tau_v \leq t, \tau_w \leq t]. \qquad (12)$$

Now, by the Markov property, for any pair $(v, w) \in V^2$ we have

$$\mathbb{P}_{\pi}[\tau_v \leq t, \tau_w \leq t] \leq \mathbb{P}_{\pi}[\tau_v \leq t]\,\mathbb{P}_v[\tau_w \leq t] + \mathbb{P}_{\pi}[\tau_w \leq t]\,\mathbb{P}_w[\tau_v \leq t]$$
$$\leq p\mathbb{P}_v[\tau_w \leq t] + p\mathbb{P}_w[\tau_v \leq t]. \qquad (13)$$

Thus by inserting the bound from (13) into (12) we have

$$\mathbb{E}_{\pi}\left[R(t)^2\right] \leq p\sum_{v \in S}\sum_{w \in S}\mathbb{P}_v[\tau_w \leq t] + p\sum_{w \in S}\sum_{v \in S}\mathbb{P}_w[\tau_v \leq t]$$

$$\leq p\sum_{v \in S}\mathbb{E}_v[R(t)] + p\sum_{w \in S}\mathbb{E}_w[R(t)]$$

$$\leq p\sum_{v \in S}\mathbb{E}_v[R(t)] \cdot \pi(v) \cdot \left(\min_{v \in S}\pi(v)\right)^{-1} + p\sum_{w \in S}\mathbb{E}_w[R(t)] \cdot \pi(w) \cdot \left(\min_{v \in S}\pi(v)\right)^{-1}$$

$$\leq 2p\mathbb{E}_{\pi}[R(t)]\left(\min_{v \in S}\pi(v)\right)^{-1}$$

$$\leq 2p^2|S|(\min_{v \in S}\pi(v))^{-1},$$

as claimed. $\qquad\qquad\square$

4. Mixing few walks to cover many vertices

In this section, we present several bounds on $t_{\mathrm{cov}}^{(k)}$, the multiple cover time from worst-case start vertices, based on $t_{\mathrm{cov}}^{(k)}(\pi)$, the multiple cover time from stationarity, and a new notion that we

call *partial* mixing time. The intuition behind this is that on many graphs such as cycles or binary trees, only a certain number, say $\tilde{k}$ out of k walks will be able to reach vertices that are "far away" from their start vertex. That means covering the whole graph G hinges on how quickly these $\tilde{k}$ "mixed" walks cover G. However, we also need to take into account the number of steps needed to "mix" those. Theorem 4.7 makes this intuition more precise and suggests that the best strategy for covering a graph might be when $\tilde{k}$ is chosen so that the time to mix $\tilde{k}$ out of k walks and the stationary cover time of $\tilde{k}$ walks are approximately equal. As in the previous section we shall first state our results before proving them in the final two subsections.

4.1. Two notions of mixing for multiple random walks

We begin by introducing the notion of *partial mixing time*. For any graph G, and any $1 \leq \tilde{k} < k$, we define the *partial mixing time*:

$$
t_{\text{mix}}^{(\tilde{k},k)}(G) = \inf \left\{ t \geq 1 : \text{there exists an SST } \tau \text{ such that } \min_{v \in V} \mathbb{P}_v[\tau \leq t] \geq \tilde{k}/k \right\}
$$

$$
= \inf \left\{ t \geq 1 : s(t) \leq 1 - \tilde{k}/k \right\}. \tag{14}
$$

where SST stands for strong stationary time, and $s(t)$ is the separation distance (see Section 2). We note that the two definitions above are equivalent by the following result.

Proposition 4.1. ([1], Proposition 3.2). *If σ is an SST then $\mathbb{P}[\sigma > t] \geq s(t)$ for any $t \geq 0$. Furthermore there exists an SST for which equality holds.*

Finally, we can rewrite the second definition based on $s(t)$ as follows:

$$
t_{\text{mix}}^{(\tilde{k},k)}(G) = \inf \left\{ t \geq 1 : \forall u, v \in V : P_{u,v}^t \geq \tilde{k}/k \cdot \pi(v) \right\}. \tag{15}
$$

This notion of mixing, based on the idea of separation distance and strong stationary times for single walks, will be useful for establishing an upper bound on the worst-case cover time. For lower bounds on the cover time by multiple walk we will now introduce another notion of mixing for multiple random walks in terms of hitting probabilities of large sets. Before doing so, we recall a fundamental connection for single random walks which links mixing times with hitting times of large sets. In particular, let

$$
t_{\text{H}}(\alpha) = \max_{u \in V, S \subseteq V : \pi(S) \geq \alpha} \mathbb{E}_u[\tau_S], \qquad \text{and} \qquad t_{\text{H}} := t_{\text{H}}(1/4),
$$

then the following theorem shows this *large-set hitting time* is equivalent to the mixing time.

Theorem 4.2. ([52] and independently [54]). *Let $\alpha < 1/2$. Then there exist positive constants $c(\alpha)$ and $C(\alpha)$ so that for every reversible chain*

$$
c(\alpha) \cdot t_{\text{H}}(\alpha) \leq t_{\text{mix}}(\alpha) \leq C(\alpha) \cdot t_{\text{H}}(\alpha).
$$

Inspired by that fundamental result, we introduce the following quantity, which will be used to lower bound the cover time of multiple random walks,

$$
t_{\text{large-hit}}^{(\tilde{k},k)}(G) = \min \left\{ t \geq 1 : \min_{u \in V, S \subseteq V : \pi(S) \geq 1/4} \mathbb{P}_u[\tau_S \leq t] \geq \frac{\tilde{k}}{k} \right\}. \tag{16}
$$

Note that both notions of mixing times are only defined for $\tilde{k} < k$. However, by the union bound, there exists a $C < \infty$ such that if we run k walks for $Ct_{\mathsf{mix}} \log k$ steps then all k walks will be close to stationarity in terms of TV-distance. Our next lemma generalises this fact.

Lemma 4.3. *There exists a constant $C < \infty$ such that for any graph and $1 \leq \tilde{k} < k$ we have*

(i) $t_{\mathsf{mix}}^{(\tilde{k},k)} \leq 2 \cdot t_{\mathsf{mix}} \cdot \left\lceil \log_2 \left(\dfrac{2k}{k - \tilde{k}} \right) \right\rceil$,

(ii) $t_{\mathsf{large-hit}}^{(\tilde{k},k)} \leq C \cdot t_{\mathsf{mix}} \cdot \left\lceil \log \left(\dfrac{k}{k - \tilde{k}} \right) \right\rceil$.

The partial mixing time can be bounded from below quite simply by mixing time.

Lemma 4.4. *For any graph and $1 \leq \tilde{k} < k$ we have*

$$t_{\mathsf{mix}}^{(\tilde{k},k)} \geq t_{\mathsf{mix}} \left(1 - \frac{\tilde{k}}{k} \right).$$

Proof. This follows since $d_{\mathrm{TV}}(t) \leq s(t)$ holds for any $t \geq 0$ by [42, Lemma 6.3]. $\qquad\square$

We would prefer a bound in terms of $t_{\mathsf{mix}} := t_{\mathsf{mix}}(1/4)$ instead of $t_{\mathsf{mix}}(1 - \tilde{k}/k)$ as the former is easier to compute for most graphs. The following lemma establishes such a lower bound for both notions of mixing time at the cost of a $\tilde{k}/k$ factor.

Lemma 4.5. *There exists some constant $c > 0$ such that for any graph and $1 \leq \tilde{k} < k$ we have*

(i) $t_{\mathsf{mix}}^{(\tilde{k},k)} \geq c \cdot \dfrac{\tilde{k}}{k} \cdot t_{\mathsf{mix}}$,

(ii) $t_{\mathsf{large-hit}}^{(\tilde{k},k)} \geq c \cdot \dfrac{\tilde{k}}{k} \cdot t_{\mathsf{mix}}$.

We leave as an open problem whether our two notions of mixing for multiple random walks are equivalent up to constants, but the next result gives partial progress in one direction.

Lemma 4.6. *For any graph and $1 \leq \tilde{k} < k/4$ we have*

$$t_{\mathsf{large-hit}}^{(\tilde{k},k)} \leq t_{\mathsf{mix}}^{(4\tilde{k},k)} + 1 \leq 2t_{\mathsf{mix}}^{(4\tilde{k},k)}.$$

4.2. Upper and lower bounds on cover time by partial mixing

Armed with our new notions of mixing time for multiple random walks from Section 4.1, we can now use them to prove upper and lower bounds on the worst-case cover time in terms of stationary cover times and partial mixing times. We begin with the upper bound.

Theorem 4.7. *For any graph G and any $1 \leq k \leq n$,*

$$t_{\mathsf{cov}}^{(k)} \leq 12 \cdot \min_{1 \leq \tilde{k} < k} \max \left(t_{\mathsf{mix}}^{(\tilde{k},k)}, t_{\mathsf{cov}}^{(\tilde{k})}(\pi) \right).$$

This theorem improves on various results in [5] and [19] which bound the worst-case cover time by mixing all k walks, and it also generalises a previous result in [20, Lemma 3.1], where most walks were mixed, that is, $\tilde{k} = k/2$.

We also prove a lower bound for cover times, however this involves the related definition of partial mixing time based on the hitting times of large sets.

Theorem 4.8. *For any graph G with $\pi_{\max} = \max_u \pi(u)$ and any $1 \leq k \leq n$,*

$$t_{\mathrm{cov}}^{(k)} \geq \frac{1}{16} \cdot \max_{1 \leq \tilde{k} < k} \min\left(t_{\mathrm{large-hit}}^{(\tilde{k},k)}, \frac{1}{\tilde{k}\pi_{\max}} \right).$$

Further, for any regular graph G any $\gamma > 0$ fixed, there is a constant $C = C(\gamma) > 0$ such that

$$t_{\mathrm{cov}}^{(k)} \geq C \cdot \max_{n^\gamma \leq \tilde{k} < k} \min\left(t_{\mathrm{large-hit}}^{(\tilde{k},k)}, \frac{n \log n}{\tilde{k}} \right).$$

As we will see later, both Theorems 4.7 and 4.8 yield asymptotically tight (or tight up to logarithmic factors) upper and lower bounds for many concrete networks. To explain why this is often the case, note that both bounds include one non-increasing function in $\tilde{k}$ and one non-decreasing in function in $\tilde{k}$. That means both bounds are optimised when the two functions are as close as possible. Then balancing the two functions in the upper bound asks for $\tilde{k}$ such that $t_{\mathrm{mix}}^{(\tilde{k},k)} \approx t_{\mathrm{cov}}^{(\tilde{k})}(\pi)$. Similarly, balancing the two functions in the first lower bound demands $t_{\mathrm{large-hit}}^{(\tilde{k},k)} \approx n/\tilde{k}$ (assuming $\pi_{\max} = O(1/n)$). Hence for any graph G where $t_{\mathrm{mix}}^{(\tilde{k},k)} \approx t_{\mathrm{large-hit}}^{(\tilde{k},k)}$, and also $t_{\mathrm{cov}}^{(\tilde{k})}(\pi) \approx n/\tilde{k}$, the upper and lower bounds will be close. This turns out to be the case for many networks, as we will demonstrate in Section 5.

One exception where Theorem 4.8 is far from tight is the cycle, we shall also prove a min-max theorem but with a different notion of partial cover time which is tight for the cycle.

For a set $S \subseteq V$, we let $\tau_{\mathrm{cov}}^{(k)}(S)$ be the first time that every vertex in S has been visited by at least one of the k walks, thus $\tau_{\mathrm{cov}}^{(k)}(V) = \tau_{\mathrm{cov}}^{(k)}$. Then we define the set cover time

$$t_{\mathrm{large-cov}}^{(k)} = \min_{S:\pi(S) \geq 1/4} \min_{\mu} \mathbb{E}_{\mu^k}\left[\tau_{\mathrm{cov}}^{(k)}(S) \right],$$

where the first minimum is over all sets $S \subseteq V$ satisfying $\pi(S) \geq 1/4$ and the second is over all probability distributions μ on the set $\partial S = \{x \in S : \text{exists } y \in S^c, xy \in E\}$.

Theorem 4.9. *For any graph G and any $1 \leq k \leq n$,*

$$t_{\mathrm{cov}}^{(k)} \geq \frac{1}{2} \cdot \max_{1 \leq \tilde{k} < k} \min\left(t_{\mathrm{large-hit}}^{(\tilde{k},k)}, t_{\mathrm{large-cov}}^{(\tilde{k})} \right).$$

4.3. Geometric lower bounds on the large-hit and large-cover times

We will now derive two useful lower bounds on $t_{\mathrm{large-hit}}^{(\tilde{k},k)}$; one based on the conductance of the graph, and a second one based on the distance to a large set the random walk needs to hit.

For two sets $A, B \subseteq V$ the *ergodic flow* $Q(A, B)$ is given by $Q(A, B) = \sum_{a \in A, b \in B} \pi(a) P_{a,b}$, where P denotes the transition matrix of a (lazy) single random walk. We define the *conductance* $\Phi(S)$ of a set $S \subseteq V$ with $\pi(S) \in (0, 1/2]$ to be

$$\Phi(S) = \frac{Q(S, S^c)}{\pi(S)} \quad \text{and let} \quad \Phi(G) = \min_{S \subseteq V, 0 < \pi(S) \leq 1/2} \Phi(S).$$

Lemma 4.10. *For any graph G with conductance $\Phi(G)$, and any $1 \leq \tilde{k} \leq k$, we have*

$$t_{\mathrm{large-hit}}^{(\tilde{k},k)} \geq \frac{\tilde{k}}{k} \cdot \frac{2}{\Phi(G)}.$$

We remark that a similar bound to that in Lemma 4.10 was used implicitly in [58, Proof of Theorem 1.1], which proved $t_{\text{cov}}^{(k)} \geq \sqrt{\frac{n}{k \cdot \Phi(G)}}$.

The next lemmas are needed to apply Theorems 4.8 & 4.9 to cycles/tori.

Lemma 4.11. *Let G be a d-dimensional torus with constant $d \geq 2$ (or cycle, $d = 1$). Then, for any $u \in V$, $S \subseteq V$ satisfying $|S| \geq n/4$, and $\tilde{k} \leq k/2$, we have*

$$t_{\text{large-hit}}^{(\tilde{k},k)} = \Omega\left(\frac{\text{dist}(u,S)^2}{\log(k/\tilde{k})}\right).$$

Lemma 4.12. *Let $S \subseteq V$ be a subset of vertices with $\pi(S) \geq 1/4$, $t \geq 2$ be an integer and $k \geq 100$ such that $\sum_{s=0}^{t} P_{u,u}^s \geq 32 \cdot t \cdot \pi(u) \cdot k$ for all $u \in S$. Then, for any distribution μ on S,*

$$\mathbb{E}_{\mu^{k/8}}\left[\tau_{\text{cov}}^{(k/8)}(S)\right] \geq t/5.$$

4.4. Proofs of lemmas in Section 4.1

Proof of Lemma 4.3. We start with Item (i). Let $\bar{d}_{\text{TV}}(t) = \max_{x,y \in V} \|P_{x,\cdot}^t - P_{y,\cdot}^t\|_{\text{TV}}$. Then

$$d_{\text{TV}}(t) \leq \bar{d}_{\text{TV}}(t) \leq 2d_{\text{TV}}(t), \qquad \bar{d}_{\text{TV}}(\ell t) \leq \bar{d}_{\text{TV}}(t)^\ell \qquad \text{and} \qquad s(2t) \leq 1 - (1 - \bar{d}_{\text{TV}}(t))^2,$$

hold for any integers $t, \ell \geq 0$ by Lemmas 4.11, 4.12 and 19.3, respectively, of [42].

Thus if we take $\ell = \lceil \log_2(2k/(k-\tilde{k})) \rceil$ and $t = t_{\text{mix}}(1/4)$ then we have

$$s(2\ell t) \leq 1 - (1 - \bar{d}_{\text{TV}}(\ell t))^2 \leq 2\bar{d}_{\text{TV}}(\ell t) \leq 2\bar{d}_{\text{TV}}(t)^\ell \leq 2 \cdot (2d_{\text{TV}}(t))^\ell \leq 2 \cdot \left(\frac{1}{2}\right)^{\log_2\left(\frac{2k}{k-\tilde{k}}\right)} \leq 1 - \frac{\tilde{k}}{k},$$

it follows that $t_{\text{mix}}^{(\tilde{k},k)} \leq 2\ell t \leq 2t_{\text{mix}}\lceil \log_2(2k/(k-\tilde{k})) \rceil$.

For Item (ii), by the Markov property for any non-negative integers ℓ and t, we have

$$\max_{x \in V, S \subseteq V, \pi(S) \geq 1/4} \mathbb{P}_x[\tau_S > \ell t] \leq \left(\max_{x \in V, S \subseteq V, \pi(S) \geq 1/4} \mathbb{P}_x[\tau_S > t]\right)^\ell.$$

By Markov's inequality we have $\max_{x \in V, S \subseteq V, \pi(S) \geq 1/4} \mathbb{P}_x[\tau_S > 2t_{\text{H}}] \leq 1/2$ and by Theorem 4.2 there exists some C such that $t_{\text{H}} \leq Ct_{\text{mix}}$. Thus, if we take $T = 2Ct_{\text{mix}} \cdot \lceil \log_2(k/(k-\tilde{k})) \rceil$, then

$$\max_{x \in V, S \subseteq V, \pi(S) \geq 1/4} \mathbb{P}_x[\tau_S > T] \leq \left(\max_{x \in V, S \subseteq V, \pi(S) \geq 1/4} \mathbb{P}_x[\tau_S > 2t_{\text{H}}]\right)^{\log_2\left(k/(k-\tilde{k})\right)} \leq 1 - \frac{\tilde{k}}{k}.$$

It follows that $t_{\text{large-hit}}^{(\tilde{k},k)} \leq T = C't_{\text{mix}}\lceil \log(k/(k-\tilde{k})) \rceil$ for some $C' < \infty$. $\qquad\square$

Proof of Lemma 4.5. For Item (i), if we let $\ell = t_{\text{mix}}^{(\tilde{k},k)} + 1$ then the separation distances satisfies $s(\ell) \leq 1 - \tilde{k}/k$. Thus, by the definition of separation distance, for any pair of vertices $x, y \in V$ we have $P_{x,y}^\ell \geq (\tilde{k}/k) \cdot \pi(y)$. Thus for any $x \in V$ and set $S \subseteq V$ satisfying $\pi(S) \geq 1/4$ we have

$$\mathbb{P}_x[\tau_S \leq \ell] \geq \sum_{y \in S} P_{x,y}^\ell \geq \sum_{y \in S} \frac{\tilde{k}}{k} \cdot \pi(y) = \frac{\tilde{k}}{k} \sum_{y \in S} \pi(y) = \frac{\tilde{k}}{k}\pi(S) \geq \frac{\tilde{k}}{4k}. \tag{17}$$

Since (17) holds for all $x \in V$ and $S \subseteq V$ where $\pi(S) \geq 1/4$, we have $t_{\text{H}} \leq \ell \cdot (4k/\tilde{k})$, as τ_S is stochastically dominated by ℓ times the number of phases of length ℓ before S is hit. By Theorem 4.2 there

exists a universal constant $C < \infty$ such that for any graph $t_{\text{mix}} \le C \cdot t_{\text{H}}$, thus

$$t_{\text{mix}} \le C \cdot t_{\text{H}} \le C \cdot \ell \cdot (4k/\tilde{k}) = \frac{4Ck}{\tilde{k}} \left(t_{\text{mix}}^{(\tilde{k},k)} + 1 \right) \le C' \cdot \frac{k}{k} \cdot t_{\text{mix}}^{(\tilde{k},k)},$$

for some universal constant $C' < \infty$ as $t_{\text{mix}}^{(\tilde{k},k)} \ge 1$.

For Item (ii), observe that if instead we set $\ell = t_{\text{large-hit}}^{(\tilde{k},k)}$ then (17) still holds (in fact the stronger bound $\mathbb{P}_x[\tau_S \le \ell] \ge \tilde{k}/k$ holds) and the rest of the proof goes through unchanged. $\square$

Proof of Lemma 4.6. If we let $\ell = t_{\text{mix}}^{(4\tilde{k},k)} + 1$ then $s(\ell) \le 1 - 4\tilde{k}/k$. So, $P_{x,y}^\ell \ge (4\tilde{k}/k) \cdot \pi(y)$ for any pair of vertices $x, y \in V$. Thus for any $x \in V$ and set $S \subseteq V$ satisfying $\pi(S) \ge 1/4$ we have

$$\mathbb{P}_x[\tau_S \le \ell] \ge \sum_{y \in S} P_{x,y}^\ell \ge \sum_{y \in S} \frac{4\tilde{k}}{k} \cdot \pi(y) = \frac{4\tilde{k}}{k} \sum_{y \in S} \pi(y) = \frac{4\tilde{k}}{k} \pi(S) \ge \frac{\tilde{k}}{k}.$$

Consequently, we have $t_{\text{large-hit}}^{(\tilde{k},k)} \le \ell = t_{\text{mix}}^{(4\tilde{k},k)} + 1 \le 2t_{\text{mix}}^{(4\tilde{k},k)}$, as claimed since $t_{\text{mix}}^{(4\tilde{k},k)} \ge 1$. $\square$

4.5. Proofs of upper and lower bounds for covering via partial mixing

Proof of Theorem 4.7. Fix any $1 \le \tilde{k} < k$. It suffices to prove that with k walks starting from arbitrary positions running for

$$t = t_{\text{mix}}^{(\tilde{k},k)} + 2t_{\text{cov}}^{(\tilde{k})}(\pi) \le 3 \cdot \max\left(t_{\text{mix}}^{(\tilde{k},k)}, t_{\text{cov}}^{(\tilde{k})}(\pi) \right)$$

steps, we cover G with probability at least $1/4$. Consider a single walk X_1 on G starting from v. From the definition of $s_v(t)$ we have that at time $T = t_{\text{mix}}^{(\tilde{k},k)}$ there exists a probability measure ν_v on V such that,

$$P_{v,w}^T = (1 - s_v(T))\pi(w) + s_v(T)\nu_v(w).$$

Now, note that (14) yields $(1 - s_v(T)) \ge \tilde{k}/k$, therefore, we can generate $X_1(T)$ as follows: with probability $1 - s_v(T) \ge \tilde{k}/k$ we sample from π, otherwise we sample from ν_v. If we now consider k independent walks, the number of walks that are sampled at time T from π has a binomial distribution $\text{Bin}(k, \tilde{k}/k)$ with k trials and probability $\tilde{k}/k$, whose expectation is $\tilde{k}$. Since the expectation $\tilde{k}$ is an integer it is equal to the median. Thus, with probability at least $1/2$, at least $\tilde{k}$ walks are sampled from the stationary distribution. Now, consider only the $\tilde{k}$ independent walks starting from π. After $2t_{\text{cov}}^{(\tilde{k})}(\pi)$ steps, these walks will cover G with probability at least $1/2$, due to Markov's inequality.

We conclude that in t time steps, from any starting configuration of the k walks, the probability we cover the graph is at least $1/4$. Hence in expectation, after (at most) 4 periods of length t we cover the graph. $\square$

Proof of Theorem 4.8. By the definition of $t_{\text{large-hit}}^{(\tilde{k},k)}$, there exists a vertex u and $S \subseteq V$ such that $\mathbb{P}_u\left[\tau_S \le t_{\text{large-hit}}^{(\tilde{k},k)} - 1 \right] < \tilde{k}/k$. For such a vertex u, we consider k walks, all started from u, which run for $t_{\text{large-hit}}^{(\tilde{k},k)} - 1$ steps. It follows that the number of walks that hit S before time $t_{\text{large-hit}}^{(\tilde{k},k)}$ is dominated by a binomial distribution with parameters k and $p = \tilde{k}/k$, whose expected value and median is $\tilde{k}$. We conclude that with probability at least $1/2$, at most $\tilde{k}$ walks hit S. Note that $|S| \ge \pi(S)/\pi_{\max} \ge 1/(4\pi_{\max})$. Hence even if all $\tilde{k}$ walks that reached S before time $t_{\text{large-hit}}^{(\tilde{k},k)}$ were

allowed to run for exactly $1/(8\tilde{k}\pi_{\max}) - 1$ steps (if a walk exits S, we can completely ignore the steps until it returns to S), then the total number of covered vertices in S would be at most

$$\tilde{k} \cdot \left(\frac{1}{8\tilde{k}\pi_{\max}}\right) < 1/(4\pi_{\max}) \leq |S|,$$

which concludes the proof of the first bound.

For the second bound, we follow the first part of the proof before and consider again at most $\tilde{k}$ walks that reach the set S. Let us denote by κ the induced distribution over S upon hitting S from u for the first time. Now, each of the $\tilde{k}$ walks continues for another $\ell = \lfloor \epsilon (n/\tilde{k}) \log(n) \rfloor$ steps, where $0 < \epsilon = \epsilon(\gamma) < 1$ is a sufficiently small constant fixed later. We now define, for any $v \in S$, the probability $p_v = \mathbb{P}_\kappa[\tau_v \leq \epsilon(n/\tilde{k}) \log n]$. Observe that since a walk of length ℓ can cover at most ℓ vertices, we have $\sum_{v \in S} p_v \leq \ell$. Further, define the set $\widetilde{S} = \{v \in S : p_v < 8\ell/n\}$.

Since for every v with $p_v \geq 8\ell/n$ we have that $np_v/(8\ell) \geq 1$ and

$$|S \setminus \tilde{S}| = \left|\{v \in S : p_v \geq 8\ell/n\}\right| \leq \sum_{v \in S} \frac{np_v}{8\ell} \leq \frac{n}{8} \leq \frac{|S|}{2},$$

where the last inequality holds since G is regular. Thus $\tilde{S} \geq |S|/2$. Now, let Z be the number of unvisited vertices in $\widetilde{S}$ after we run $\tilde{k}$ random walks starting from κ. If $p_* = \max_{v \in \tilde{S}} p_v < 8\ell/n$ then, by definition and recalling that $\ell = \lfloor \epsilon (n/\tilde{k}) \log(n) \rfloor$ where $0 < \varepsilon < 1$, we have

$$\mathbb{E}\left[Z\right] \geq |\tilde{S}| \cdot \left(1 - p_*\right)^{\tilde{k}} \geq (n/2) \cdot e^{-\tilde{k} \cdot p_*} \cdot \left(1 - \tilde{k}p_*^2\right) \geq n/2 \cdot e^{-8\epsilon \log(n)} \cdot \left(1 - \frac{(8 \log n)^2}{\tilde{k}}\right) \geq n^{1-9\varepsilon},$$

where the second inequality holds by (1) and the last since $\tilde{k} \geq n^\gamma$ where $\gamma > 0$ fixed.

Finally, since each of these $\tilde{k}$ random walks can change Z by at most ℓ vertices, by the method of bounded differences [18, Theorem 5.3],

$$\mathbb{P}\left[Z < \mathbb{E}\left[Z\right]/2\right] \leq \exp\left(-2\frac{(Z/2)^2}{\tilde{k}\ell^2}\right) \leq \exp\left(-\frac{n^{2-18\epsilon}\tilde{k}}{8\varepsilon^2 n^2 \log^2 n}\right) < 1/2,$$

provided we have $\epsilon < \gamma/18$ as $\tilde{k} \geq n^\gamma$. This implies that $\mathbb{P}[Z \geq 1] \geq 1/2$. $\square$

Proof of Theorem 4.9. Since we are bounding the worst-case cover time from below we can assume that all walks start from a single vertex u. First, consider the k walks running for $t^{(\tilde{k},k)}_{\text{large-hit}} - 1$ steps. Then, by the definition of $t^{(\tilde{k},k)}_{\text{large-hit}}$, there exists a vertex u and $S \subseteq V$ such that $\mathbb{P}_u[\tau_S \leq t] < \tilde{k}/k$, therefore, the number of walks that, starting from u, hit S before time t is dominated by a binomial distribution with parameters k and $p = \tilde{k}/k$, whose expected value and median is $\tilde{k}$. We conclude that, if $\mathcal{E}$ is the event that at most $\tilde{k}$ walks hit S by time $t^{(\tilde{k},k)}_{\text{large-hit}} - 1$, then $\mathbb{P}[\mathcal{E}] \geq 1/2$.

Although, conditional on $\mathcal{E}$, we know at most $\tilde{k}$ walks hit S by time $t^{(\tilde{k},k)}_{\text{large-hit}} - 1$, we do not know when they arrived or which vertices of S they hit first. For a lower bound we assume these $\tilde{k}$ walks arrived at time 0 and then take the minimum over all sets S such that $\pi(S) \geq 1/4$ and starting distributions μ on ∂S, the vertex boundary of S (note all particles started from u, so they have the same distribution when they enter S for first time). It follows that, conditional on $\mathcal{E}$, the expected time for $\tilde{k}$ walks which hit S before time $t^{(\tilde{k},k)}_{\text{large-hit}}$ to cover S is at least $t^{(\tilde{k})}_{\text{large-cov}}$ and so the result follows. $\square$

4.6. Proofs of bounds on the large-hit and large-cover times

Proof of Lemma 4.10. Let $S \subseteq V$ be a set such that $\Phi(G) = \Phi(S)$ and $\pi(V \setminus S) \geq 1/2$ (such a set exists by symmetry of $\Phi(S) = \Phi(V \setminus S)$). Let π_S be the stationary distribution restricted to S, that is $\pi_S(s) = \pi(s)/\pi(S)$ for $s \in S$ and $\pi_S(x) = 0$ for $x \notin S$. As shown in [24, Proposition 8], the probability that a (single) random walk X_t remains in a set S when starting from a vertex in S sampled from π_S, within t steps is at least $(1 - \Phi(S)/2)^t \geq 1 - \Phi(S)t/2$. Let $t < \frac{2\tilde{k}}{k\Phi(S)}$, then $\mathbb{P}_{\pi_S}[\tau_{S^c} \leq t] < \tilde{k}/k$, and thus by taking u as the vertex that minimises the escape probability from S, we conclude that $t_{\text{large-hit}}^{(\tilde{k},k)} \geq \frac{\tilde{k}}{k} \cdot \frac{2}{\Phi(S)}$. $\qquad\square$

Before proving Lemma 4.11 we first establish an elementary result.

Lemma 4.13. (cf. [47], Theorem 13.4). *Let $G = (V, E)$ be a d-dimensional torus ($d \geq 2$) or cycle ($d = 1$). Then for any $D \geq 0$ and $t \geq 1$,*

$$\mathbb{P}\left[\max_{1 \leq s \leq t} \text{dist}(X_0, X_s) \geq D\right] \leq 2d \cdot \exp\left(-\frac{D^2}{2td^2}\right).$$

Proof of Lemma 4.13. Consider a random walk for t steps on a d-dimensional torus (or cycle), where $d \geq 1$. Let $Z_1, Z_2, \ldots, Z_t \in \{-1, 0, +1\}$ be the transitions along the first dimension, and let $S_i = Z_1 + \ldots + Z_i$. Note that S_i is a zero-mean martingale with respect to the Z_i. Define $\tau = \min\{i : |S_i| \geq D/d\} \wedge t$, which is a bounded stopping time, and thus $S_{\tau \wedge i}$ is another martingale with increments bounded by 1. Then, by Azuma's inequality [49, Theorem 13.4],

$$\mathbb{P}[\tau \leq t] = \mathbb{P}\left[|S_{\tau \wedge t}| \geq D/d\right] \leq 2\exp\left(-\frac{(D/d)^2}{2t}\right)$$

Now, for the random walk, in order to overcome a distance D during t steps, the above event must occur for at least one of the d dimensions, so by the union bound

$$\mathbb{P}\left[\max_{1 \leq s \leq t} \text{dist}(X_0, X_s) \geq D\right] \leq 2d \cdot \exp\left(-\frac{D^2}{2td^2}\right),$$

as claimed. $\qquad\square$

Proof of Lemma 4.11. Recall from the definition of $t_{\text{large-hit}}^{(\tilde{k},k)}$ that there must exist a vertex u and set S with $\pi(S) \geq 1/4$ such that the probability a random walk of length $t_{\text{large-hit}}^{(\tilde{k},k)}$ started from u has hit S is at most $\tilde{k}/k$. Notice that since the d-dimensional torus is regular any set S of size at least $n/4$ satisfies $\pi(S) \geq 1/4$. In order for a random walk to hit the set S within t steps, it must reach a distance $D = \text{dist}(u, S)$ from u at least once during t steps. Let t be given by

$$t = \left\lfloor \frac{D^2}{2d^2 \log\left(2dk/\tilde{k}\right)} \right\rfloor.$$

Then by Lemma 4.13,

$$\mathbb{P}_u[\tau_S \leq t] \leq \mathbb{P}\left[\max_{1 \leq s \leq t} \text{dist}(u, X_t) \geq D\right] \leq 2d \cdot \exp\left(-\frac{D^2}{2td^2}\right) \leq \frac{\tilde{k}}{k},$$

and the result follows. $\qquad\square$

Proof of Lemma 4.12. In the first part of the proof we will work with random walks whose lengths are independent samples from Geo$(1/t)$. That is, we consider walks $(X_0, X_1, \ldots, X_{L-1})$ where $L \geq 1$ is a geometric random variable with mean t, which is independent of the trajectory $(X_s)_{s \geq 0}$. Let $\tilde{P}_{w,u}^s = \mathbb{P}_w[X_s = u, s < L]$. We call the above a geometric random walk of expected length t.

We consider a collection of k independent geometric random walks of length t, in particular the lengths of these walks are also independent of each other.

Let us lower bound the expected number of unvisited vertices in S by k independent geometric random walks of expected length t. Define a subset $S' \subseteq S$ as

$$S' = \left\{ u \in S : \sum_{s=0}^{\infty} \sum_{w \in V} \mu(w) \tilde{P}_{w,u}^s \leq 8t \cdot \pi(u) \right\}.$$

Since the geometric random walk has an expected length of t, it visits at most t vertices in expectation and thus

$$\sum_{u \in S} \sum_{s=0}^{\infty} \sum_{w \in V} \mu(w) \tilde{P}_{w,u}^s \leq t.$$

It follows by definition of S' that

$$t \geq \sum_{u \in V \setminus S'} \sum_{s=0}^{\infty} \sum_{w \in V} \mu(w) \tilde{P}_{w,u}^s \geq \sum_{u \in V \setminus S'} 8t \cdot \pi(u),$$

and thus $\sum_{u \in V \setminus S'} \pi(u) \leq 1/8$. Hence $\sum_{u \in S'} \pi(u) \geq \pi(S) - 1/8 \geq 1/4 - 1/8 = 1/8$.

Let $Z_t = Z_t(u)$ denote the number of visits to u by a geometric random walk of expected length t. The probability a single walk starting from μ visits a vertex $u \in S'$ before being killed is at most

$$\mathbb{P}_\mu[Z_t \geq 1] = \frac{\mathbb{E}_\mu[Z_t]}{\mathbb{E}_\mu[Z_t \mid Z_t \geq 1]} = \frac{\sum_{s=0}^{\infty} \sum_{w \in V} \mu(w) \tilde{P}_{w,u}^s}{\sum_{s=0}^{\infty} \tilde{P}_{u,u}^s},$$

since conditional on the walk having reached a vertex u, the expected remaining returns before getting killed is equal to $\sum_{s=0}^{\infty} \tilde{P}_{u,u}^s$. Now observe that $\tilde{P}_{u,u}^s \geq \frac{1}{4} \cdot P_{u,u}^s$, which follows since $\mathbb{P}\left[\text{Geo}(1/t) > s\right] = \left(1 - \frac{1}{t}\right)^s \geq \left(1 - \frac{1}{t}\right)^t \geq 1/4$, for any $t \geq 2$ and any $s \leq t$. Thus we have

$$\mathbb{P}_\mu[Z_t \geq 1] \leq \frac{\sum_{s=0}^{\infty} \sum_{w \in V} \mu(w) \tilde{P}_{w,u}^s}{\frac{1}{4} \cdot \sum_{s=0}^{t} P_{u,u}^s} \leq \frac{8t \cdot \pi(u)}{\frac{1}{4} \cdot (32 \cdot t \cdot \pi(u) \cdot k)} = \frac{1}{k},$$

by hypothesis. Let Y the stationary mass of the unvisited vertices in S', then

$$\mathbb{E}[Y] \geq \sum_{u \in S'} \pi(u) \cdot \left(1 - \frac{1}{k}\right)^k \geq \frac{1}{4} \cdot \pi(S').$$

Hence with probability at least $1/4$, at least one vertex in S' remains unvisited by the k random walks whose length is sampled from $\text{Geo}(1/t)$. Finally, since each walk is independent the number of walks which run for more than t steps is binomially distributed with parameters k and $p = (1 - 1/t)^t \geq 1/4$. Thus by a Chernoff bound the probability that less than $1/8$ of the k random walks run for more than t steps is at most $\exp\left(-\frac{(k/8)^2}{(2k/4)}\right) = e^{-k/32}$. Hence, by coupling, $k/8$ random walks of length t do not visit all vertices in S with probability at least $1/4 - e^{-k/32} \geq 1/5$, provided $k \geq 100$. □

5. Applications to standard graphs

In this section, we apply the results of the previous sections to determine the stationary and worst-case multiple walk cover times. Firstly, we determine the stationary cover times for many fundamental networks using results from Sections 3 and 4. Secondly, using our results for the stationary cover times, we then apply them to the min-max (and max-min) characterisations from

Section 4. Along the way, we also have to derive bounds for the partial mixing time and the time to hit a large set. Due to this section being large, the proofs of all results are located in the same subsections as the statements (unlike Sections 3 and 4). For a quick reference and comparison of the results of this section the reader is encouraged to consult Table 1.

5.1. The cycle

Our first result determines the stationary cover time of the cycle up to constants. This result comes from Theorem 3.1 and Lemma 3.9 along with some additional results and arguments.

Theorem 5.1. *For the n-vertex cycle C_n, and any integer $k \geq 2$, we have*

$$t_{\mathsf{cov}}^{(k)}(\pi) = \Theta\left(\left(\frac{n}{k}\right)^2 \log^2 k\right).$$

The lower bound for $t_{\mathsf{cov}}^{(k)}(\pi)$ provided $k \geq n^{1/20}$ was already known [37, Lemma 18]. Here, we prove Theorem 5.1, which holds for any $k \geq 2$, by extending the applicable range of k for the lower bound and supplying a new upper bound. We also demonstrate how to fully recover the worst-case cover time below using our new methods from Section 4.

Theorem 5.2. ([5], Theorem 3.4). *For the n-vertex cycle C_n, and any $2 \leq k \leq n$, we have*

$$t_{\mathsf{cov}}^{(k)} = \Theta\left(\frac{n^2}{\log k}\right).$$

We begin with the proof for the stationary case.

Proof of Theorem 5.1. We split the analysis for $k \geq 2$ into two cases.
Case (i) $[2 < k \leq n^{1/20}]$: The lower bound is covered by [37, Lemma 18] so we just prove the upper bound. For a single walk $t_{\mathsf{cov}}(\pi) = \mathcal{O}(n^2)$ [2, Proposition 6.7] and so since k walks take at most as long to cover as a single walk, we can assume that $k \geq 10000$ when proving the upper bound.

To begin, divide the cycle as evenly as possible into k disjoint intervals $\mathcal{I}_1, \ldots, \mathcal{I}_k$ of consecutive vertices, each of size $\lfloor n/k \rfloor$ or $\lceil n/k \rceil$. For $2 \leq c \leq k/\log k$ let $t^*(c) = t^* = \lceil (cn/k)^2 \log^2 k \rceil$. Now, for each interval $\mathcal{I}_i$ we let $\mathcal{J}_i(c)$ be an interval of length $\ell = \lfloor \sqrt{t^*(c)} \rfloor \leq n$ centred around $\mathcal{I}_i$. Note that since $c \geq 2$ and n is large we have $\mathcal{I}_i \subset \mathcal{J}_i$ for each $1 \leq i \leq k$.

Claim 5.3. *For any $2 \leq c \leq k/\log k$, a walk of length $t^*(c)$ starting at any vertex in the interval $\mathcal{J}_i(c)$ will visit all vertices of $\mathcal{I}_i$ with probability at least $1/250$ when n is suitably large.*

Proof of Claim 5.3. Let $\mathcal{N}(0, 1)$ be the standard normal distribution, then by [26] for any $x > 0$:

$$\mathbb{P}[\mathcal{N}(0, 1) > x] \geq \frac{x}{x^2 + 1} \cdot \frac{1}{\sqrt{2\pi}} e^{-x^2/2}. \tag{18}$$

Note that all vertices of $\mathcal{I}_i$ will have been covered if the walk has travelled from one endpoint of $\mathcal{J}_i$ to the other via a path contained within $\mathcal{J}_i$. Let S_j be the distance of a random walk at time j from its start point. Then, for large n, by the Central Limit Theorem and (18) we have

$$\mathbb{P}\left[\frac{S_{t^*/2}}{\sqrt{t^*/2}} > \frac{\ell}{\sqrt{t^*/2}}\right] \geq (1 - o(1))\,\mathbb{P}\left[\mathcal{N}(0, 1) > \sqrt{2}\right] \geq (1 - o(1))\frac{\sqrt{2}}{3} \cdot \frac{e^{-1}}{\sqrt{2\pi}} > \frac{1}{15}. \tag{19}$$

Now, by symmetry and (19), regardless of its start point within the interval $\mathcal{J}_i$, with probability at least $1/15$ the walk will have hit the 'left' end of $\mathcal{J}_i$ within at most $t^*/2$ steps. Once at the 'left' end of $\mathcal{J}_i$, then again by (19) with probability at least $1/15$, it will have reached the right end via a path though $\mathcal{J}_i$ within at most $t^*/2$ additional steps. Thus, for suitably large n, with probability at least $(1/15)^2 > 1/250$ a walk of length t^* starting in $\mathcal{J}_i$ will cover $\mathcal{I}_i$. □

Let w_i be the number of walks which start in $\mathcal{J}_i$. By Chernoff's bound [49, Theorem 4.5]:

$$\mathbb{P}_{\pi^k}\left[w_i < \frac{3}{4} \cdot c \log k\right] = \mathbb{P}\left[\text{Bin}\left(k, \frac{c \log k}{k}\right) < \frac{3}{4} \cdot c \log k\right] \leq e^{-\frac{(1/4)^2}{2} \cdot c \log k} = k^{-c/32}. \quad (20)$$

By Claim 5.3, conditional on w_i, none of the walks in $\mathcal{J}_i$ cover $\mathcal{I}_i$ w.p. at most $\left(\frac{249}{250}\right)^{w_i}$ when n is large. Hence, by (20) a fixed interval $\mathcal{I}_i$ is not covered w.p. at most $(249/250)^{(3/4) \cdot c \log k} + k^{-c/32}$. As $(3/4) \ln (249/250) < -3/1000$ and $k \geq 10000$, for any $1000 \leq c \leq k/\log k$, we have

$$\mathbb{P}_{\pi^k}\left[\tau_{\text{cov}}^{(k)} > t^*(c)\right] \leq k\left(k^{-3c/1000} + k^{-c/32}\right) \leq 2k^{1-3c/1000} \leq k^{-c/1000}, \quad (21)$$

by the union bound. Now, as $t^*(k/\log k) = n^2$, observe that we have

$$\mathbb{E}_{\pi^k}\left[\tau_{\text{cov}}^{(k)}\right] \leq t^*(1000) + \sum_{i=1000}^{\lfloor \frac{k}{\log k}\rfloor} [t^*(i+1) - t^*(i)] \cdot \mathbb{P}_{\pi^k}\left[\tau_{\text{cov}}^{(k)} > t^*(i)\right] + \sum_{t=n^2}^{\infty} \mathbb{P}_{\pi^k}\left[\tau_{\text{cov}}^{(k)} > t\right]. \quad (22)$$

Using $\mathbb{P}_{\pi^k}[\tau_{\text{cov}}^{(k)} > t^*(i)] \leq k^{-i/1000}$ for $1000 \leq i \leq k/\log k$ by (21), gives

$$\sum_{i=1000}^{\frac{k}{\log k}} [t^*(i+1) - t^*(i)] \cdot \mathbb{P}_{\pi^k}\left[\tau_{\text{cov}}^{(k)} > t^*(i)\right] \leq \left(\left(\frac{n}{k}\right)^2 \log^2 k + 1\right) \sum_{i=0}^{\infty} (2i+1) \cdot k^{-i/1000}$$

$$= \mathcal{O}\left(\left(\frac{n}{k}\right)^2 \log^2 k\right). \quad (23)$$

Recall that $\mathbb{P}_{\pi^k}\left[\tau_{\text{cov}}^{(k)} > t^*(k/\log k)\right] \leq k^{-\frac{1}{1000} \cdot k/\log k} = e^{-k/1000}$ by (21). If $\mathbf{v}$ is the worst-case start position vector for k walks to cover a cycle then for any $t = i \cdot n^2$ and $i \geq 1$,

$$\mathbb{P}_{\pi^k}\left[\tau_{\text{cov}}^{(k)} > t\right] \leq \mathbb{P}_{\pi^k}\left[\tau_{\text{cov}}^{(k)} > t^*(k/\log k)\right] \cdot \mathbb{P}_{\mathbf{v}}\left[\tau_{\text{cov}}^{(k)} > t - n^2\right] \leq e^{-k/1000} \cdot \mathbb{P}_{\mathbf{v}}\left[\tau_{\text{cov}}^{(k)} > (i-1)n^2\right],$$

by the Markov property. Note that $\mathbb{P}_{\mathbf{v}}[\tau_{\text{cov}}^{(k)} > (i-1)n^2] \leq (1/2)^{i-1}$ by Markov's inequality since $t_{\text{cov}} = n(n-1)/2$. Thus the second sum on the RHS of (22) satisfies

$$\sum_{t=n^2}^{\infty} \mathbb{P}_{\pi^k}\left[\tau_{\text{cov}}^{(k)} > t\right] \leq n^2 \cdot \sum_{i=1}^{\infty} \mathbb{P}_{\pi^k}\left[\tau_{\text{cov}}^{(k)} > i \cdot n^2\right] \leq n^2 \cdot \sum_{i=1}^{\infty} e^{-k/1000} 2^{-i+1} = 2e^{-k/1000} n^2. \quad (24)$$

Case (i) then follows by inserting (23) and (24) into (22).

Case (ii) $[k > n^{1/20}]$: The upper bound follows from Theorem 3.1 since in this case $\log k = \Omega(\log n)$. By Lemma A.2 we have $\sum_{i=0}^{\lfloor t/2 \rfloor} P_{v,v}^{2i} = \Omega(\sqrt{t})$ so, by Lemma 3.10 (ii),

$$\mathbb{P}_{\pi}[\tau_v \leq t] = \mathcal{O}\left(\sqrt{t}/n\right). \quad (25)$$

We shall now apply Lemma 3.9 with $S = V$ and so if we set $t = (cn/k)^2 \log^2 k$ for a suitably small constant $c > 0$ then $p = \max_{v \in S} \mathbb{P}_{\pi}[\tau_v \leq t] \leq \frac{\log n}{100k}$, by (25). Observe that $|S| \min_{v \in S} \pi(v) = 1$ and also $2p^2 k < 1$ since $k \geq n^{1/20}$ and n is large. Thus Lemma 3.9 gives

$$\mathbb{P}_{\pi^k}\left[\tau_{\text{cov}}^{(k)} \leq t\right] \leq \frac{4kp^2 e^{2kp}}{|S| \min_{v \in S} \pi(v)} \leq 4k\left(\frac{\log n}{100k}\right)^2 \cdot e^{2k \cdot \frac{\log n}{100k}} = \frac{4(\log n)^2}{100^2 k} \cdot n^{1/50} = o(1),$$

as $k \geq n^{1/20}$, which completes Case (ii) and finishes the proof. $\qquad\square$

The final element we need for our analysis is to identify the partial mixing time of the cycle. For such, we provide bounds for the partial mixing time for all d-dimensional torus $\mathbb{T}_d$ (which are going to be used later), and we recall the cycle is the 1-dimensional torus $\mathbb{T}_1$.

Lemma 5.4. *For any integer $d \geq 1$ there exists a constant $C_d < \infty$ such that for any $1 \leq \tilde{k} \leq k/2 = \mathcal{O}(n)$ the partial mixing time of the n-vertex torus $\mathbb{T}_d$ satisfies $t_{\mathrm{mix}}^{(\tilde{k},k)} \leq C_d \cdot n^{2/d}/\log(k/\tilde{k})$.*

Proof. Let Q and P be the transition matrices of the lazy walk on the (infinite) d-dimensional integer lattice $\mathbb{Z}^d$ and the (finite) d-dimensional n-vertex torus $\mathbb{T}_d$, respectively. By [28, Theorem 5.1 (15)], for each $d \geq 1$ there exist constants $C, C', C'' > 0$ such that for any $t \geq 1$ and $u, v \in \mathbb{Z}^d$ satisfying $||u - v||_2 \leq t/C''$ we have

$$Q_{u,v}^t \geq \left(\frac{C}{t}\right)^{d/2} \exp\left(-C' \cdot \frac{||u-v||_2^2}{t}\right). \tag{26}$$

For any $t \geq 0$ and $u, v \in V(\mathbb{T}_d)$ we have $P_{u,v}^t \geq Q_{u,v}^t$ and $||u - v||_\infty \leq n^{1/d}/2$. Therefore,

$$\max_{u,v \in V(\mathbb{T}_d)} ||u - v||_2^2 \leq d \cdot (n^{1/d}/2)^2 = dn^{2/d}/4, \tag{27}$$

and so if we set

$$t = \left\lceil \frac{dC' \cdot n^{2/d}}{4\log(k/\tilde{k})} \right\rceil \leq \frac{dC' \cdot n^{2/d}}{2\log(k/\tilde{k})}. \tag{28}$$

then, for large n and any $u, v \in V(\mathbb{T}_d)$, we have $||u - v||_2 \leq t/C''$ and thus, by (26) and (27),

$$P_{u,v}^t \geq \left(\frac{C}{t}\right)^{d/2} \cdot \exp\left(-C' \cdot \frac{dn^{2/d}/4}{t}\right) \geq \frac{1}{n}\left(\frac{C2\log(k/\tilde{k})}{dC'}\right)^{d/2} \cdot \exp\left(-\log\left(\frac{k}{\tilde{k}}\right)\right) \geq c' \cdot \frac{\tilde{k}}{nk}$$

for some $c' := c'(d)$, as $\log(k/\tilde{k}) \geq \log 2$, which holds by hypothesis. It follows from the definition of separation distance that $s(t) \leq 1 - c' \cdot \frac{\tilde{k}}{nk}$ for t given by (28). Note that if $c' \geq 1$ then the statement of the lemma follows by (14), the definition of $t_{\mathrm{mix}}^{(\tilde{k},k)}$. Otherwise, assuming $c' < 1$, if we take $t' = \lceil 2/c' \rceil \cdot t$ then as separation distance is sub-multiplicative [42, Ex. 6.4] we have

$$s(t') \leq s(t)^{\lceil 2/c' \rceil} \leq \left(1 - \frac{c'\tilde{k}}{nk}\right)^{\lceil 2/c' \rceil} \leq \frac{1}{1 + \lceil 2/c' \rceil \cdot \frac{c'\tilde{k}}{nk}} = 1 - \frac{\lceil 2/c' \rceil \cdot \frac{c'\tilde{k}}{nk}}{1 + \lceil 2/c' \rceil \cdot \frac{c'\tilde{k}}{nk}} \leq 1 - \frac{\tilde{k}}{nk},$$

for suitably large n where in the second to last inequality we have used the fact that $(1 + x)^r \leq \frac{1}{1-rx}$ for any $r \geq 0$ and $x \in [-1, 1/r]$. $\qquad\square$

Next we apply our new methodology to the cycle to recover the worst-case cover time from the stationary case.

Proof of Theorem 5.2. We can assume that $k \geq C$ for a large fixed constant C (in particular one satisfying $\log C \geq 1$), as otherwise the result holds since $t_{\mathrm{cov}} = \Theta(n^2)$, and by [19, Theorem 4.2] the speed-up of the cover time on any graph is $\mathcal{O}(k^2)$.

For $k \geq C$, define $\tilde{k} = \lfloor \log(k) \rfloor \geq 1$. Recall that $t_{\mathrm{mix}}^{(\tilde{k},k)} = \mathcal{O}(n^2/\log(k/\log k)) = \mathcal{O}(n^2/\log k)$ by Lemma 5.4, and $t_{\mathrm{cov}}^{(\tilde{k})}(\pi) = \mathcal{O}\left(\frac{n^2}{(\log k)^2}(\log\log k)^2\right)$ by Theorem 5.1. Thus, Theorem 4.7 yields $t_{\mathrm{cov}}^{(k)} = \mathcal{O}\left(\frac{n^2}{\log k}\right)$.

We will use Theorem 4.9 to prove the lower bound, and for such, we need lower bounds for $t_{\mathrm{large-hit}}^{(\tilde{k},k)}$, and $t_{\mathrm{large-cov}}^{(k)}$ for an appropriate choice of $\tilde{k}$. We will indeed prove that if we choose $\tilde{k}$

as a constant, then $t_{\text{large-hit}}^{(\tilde{k},k)} = \Omega(n^2/\log k)$ and $t_{\text{large-cov}}^{(\tilde{k})} = \Omega(n^2)$, leading to the desired result as $t_{\text{cov}}^{(k)} = \Omega(\min(t_{\text{large-hit}}^{(\tilde{k},k)}, t_{\text{large-cov}}^{(\tilde{k})}))$.

For $t_{\text{large-hit}}^{(\tilde{k},k)}$ we note that for any vertex u of the cycle we can find a set of vertices of size at least $n/2$ with minimum distance at least $\lfloor n/4 \rfloor$ from u. Thus by Lemma 4.11 for any $\tilde{k} \le k/2$ we have

$$t_{\text{large-hit}}^{(\tilde{k},k)} = \Omega\left(\frac{n^2}{\log(k/\tilde{k})}\right).$$

To find a lower bound for $t_{\text{large-cov}}^{(\tilde{k})}$, by Lemma A.2 there exists some constant $c > 0$ such that for any $1 \le t \le n^2$ and $u \in V$ the return probabilities in a cycle satisfy $\sum_{s=0}^{t} P_{u,u}^s \ge c\sqrt{t}$. Thus if we take $\tilde{k} \ge \min(100, c/8)$ and let $t = \lfloor (cn/(256\tilde{k}))^2 \rfloor$ then $c\sqrt{t} \ge 32 \cdot t \cdot \pi(u) \cdot (8\tilde{k})$ is satisfied. Thus by Lemma 4.12 (with $k = 8\tilde{k}$) we have $t_{\text{large-cov}}^{(\tilde{k})} \ge t/5 = \Omega(n^2)$. □

5.2. Complete binary tree and two-dimensional torus

In this section, we derive the multiple cover times for the Complete Binary Tree and 2-Dimensional Torus. We treat them together as their proofs have several common elements. Some standard estimates such as return probabilities and other elementary results on trees can be found in Section A.2.

Theorem 5.5. *Let G be the two-dimensional torus $\mathbb{T}_2$ or the complete binary tree $\mathcal{T}_n$. Then there exists a constant $c > 0$ (independent of n) such that for any $1 \le k \le cn \log n$,*

$$t_{\text{cov}}^{(k)}(\pi) = \Theta\left(\frac{n \log n}{k} \log\left(\frac{n \log n}{k}\right)\right).$$

For worst-case cover time of the binary tree the best previously known bounds differ by multiplicative poly($\log n$) factors [23]. Using our new min-max and max-min characterisations, and some additional calculations, we can now determine $t_{\text{cov}}^{(k)}$ up to constants for any $1 \le k \le n$.

Theorem 5.6. *For the complete binary tree $\mathcal{T}_n$:*

$$t_{\text{cov}}^{(k)} = \begin{cases} \Theta\left(\dfrac{n}{k} \log^2 n\right) & \text{if } 1 \le k \le \log^2 n, \\[2ex] \Theta\left(\dfrac{n}{\sqrt{k}} \log n\right) & \text{if } \log^2 n \le k \le n. \end{cases}$$

The worst-case cover time of the 2d-torus was shown in [31]:

Theorem 5.7. ([31]). *For the two-dimensional torus $\mathbb{T}_2$:*

$$t_{\text{cov}}^{(k)} = \begin{cases} \Theta\left(\dfrac{n}{k} \log^2 n\right) & \text{if } 1 \le k \le \log^2 n, \\[2ex] \Theta\left(\dfrac{n}{\log(k/\log^2 n)}\right) & \text{if } \log^2 n \le k \le n. \end{cases}$$

Using the tools introduced in Section 4 we can recover the upper bounds in Theorem 5.7 fairly efficiently. However, for the lower bounds in Theorem 5.7 we did not find a way to apply our (or any other) general techniques to give a tight bound easily for all k. The methods presented in this

work give a lower bound tight up to a $\log n$ factor however we do not give the details as recovering loose bounds on known quantities is not the goal of this work.

5.2.1. Stationary cover time of the binary tree & 2d-torus

In this section, we prove Theorem 5.5. The upper bounds are established by applying Lemma 3.5 to both graphs. A matching lower bound is proved by showing that both graphs have a set which is particularly hard to cover.

Proof of Theorem 5.5. For the upper bound, in either graph we have, $\sum_{i=0}^{t} P_{v,v}^{i} = \mathcal{O}(1 + \log t)$ for any $v \in V$ and $t \le t_{\text{rel}}$ by Lemmas A.4, A.5 and A.2, and that $t_{\text{mix}} = \mathcal{O}(n)$ for both graphs by [2, Eq. 5.59] and [42, Eq. 5.6]. Thus the upper bound follows directly from Lemma 3.5.

We split the lower bound into three cases depending on the value of k. First, set

$$t^* = \frac{n \log n}{k} \cdot \log\left(\frac{n \log n}{k}\right), \tag{29}$$

and observe that we aim to prove $t_{\text{cov}}^{(k)}(\pi) = \Omega(t^*)$. We now show this in each case.

Case (i) $[1 \le k \le (\log n)^{5/3}]$: First recall the following bound by [19, Theorem 4.8]:

$$t_{\text{cov}} \le k t_{\text{cov}}^{(k)}(\pi) + \mathcal{O}\left(k t_{\text{mix}} \log k\right) + \mathcal{O}\left(k \sqrt{t_{\text{cov}}^{(k)}(\pi) t_{\text{mix}}}\right). \tag{30}$$

For both $\mathbb{T}_2$ and $\mathcal{T}_n$ we have $t_{\text{mix}} = \Theta(n)$, $t_{\text{hit}} = \Theta(n \log n)$ and $t_{\text{cov}} = \Theta(n \log^2 n)$ by [42, Section 11.3.2] and [2, Theorem 6.27], respectively. Recall also that $t_{\text{cov}}^{(k)}(\pi) \le \mathcal{O}\left(\frac{t_{\text{hit}} \log n}{k}\right) = \mathcal{O}\left(\frac{n \log^2 n}{k}\right)$ by Theorem 3.2, and so plugging these bounds into (30) gives

$$t_{\text{cov}} \le k t_{\text{cov}}^{(k)}(\pi) + \mathcal{O}\left(kn \log \log n\right) + \mathcal{O}\left(\sqrt{k} n \log n\right).$$

Thus for either graph, if $1 \le k \le (\log n)^{5/3}$, we have

$$t_{\text{cov}}^{(k)}(\pi) \ge t_{\text{cov}}/k - \mathcal{O}\left(n \log \log n\right) - \mathcal{O}\left(\frac{n}{\sqrt{k}} \log n\right) = \Omega\left(\frac{n \log^2 n}{k}\right) = \Omega(t^*).$$

Case (ii) $[(\log n)^{5/3} \le k \le n^{1/2}]$: Let t^* be as (29) and observe that in this case

$$\frac{n(\log n)^2}{2k} \le \frac{n \log n}{k} \cdot \log\left(n^{1/2} \log n\right) \le t^* \le \frac{n \log n}{k} \cdot \log\left(\frac{n}{(\log n)^{2/3}}\right) \le \frac{n(\log n)^2}{k}.$$

Let $\widehat{G} := \widehat{G}(x)$ be the geometric reset graph from Definition 3.7 where $x = k/(n(\log n)^2)$. We use $\mathbb{P}_{u,H}[\cdot]$ to denote the law of the (non-lazy) random walk on H started from $u \in V(H)$. For ease of reading, we prove the next claim on Page 56 after concluding the current proof.

Claim 5.8. *Let* $G = \mathcal{T}_n$ *or* $\mathbb{T}_2$, $(\log n)^{5/3} \le k \le n^{1/2}$, *and* $\widehat{G} := \widehat{G}(x)$, *where* $x = \kappa \cdot k/(n(\log n)^2)$ *for a fixed constant* $\kappa > 0$. *Then, there exists a subset* $S \subseteq V$ *such that* $\log |S| \ge (\log n)/100$ *and a constant* $\kappa' > 0$ *(independent of* κ*) such that* $\mathbb{E}_{u,\widehat{G}}[\tau_v] \ge \kappa' \cdot n \cdot \log\left(\frac{n \log n}{k}\right)$ *for all* $u, v \in S$.

In light of Claim 5.8 it follows from [33, Theorem 1.4] that, for any $x = \kappa \cdot k/(n(\log n)^2)$,

$$\mathbb{E}_{\widehat{\pi},\widehat{G}(x)}[\tau_{\text{cov}}] \ge \frac{\log |S|}{2} \cdot \min_{u,v \in S} \mathbb{E}_{u,\widehat{G}}[\tau_v] = \frac{1}{2} \cdot \frac{\log n}{100} \cdot \kappa' n \log\left(\frac{n \log n}{k}\right) \ge \frac{\kappa' n(\log n)^2}{400}, \tag{31}$$

where we note that although [33, Theorem 1.4] is stated only for the simple random walk, it holds for all reversible Markov chains, see [11, Page 4]. Thus, by Lemma 3.11 and (31), there exists some

constant $c > 0$, independent of κ, such that for all suitably large n,

$$\mathbb{P}_{\widehat{\pi}, \widehat{G}(x)}\big[\tau_{\text{cov}} > cn(\log n)^2\big] > 1/3. \tag{32}$$

We now aim to apply Lemma 3.8 where to begin with we choose the values $T = cn \log^2 n$ and $C = 100$. Observe that $T \geq 5Ck$ for large enough n since $k \leq n^{1/2}$. Finally, since c is independent of κ, we can set $\kappa = 100/c$ such that $x = \kappa \cdot k/(n(\log n)^2) = Ck/T$. Then, by Lemma 3.8 and (32),

$$\mathbb{P}_{\pi^k, G}\left[\tau_{\text{cov}}^{(k)} > \frac{cn(\log n)^2}{1000k}\right] > \mathbb{P}_{\widehat{\pi}, \widehat{G}(x)}\big[\tau_{\text{cov}} > cn(\log n)^2\big] - \exp\left(-\frac{100k}{50}\right) \geq \frac{1}{3} - e^{-2} \geq \frac{1}{10}.$$

Case (iii) $[n^{1/2} \leq k \leq cn \log n]$: Let $\delta \in (0, 1)$ be a small constant to be chosen later, and let $c = (\delta/2)^2$, and t^* be as given by (29). Let u be a leaf of the tree $\mathcal{T}_n$, or any vertex of $\mathbb{T}_2$, then, $\sum_{i=0}^{\delta t^*} P_{u,u}^i = \Omega(\log(\delta t^*))$ by Lemmas A.4 and A.2, respectively. Then, by an application of Lemma 3.10 (ii) there exists a non-negative constant C, such that

$$\begin{aligned}
\mathbb{P}_{\pi}\big[\tau_u \leq \delta \cdot t^*\big] &\leq \frac{2\delta \cdot (t^* + 1) \cdot \pi(u)}{\sum_{s=0}^{\delta t^*/2} P_{u,u}^s} \\
&\leq \frac{C \cdot \delta \log n \cdot \log((n/k) \cdot \log n)}{k \cdot \log\big((\delta/2) \cdot (n/k) \log n \cdot \log\big((n/k) \cdot \log n\big)\big)} \\
&= \frac{C \cdot \delta \log n \cdot \log((n/k) \cdot \log n))}{k \cdot \big(\log(\delta/2) + \log((n/k) \cdot \log n) + \log(\log((n/k) \cdot \log n))\big)} \\
&\leq \frac{2C\delta \log n \cdot \log((n/k) \cdot \log n)}{k \cdot \log((n/k) \cdot \log n)} \\
&= \frac{2C\delta \log n}{k}, \tag{33}
\end{aligned}$$

where in the last inequality we use that $k \leq c(n \log n) = (\delta/2)^2(n \log n)$, and thus we have that $\log((n/k) \cdot \log n) \geq 2\log(\delta/2)$.

We will now apply Lemma 3.9, where for the binary tree $\mathcal{T}_n$ we choose S as the set of all leaves, and for $\mathbb{T}_2$ choose $S = V(\mathbb{T}_2)$. Thus in either case we have $|S| \min_{v \in S} \pi(v) \geq 1/3$, $p \leq \frac{2C\delta \log n}{k}$ by (33) and $2p^2 k < 1$ since $k \geq n^{1/2}$. Thus by Lemma 3.9

$$\mathbb{P}_{\pi^k}\left[\tau_{\text{cov}}^{(k)} \leq t\right] \leq \frac{4kp^2 e^{2kp}}{|S| \min_{v \in S} \pi(v)} \leq 12k\left(\frac{2C\delta \log n}{k}\right)^2 \cdot e^{2k \cdot \frac{2C\delta \log n}{k}} = \frac{48C^2\delta^2(\log n)^2}{k} \cdot n^{4C\delta} = o(1),$$

where the last equality follows by taking $\delta = \frac{1}{12C}$ since $k \geq n^{1/2}$. $\qquad\square$

It remains to prove Claim 5.8.

Proof of Claim 5.8. For $\mathcal{T}_n$ let S be the set of leaves at pairwise distance at least $(\log n)/2$ and in $\mathbb{T}_2$ we take an (almost) evenly spaced sub-lattice where the distance between points next to each other is $\Theta(n^{1/4})$. It is easy to see that one can find such an S of polynomial size, in particular we can take $\log|S| \geq (\log n)/100$.

We first consider a walk $\widehat{W}_t$ in $\widehat{G}(x)$ from $\widehat{\pi}$ (rather than u). Let $N_v(T)$ be the number of visits to $v \in S$ in the interval $[0, T)$, where, for some $\delta > 0$,

$$T = \left\lfloor \delta \cdot n \cdot \log\left(\frac{n \log n}{k}\right) \right\rfloor. \tag{34}$$

Let the random variable Y be the first time that the walk in $\widehat{G}(x)$ started from a vertex in $V(G)$ leaves $V(G)$ to visit z. Observe that $Y \sim \mathrm{Geo}(x)$ regardless of the start vertex, and thus $\mathbb{P}[Y \geq 1/x] = (1-x)^{1/x} \geq e^{-1}(1-x^2) \geq e^{-2}$ by (1). Conditional on $\{Y \geq 1/x\} \cap \{\widehat{W}_0 \neq z\}$ the walk has the same law as a walk P on G until time $1/x$, as edges not leading to z all have the same weight in $\widehat{G}$. Thus, if Q is the transition matrix of the walk on $\widehat{G}(x)$ then for any vertex $v \neq z$, we have

$$Q_{v,v}^i \geq e^{-2} \cdot P_{v,v}^i.$$

Now, as $\min{(T/2 - 1, 1/x)} = 1/x$, by Lemma 3.10 (ii),

$$\mathbb{E}_{\widehat{\pi},\widehat{G}}\left[N_v(T) \mid N_v(T) \geq 1\right] \geq \frac{1}{2} \cdot \sum_{i=0}^{\min{(T/2-1,1/x)}} Q_{v,v}^i \geq \frac{e^{-2}}{2} \cdot \sum_{i=0}^{1/x} P_{v,v}^i \geq C \cdot \log(1/x),$$

for some $C > 0$ fixed by Lemmas A.4 and A.2. Now, for any fixed $\kappa > 0$, by Lemma 3.10 (ii),

$$\mathbb{P}_{\widehat{\pi},\widehat{G}}[N_v(T) \geq 1] = \frac{\mathbb{E}_{\widehat{\pi},\widehat{G}}[N_v(T)]}{\mathbb{E}_{\widehat{\pi},\widehat{G}}[N_v(T) \mid N_v(T) \geq 1]} \leq \frac{T/n}{C \cdot \log(1/x)} \leq \frac{\delta \log\left(\frac{n \log n}{k}\right)}{C \log\left(\frac{n(\log n)^2}{\kappa \cdot k}\right)} \leq \frac{\delta}{C},$$

since $\mathbb{E}_{\widehat{\pi},\widehat{G}}[N_v(T)] = \widehat{\pi}(v)T \leq T/n$. Thus, by taking $\delta = C/2 > 0$, we have

$$\mathbb{E}_{\widehat{\pi},\widehat{G}}[\tau_v] \geq \left(1 - \mathbb{P}_{\widehat{\pi},\widehat{G}}[N_v(T) \geq 1]\right) \cdot T \geq \frac{1}{2} \cdot T. \tag{35}$$

Recall $Y \sim \mathrm{Geo}(x)$ and observe that $1/x \leq n(\log n)^{1/3}/\kappa = o(n\sqrt{n})$ since $k \geq (\log n)^{5/3}$ and $\kappa > 0$ is fixed. Thus $\mathbb{P}_{u,\widehat{G}}[Y > n\sqrt{\log n}] = (1-x)^{n\sqrt{\log n}} = o(1)$. Hence, for any $u, v \in S$,

$$\mathbb{P}_{u,\widehat{G}}[\tau_v < Y] = \sum_{i=0}^{\infty} \mathbb{P}_{u,\widehat{G}}[\tau_v < i, Y = i] = o(1) + \sum_{i=0}^{n\sqrt{\log n}} \mathbb{P}_{u,\widehat{G}}[\tau_v < i \mid Y = i]\, \mathbb{P}_{u,\widehat{G}}[Y = i].$$

Now, if we condition on the walk in $\widehat{G}$ not taking an edge to z up to time i then, since the weights on edges not going to z are all the same, this has the same law as a trajectory in G of length i. It follows that $\mathbb{P}_{u,\widehat{G}}[\tau_v < i \mid Y = i] = \mathbb{P}_{u,G}[\tau_v < i]$, and so we have

$$\mathbb{P}_{u,\widehat{G}}[\tau_v < Y] = o(1) + \sum_{i=0}^{n\sqrt{\log n}} \mathbb{P}_{u,G}[\tau_v < i]\, \mathbb{P}_{u,\widehat{G}}[Y = i] \leq \mathbb{P}_{u,G}\left[\tau_v < n\sqrt{\log n}\right] + o(1). \tag{36}$$

For either graph $G = \mathcal{T}_n, \mathbb{T}_2$ we have $t_{\mathrm{hit}}(G) = \mathcal{O}(n \log n)$ and $\mathbb{E}_{u,G}[\tau_v] = \Omega(n \log n)$, where the latter bound is by symmetry as the effective resistance between any two vertices in S is $\Omega(\log n)$. Thus we can apply Lemma A.1 to give $\mathbb{P}_{u,G}[\tau_v > n\sqrt{\log n}] > 1/3$. Therefore $\mathbb{P}_{u,\widehat{G}}[\tau_v > Y] > 1/3 - o(1) \geq 1/4$ by (36). We now extend the distribution π on G to $\widehat{G}$ by setting $\pi(z) = 0$, and observe that $\widehat{\pi}(y) \leq \pi(y)$ for all $y \neq z$. A walk on $\widehat{G}$ at z moves to a vertex of $V(\widehat{G})$ distributed according to π in the next step, thus

$$\mathbb{E}_{\widehat{\pi},\widehat{G}}[\tau_v] = \sum_{y \in V(\widehat{G})} \widehat{\pi}(y)\mathbb{E}_{y,\widehat{G}}[\tau_v] \leq \sum_{y \neq z} \pi(y)\mathbb{E}_{y,\widehat{G}}[\tau_v] + \widehat{\pi}(z)\left(\mathbb{E}_{\pi,\widehat{G}}[\tau_v] + 1\right) \leq 3\mathbb{E}_{\pi,\widehat{G}}[\tau_v].$$

For any $i \geq 0$, $\mathbb{E}_{u,\widehat{G}}[\tau_v \mid \tau_v > Y, Y = i] = i + 2 + \mathbb{E}_{\pi,\widehat{G}}[\tau_v] \geq \mathbb{E}_{\widehat{\pi},\widehat{G}}[\tau_v]/3$. To see the first equality note that at time i the walk has not yet hit v and takes a step to z, thus two steps later the walk is at a vertex sampled from π. Hence, when n is large, for any $u, v \in S$ by (35)

$$\mathbb{E}_{u,\widehat{G}}[\tau_v] \geq \mathbb{E}_{u,\widehat{G}}\left[\tau_v \mid \tau_v > Y\right] \cdot \mathbb{P}_{u,\widehat{G}}[\tau_v > Y] \geq \frac{\mathbb{E}_{\widehat{\pi},\widehat{G}}[\tau_v]}{3} \cdot \frac{1}{4} \geq \frac{T}{24}.$$

The claim follows from (34) since $\delta = C/2 > 0$ and $C > 0$ is fixed and independent of κ. $\qquad\square$

5.2.2. Worst-case binary tree

The following result, needed for Theorem 5.6, gives a bound on the partial mixing time.

Lemma 5.9. *For the complete binary tree $\mathcal{T}_n$ and any $1 \leq \tilde{k} \leq k/100$,*

$$t_{\mathsf{mix}}^{(\tilde{k},k)} = \mathcal{O}\left(\frac{\tilde{k}}{k} \cdot n + \log n \right).$$

Proof. Let r be the root and $s_r(t)$ be the separation distance from r. We begin with two claims, which will be verified once we finish the current proof.

Claim 5.10. *There exists some $C < \infty$ such that $s_r(t) \leq n^{-10}$ for any $t \geq t_0 = C \log n$.*

Claim 5.11. *For any $u \in V$, $\frac{50k \log n}{n} \leq \tilde{k} \leq \frac{k}{100}$ and $t_1 = \frac{50\tilde{k}n}{k}$ we have $\mathbb{P}_u[\tau_r \leq t_1] \geq 2\tilde{k}/k$.*

From these two claims, we conclude that, for any pair of vertices v, u, and $\frac{50k \log n}{n} \leq \tilde{k} \leq \frac{k}{100}$,

$$P_{u,v}^{t_1+t_0} \geq \sum_{s=0}^{t_1} \mathbb{P}_u[\tau_r = s] \, \mathbb{P}_r\big[X_{t_1+t_0-s} = v\big] \geq \sum_{s=0}^{t_1} \mathbb{P}_v[\tau_r = s] \, \pi(v)(1 - n^{-10}) \geq \frac{2\tilde{k}}{k} \cdot \frac{\pi(v)}{2} = \frac{\tilde{k}}{k}\pi(v),$$

thus for $\tilde{k}$ as above the result follows from definition (15) of $t_{\mathsf{mix}}(\tilde{k}, k)$. For $\tilde{k} \leq \frac{50k \log n}{n}$ the result follows since $t_{\mathsf{mix}}(\tilde{k}, k)$ is increasing in $\tilde{k}$. $\qquad\square$

Proof of Claim 5.10. Let Y_t be the distance from X_t to the root r, where X_t is a lazy random walk starting from r. Then Y_t is a biased random walk (towards root (left) w.p. 1/6, towards leaves (right) w.p. 2/6 and stay put w.p. 1/2) on the path $0, \ldots, h-1$ with reflecting barriers, and $Y_0 = 0$. Consider Y_t' to be an independent random walk with the same transition matrix as Y_t but starting from distribution μ that denotes the stationary distribution of the biased walk on the path. To couple the walks we assume that Y' starts at a vertex $i > 0$ (or else it has already met Y), then both walks move independently unless Y' is next to Y (thus to the right of it). In this case we sample Y' first then if Y' moves left then Y stays put (and they meet) otherwise, Y moves either left, right or stays with probabilities 1/5, 2/5 and 2/5, respectively.

Now, notice that Y and Y' must have met by the time that Y reaches $h-1$. We can upper bound $\mathbb{P}[Y_t \leq h-1]$ by $\mathbb{P}[\sum_{i=1}^{t} Z_i \leq h-1]$, where Z_i are i.i.d. random variable that take value 1 w.p 2/6, value -1 w.p 1/6, and value 0 w.p 1/2. Since $h = \log_2 n$ and by choosing $t_0 = C \log_2 n$ with C large enough, by a simple application of Chernoff's bound, we have that $\mathbb{P}[\sum_{i=1}^{t} Z_i \leq h-1] \leq n^{-12}$. We conclude that the probability that Y_t and Y_t' do not meet in $\mathcal{O}(\log n)$ steps is n^{-12}. By the standard coupling characterisation of the total variation distance ([42, Proposition 4.7]), we have $||\mathbb{P}_0[Y_t = \cdot] - \mu(\cdot)||_{\mathsf{TV}} \leq n^{-12}$, and thus for all i, $\mathbb{P}_0[Y_t = i] - \mu(i) \geq -n^{-12}$. By symmetry, for any vertex at height i in the binary tree we have $\mathbb{P}_r[X_t = v] - \mu(i)/2^i \geq -n^{-12}$, since $\mu(i)/2^i = \pi(v)$. We conclude that $\mathbb{P}_r[X_t = v] \geq \pi(v) \cdot (1 - n^{-10})$ for any $v \in V(\mathcal{T}_n)$ as claimed. $\qquad\square$

Proof of Claim 5.11. We first bound the probability started from π. By Lemma A.3 we have $\sum_{s=0}^{t_1} P_{r,r}^s \leq 2 + 8t_1/n$, where r is the root. Therefore, by Lemma 3.10 (ii), we have

$$\mathbb{P}_\pi[\tau_r \leq t_1] \geq \frac{t_1}{2n + 8t_1}.$$

Note that the worst-case for our claim is when $v \in \mathcal{L}$ is a leaf, so we assume this. Denote by $\tau_\mathcal{L}$ the first time the random walk hits a leaf, then $\mathbb{P}_\pi[\tau_r < \tau_\mathcal{L}] \leq \frac{8 \log n}{n}$ by Lemma A.3, thus

$$\mathbb{P}_\pi[\tau_r \le t_1] \le \mathbb{P}_\pi[\tau_r \le t_1, \tau_r > \tau_\mathcal{L}] + \mathbb{P}_\pi[\tau_r < \tau_\mathcal{L}]$$

$$\le \mathbb{P}_\pi\big[\tau_r \le t_1, \tau_r > \tau_\mathcal{L}, \tau_\mathcal{L} \le c \log n\big] + \mathbb{P}_\pi\big[\tau_\mathcal{L} > c \log n\big] + \frac{8 \log n}{n}$$

$$\le \sum_{s=0}^{\lfloor c \log n \rfloor} \mathbb{P}_v[\tau_r \le t_1 - s]\, \mathbb{P}_\pi[\tau_\mathcal{L} = s] + \frac{9 \log n}{n} \tag{37}$$

the last inequality holds as $\mathbb{P}_\pi[\tau_\mathcal{L} > c \log n] \le (\log n)/n$, for c large by using the Chernoff's bound and the fact the height of a walk on the tree is a biased walk on a path with reflective barriers (as was used in the proof of Claim 5.10). Thus, by (37) we have

$$\mathbb{P}_\pi[\tau_r \le t_1] \le \mathbb{P}_u[\tau_r \le t_1]\, \mathbb{P}_\pi\big[\tau_\mathcal{L} \le c \log n\big] + \frac{9 \log n}{n} \le \mathbb{P}_u[\tau_r \le t_1] + \frac{9 \log n}{n}.$$

Since $(50 \log n)/n \le \tilde{k}/k \le 1/100$ we conclude that

$$\mathbb{P}_u[\tau_r \le t_1] \ge \frac{t_1}{2n + 8t_1} - \frac{9 \log n}{n} \ge \frac{1}{3} \cdot \frac{t_1}{2n + 8t_1} = \frac{50\tilde{k}/k}{6\left(1 + 4 \cdot 50\tilde{k}/k\right)} \ge \frac{2\tilde{k}}{k},$$

as desired. □

We now prove a lemma which may be regarded as a large-hitting time – since the random walk starts at a leaf in the left-subtree, and the goal is to hit any vertex in the right-subtree.

Lemma 5.12. *Let $\mathcal{T}_n$ be a complete binary tree rooted at r, then for any leaf ℓ and any $t \ge 1$*

$$\mathbb{P}_\ell[\tau_r \le t] \le 6t/n.$$

Proof. We first would like to prove that $P_{\ell,r}^t \le 6/n$ for any $t \ge 1$. However, this follows immediately by reversibility since $P_{r,\ell}^t \le 3/n$, as a random walk from r will have a uniform probability over all $2^h \ge n/3$ leaves by symmetry. Hence, for any $t \le n$, we have $\mathbb{P}_\ell[\tau_r \le t] \le \sum_{s=0}^t P_{\ell,r}^s \le 6t/n$ by Markov's inequality. □

Finally we are ready to prove the worst-case cover time for the binary tree.

Proof of Theorem 5.6. The results holds immediately for $k = 1$ by known results of cover times of binary trees. We proceed by a case analysis.

Case (i) $[2 \le k \le (\log n)^2]$: Choose $\tilde{k} = \lfloor k/2 \rfloor$, then by Lemma 4.3 the time for $\tilde{k}$ walks to mix is bounded by a constant times the single walk mixing time, which is $\Theta(n)$ for the case of the binary tree. Also, we have $t_{\mathsf{cov}}^{(k)} = \mathcal{O}(t_{\mathsf{cov}}^{(\tilde{k})}(\pi)) = \mathcal{O}((n/\tilde{k})(\log n)^2)$ by Theorem 5.5. Therefore, by Theorem 4.7 the upper bound follows. For the lower bound we have $t_{\mathsf{cov}}^{(k)} \ge t_{\mathsf{cov}}^{(k)}(\pi)$, thus the first part of the formula has been shown.

Case (ii) $[(\log n)^2 \le k \le n]$: Again, as shown in Theorem 5.5, for any $1 \le \tilde{k} \le k$

$$t_{\mathsf{cov}}^{(\tilde{k})}(\pi) = \mathcal{O}\left(\frac{n \log n}{\tilde{k}} \log\left(\frac{n \log n}{\tilde{k}}\right)\right).$$

Also by Lemma 5.9, for any $1 \le \tilde{k} \le k/100$, we have

$$t_{\mathsf{mix}}^{(\tilde{k},k)} = \mathcal{O}\left(\frac{\tilde{k}}{k} \cdot n + \log n\right).$$

To balance the last two upper bounds, we choose $\tilde{k} = \lfloor (\sqrt{k} \cdot \log n)/100 \rfloor$ so that

$$\max\left\{ t_{\text{cov}}^{(\tilde{k})}(\pi), t_{\text{mix}}^{(\tilde{k},k)} \right\} = \mathcal{O}\left(\frac{n \log n}{\sqrt{k}} \right),$$

and the upper bound follows from Theorem 4.7.

To prove a matching lower bound, let $\tilde{k} = \lfloor \sqrt{k} \cdot \log n \rfloor$. Assume that all k random walks start from an arbitrary but fixed leaf from the left subtree of r. Let $t = (n/6) \cdot \tilde{k}/k$. The number of walks that reach the root by time t has binomial distribution $\text{Bin}(k, p)$ with parameters k and p where $p = \mathbb{P}_l[\tau_r \leq t] \leq 6t/n = \tilde{k}/k$ by Lemma 5.12. Thus the expected number of walks to reach to root within t steps is upper bounded by the integer $\tilde{k}$ and so is the median. Therefore with probability at least $1/2$, at most $\tilde{k}$ out of the k walks reach the root vertex r by time t. Once we have $\tilde{k}$ walks at the root r, we consider the problem of covering the right sub-tree of r with root r_1 (which has $2^{d-1} - 1 = \Theta(n)$ vertices), assuming that $\tilde{k}$ walks start at the root r_1 (at step 0). Since we are looking for a lower bound we can assume that no walks leave the sub-tree and so the problem reduces to compute a lower bound of the cover time of a binary tree with $\tilde{k}$ walks from the root of the tree.

Since the $\tilde{k}$ walks start from the root, the time it takes to cover the whole set of vertices of the binary tree is lower-bounded by the time it takes to cover the set of leaves (starting from the root). We claim that the previous quantity is, again, lower-bounded by starting the walks from the stationary distribution. To see this, for each walk, independently sample a height H with probability proportional to the sum of the degrees in such a height, then stop the walk when it reaches height H for first time. A simple analysis shows that the distribution of the vertex where the walk stops is the stationary distribution. Note that before stopping the walk cannot have reached a leaf. Hence, we can ignore the time it takes to stop the walks, start all the random walks from the stationary distribution, and, for a lower bound, only consider the expected time to cover the leaves. We can apply the same argument as in the proof of Theorem 5.5 (an application of Lemma 3.9 with S taken to be a well spaced subset of the leaves) to lower bound the time taken to cover the leaves. Therefore, we conclude that there exists a constant $c > 0$ such that the expected time it takes to cover the leaves with $\tilde{k}$ walks starting from π is bounded from below by

$$c \cdot \frac{n \log n}{\tilde{k}} \log\left(\frac{n \log n}{\tilde{k}} \right).$$

Recall that with probability at least $1/2$, at most $\tilde{k}$ walks reach the root by time $(n/6) \cdot \tilde{k}/k$, thus,

$$t_{\text{cov}}^{(k)} \geq \frac{1}{2} \cdot \min\left(\frac{n}{6} \cdot \frac{\tilde{k}}{k}, \ c \cdot \frac{n \log n}{\tilde{k}} \log\left(\frac{n \log n}{\tilde{k}} \right) \right).$$

Since we set $\tilde{k} = \lfloor \sqrt{k} \cdot \log n \rfloor$ earlier, we obtain $t_{\text{cov}}^{(k)} = \Omega\left(\frac{n \log n}{\sqrt{k}} \right)$.   □

5.2.3. Worst-case 2d-torus
Recall that Lemma 5.4 in Section 5.1 bounds the partial mixing time of the d-dim torus. We now use this and our stationary cover time bounds to prove the upper bound in Theorem 5.7.

Proof of the Upper Bound in Theorem 5.7. Observe that the case $k = 1$ is immediate by known results on the cover time of the torus. Now, by Lemma 5.4, for any $1 \leq \tilde{k} \leq k/2$ we have

$$t_{\text{mix}}^{(\tilde{k},k)} = \mathcal{O}\left(\frac{n}{\log(k/\tilde{k})} \right). \tag{38}$$

Further, by Theorem 5.5, we have

$$t_{\text{cov}}^{(\tilde{k})}(\pi) = \mathcal{O}\left(\frac{n \log n}{\tilde{k}} \cdot \log\left(\frac{n \log n}{\tilde{k}}\right)\right).$$

Case (i) $[2 \leq k \leq 2(\log n)^2]$: Choose $\tilde{k} = \lfloor k/2 \rfloor$, and the bound on $t_{\text{cov}}^{(k)}(\pi)$ dominates, and we obtain $t_{\text{cov}}^{(k)} = \mathcal{O}((n/k)\log^2 n)$, by Theorem 4.7. The lower bound follows by $t_{\text{cov}}^{(k)} \geq t_{\text{cov}}^{(k)}(\pi)$.

Case (ii) $[2(\log n)^2 \leq n$, upper bound only]: We now prove (only) the upper bound in the remaining case $2(\log n)^2 \leq k \leq n$. We choose $\tilde{k} = \lfloor (\log n)^2 \cdot \log(\log k/(\log n)^2) \rfloor$ and obtain

$$t_{\text{cov}}^{(\tilde{k})}(\pi) = \mathcal{O}\left(\frac{n}{\log n \cdot \log(k/(\log n)^2)} \cdot \log n\right) = \mathcal{O}\left(\frac{n}{\log(k/(\log n)^2)}\right),$$

which is of the same order as the upper bound of $t_{\text{mix}}^{(\tilde{k},k)}$ in (38), since $\log((a/b)\log(a/b)) = \Theta(\log(a/b))$, thus the result follows from Theorem 4.7. $\square$

5.3. Expanders and preferential attachment

Formally, an expander is a (sequence of) graphs $(G_n)_{n \geq 1}$ such that for all $n \geq 1$: (i) G_n is connected, (ii) G_n has n vertices, and (iii) $t_{\text{rel}}(G_n) = 1/(1 - \lambda_2) \leq C$ for some constant $C > 0$ independent of n, where λ_2 is the second largest eigenvalue of the transition matrix. Equivalently, due to Cheeger's inequality, a graph is an expander if $\inf_n \Phi(G_n) > 0$.

All previous works [5, 19, 20] on multiple random walks required expanders to be regular (or regular up to constants). Here, we allow a broader class of expanders — our methods can treat any graph with bounded relaxation time provided it satisfies $\pi_{\text{min}} = \Omega(1/n)$. This class includes some graphs with heavy-tailed degree distributions as long as they have a constant average degree. Such non-regular expanders are quite common, as they include graph models for the internet such as preferential attachment graphs [48].

Theorem 5.13. *For any expander with* $\pi_{\text{min}} = \Omega(1/n)$, *for any* $1 \leq k \leq n$,

$$t_{\text{cov}}^{(k)} = \Theta\left(t_{\text{cov}}^k(\pi)\right) = \Theta\left(\frac{n}{k} \log n\right).$$

Proof. Note that the case $k = 1$ follows immediately from known results about cover times in expanders. For $k \geq 2$, consider $\tilde{k} = \lfloor k/2 \rfloor$, and recall that $t_{\text{mix}}^{(\tilde{k},k)} = \mathcal{O}(t_{\text{mix}}) = \mathcal{O}(\log n)$ by Lemma 4.3. By hypothesis $m/d_{\text{min}} = \mathcal{O}(n)$ and since the graph is an expander $t_{\text{rel}} = \mathcal{O}(1)$, hence by Corollary 3.3 we have $t_{\text{cov}}^{(\tilde{k})}(\pi) = \mathcal{O}((n/\tilde{k})\log n)$. Thus by Theorem 4.7 we have $t_{\text{cov}}^{(k)} = \mathcal{O}((n/k)\log n)$, proving the upper bound.

The lower bound follows by Theorem 3.6 since $t_{\text{cov}}^{(k)} \geq t_{\text{cov}}^{(k)}(\pi) = \Omega((n/k)\log n)$. $\square$

5.4. The hypercube

The hypercube is not covered by the results in the previous section, since it is not an expander. However, we will show that the same bound on stationary cover times holds nevertheless:

Theorem 5.14. *Let G be the hypercube with n vertices, then for any $k \geq 1$,*

$$t_{\text{cov}}^{(k)}(\pi) = \Theta\left(\frac{n}{k} \log n\right).$$

Proof. We wish to apply Lemma 3.4. This is applicable since the hypercube is regular and by [13], we have that for any vertex v, $\sum_{t=0}^{t_{\text{rel}}} P_{vv}^t \leq 2 + o(1)$ and $t_{\text{rel}} = \mathcal{O}(\log n)$. $\square$

We will also derive the result below in a more systematic way than the original proof [20] using our new characterisations involving partial mixing time and hitting times of large sets.

Theorem 5.15. ([20], Theorem 5.4). *For the hypercube with n vertices,*

$$
t_{\text{cov}}^{(k)} = \begin{cases} \Theta\left(\dfrac{n}{k} \log n\right) & \text{if } 1 \leq k \leq n/\log \log n, \\ \Theta(\log n \log \log n) & \text{if } n/\log \log n \leq k \leq n. \end{cases}
$$

In order to prove this theorem, we need to bound $t_{\text{large-hit}}^{(\alpha,n)}$ from below which is done with the following lemma.

Lemma 5.16. *For the hypercube with n vertices and any $1 \leq \tilde{k} \leq k$ satisfying $\tilde{k} \geq k \cdot e^{-\sqrt{\log n}} \geq 1$ we have*

$$
t_{\text{large-hit}}^{(\tilde{k},k)} \geq \frac{1}{100} \cdot (\log n) \log \log n.
$$

Proof of Lemma 5.16. Let d denotes the dimension of the hypercube with $n = 2^d$ vertices. Fix the vertex $u = 0^d$ and consider the set $S_u = \{v \in V : d_H(u,v) \leq d/2\}$, so $|S_u| \leq 3n/4$. Here d_H is the Hamming distance. We will estimate the probability of a random walk leaving S_u in $\ell = (1/100)d \log d$ steps. Recall that a lazy random walk on the hypercube can be considered as performing the following two-step process in each round: (1) Choose one of the d bits uniformly at random, (2) Independently, set the bit to $\{0, 1\}$ uniformly at random.

Let us denote by C_t the set of chosen coordinates, and U_t the set of unchosen coordinates in any of the first $t \leq \ell$ steps in the process above. Note that the unchosen coordinates are zero, while the chosen coordinates are in $\{0, 1\}$ independently and uniformly. By linearity of expectations and since $|U_t|$ is non-increasing in t, we have

$$
\mathbb{E}[|U_t|] \geq \mathbb{E}[|U_\ell|] = d \cdot \left(1 - \frac{1}{d}\right)^\ell \geq d \cdot \left(e^{-1}\left(1 - \frac{1}{d}\right)\right)^{(\log d)/100} \geq d \cdot \left(e^{-1}/2\right)^{(\log d)/100} \geq d^{0.9},
$$

where the first inequality is by (1). Using the Method of Bounded Differences [18, Theorem 5.3], we conclude that for any $t \leq \ell$,

$$
\mathbb{P}\left[|U_t| \leq d^{0.8}\right] \leq \exp\left(-\frac{2(d^{0.9} - d^{0.8})^2}{t \cdot 1^2}\right) \leq \exp(-d^{0.7}). \tag{39}
$$

Next consider the sum of the values at the chosen coordinates C_t at time $t \leq \ell$, which is given by

$$
Z_t = \sum_{i \in C_t} Y_i,
$$

where the $Y_i \in \{0, 1\}$ are independent and uniform variables, representing the coordinates of the random walk. Note that $\mathbb{E}[Z_t \mid |C_t|] = |C_t|/2$, and so Hoeffding's bound implies

$$
\mathbb{P}\left[|Z_t - |C_t|/2| \geq d^{0.75} \,\Big|\, |C_t|\right] \leq 2\exp(-2d^{1.5}/d) = 2\exp\left(-2d^{0.5}\right). \tag{40}
$$

Conditional on the events $|U_t| \geq d^{0.8}$ and $Z_t \leq \mathbb{E}[Z_t] + 2d^{0.75}$ we have $Z_t \leq (d - d^{0.8})/2 + 2d^{0.75} < d/2$, and so the random walk is still in the set S_u at step $t \leq \ell$. By the Union bound over the all steps $t = 1, \ldots, \ell$, and (39) and (40), for large d, these events hold with probability at least

$$
1 - \ell \cdot \left(\exp(-d^{0.7}) + 2\exp(-2d^{0.5})\right) \geq 1 - \exp\left(-\sqrt{d}\right) \geq 1 - \exp\left(-\sqrt{\log n}\right).
$$

Thus a random walk of length ℓ escapes the set S_u with probability at most $e^{-\sqrt{\log n}}$. Since S_u must be escaped to hit a worst-case set with stationary mass at least $1/4$, it follows that from the

definition (16) of $t_{\text{large}-\text{hit}}^{(\tilde{k},k)}$ and monotonicity that for any $\tilde{k} \geq k \cdot e^{-\sqrt{\log n}} \geq 1$ we have $t_{\text{large}-\text{hit}}^{(\tilde{k},k)} \geq \ell$. □

We can now apply our characterisation to find the worst-case cover time of the hypercube.

Proof of Theorem 5.15. We can assume that $k \geq 2$ by known results for the cover time of the hypercube. Now, observe that Lemma 4.3 and [42, (6.15)] yield

$$t_{\text{mix}}(\lfloor k/2 \rfloor, k) = \mathcal{O}(t_{\text{mix}}) = \mathcal{O}(\log n \cdot \log \log n). \tag{41}$$

Case (i) [$2 \leq k \leq n/\log \log n$]: By Theorem 5.14, for any $1 \leq \tilde{k} \leq k$ we have

$$t_{\text{cov}}^{(\tilde{k})}(\pi) = \Theta\left(\frac{n}{\tilde{k}} \cdot \log n\right).$$

Let $\tilde{k} = k/2$ and then by (41), $t_{\text{mix}}(\tilde{k}, k)$ is always at most $\mathcal{O}(t_{\text{cov}}^{(\tilde{k})}(\pi))$. Hence Theorem 4.7 implies $t_{\text{cov}}^{(k)} = \mathcal{O}((n/k) \log n)$. The lower bound follows by Theorem 3.6 since

$$t_{\text{cov}}^{(k)} \geq t_{\text{cov}}^{(k)}(\pi) = \Omega\left(\frac{n}{k} \cdot \log n\right).$$

Case (ii) [$n/\log \log n \leq k \leq n$]: If we choose $\tilde{k} = \lfloor \frac{n}{2 \log \log n} \rfloor$, then by monotonicity

$$t_{\text{cov}}^{(k)} \leq t_{\text{cov}}^{(\tilde{k})} = \mathcal{O}(\log n \cdot \log \log n).$$

Also by monotonicity and (41) we have $t_{\text{mix}}(\tilde{k}, k) \leq t_{\text{mix}}(\lfloor k/2 \rfloor, k) = \mathcal{O}(\log n \cdot \log \log n)$ thus the results follows from Theorem 4.7.

To prove a matching lower bound, recall that Lemma 5.16 states

$$t_{\text{large}-\text{hit}}^{(n \exp(-\sqrt{\log n}), n)} \geq \frac{1}{100} \cdot (\log n) \cdot \log \log n.$$

Again, by monotonicity, we can assume $k = n$ and choose $\tilde{k} = \lfloor n \exp(-\sqrt{\log n}) \rfloor \leq k$, giving

$$t_{\text{cov}}^{(k)} \geq \min\left(t_{\text{large}-\text{hit}}^{(\tilde{k}, n)}, \frac{1}{\tilde{k}\pi_{\min}}\right) \geq \min\left(\frac{\log n}{100} \cdot \log \log n, \exp\left(\sqrt{\log n}\right)\right)$$

by an application of the first bound in Theorem 4.8. □

5.5. Higher dimensional tori

The proof of the stationary cover time of higher dimensional tori is similar to the hypercube.

Theorem 5.17. *For d-dimensional torus $\mathbb{T}_d$, where $d \geq 3$, and any $1 \leq k \leq n$ we have*

$$t_{\text{cov}}^{(k)}(\pi) = \Theta\left(\frac{n}{k} \log n\right).$$

Proof. For the d-dimensional torus, where $d \geq 3$ we have $\sum_{t=0}^{t_{\text{rel}}} P_{vv}^t = \mathcal{O}(1)$ by Lemma A.2. Also $t_{\text{rel}} = \mathcal{O}(n^{2/d}) = o(n)$, see [2, Section 5.2]. Thus, we can apply Lemma 3.4. □

Using our machinery, we can recover the following result in full quite easily.

Theorem 5.18. ([31]). *For the d-dimensional torus, where $d \geq 3$ is constant:*

$$t_{\text{cov}}^{(k)} = \begin{cases} \Theta\left(\dfrac{n}{k} \cdot \log n\right) & \text{if } 1 \leq k \leq 2n^{1-2/d} \log n, \\[2ex] \Theta\left(n^{2/d} \cdot \dfrac{1}{\log(k/(n^{1-2/d} \log n))}\right) & \text{if } 2n^{1-2/d} \log n < k \leq n. \end{cases}$$

Proof of Theorem 5.18. As before, we can assume $k \geq 2$ by known results for the (single walk) cover time of the d-dim torus. By Theorem 5.17 and Lemma 5.4, respectively, we have

$$t_{\text{cov}}^{(\tilde{k})}(\pi) = \Theta\left((n/\tilde{k}) \log n\right), \quad \text{and} \quad t_{\text{mix}}^{(\tilde{k},k)} = \mathcal{O}\left(n^{2/d} / \log(k/\tilde{k})\right) \quad \text{if } 1 \leq \tilde{k} \leq k/2. \tag{42}$$

Case (i) $[2 \leq k \leq 2n^{1-2/d} \log n]$: For the upper bound, we can choose $\tilde{k} = \lfloor k/2 \rfloor$, then by Theorem 4.7, the expected stationary cover time by k walks is $O(n/k \cdot \log n)$. To obtain a matching lower bound we can simply use $t_{\text{cov}}^{(k)} \geq t_{\text{cov}}^{(k)}(\pi)$.

Case (ii) $[2n^{1-2/d} \log n < k \leq n]$: Beginning with the upper bound, set

$$\tilde{k} = \left\lfloor n^{1-2/d} \cdot \log\left(\frac{k}{n^{1-2/d} \log n}\right) \cdot \log n \right\rfloor \leq \frac{k}{2}.$$

Then inserting this value for $\tilde{k}$ into the bounds from (42) gives

$$t_{\text{cov}}^{(\tilde{k})}(\pi) = \mathcal{O}\left(\frac{n^{2/d}}{\log \frac{k}{n^{1-2/d} \log n}}\right), \quad \text{and} \quad t_{\text{mix}}^{(\tilde{k},k)} = \mathcal{O}\left(\frac{n^{2/d}}{\log(\frac{k}{n^{1-2/d} \log n} \cdot \frac{1}{\log(k/(n^{1-2/d} \log n))})}\right).$$

These bounds are both of the same order and so the upper bound follows from Theorem 4.7.

For the lower bound set $\tilde{k} = n^{1-2/d} \log n \leq k/2$. Then since $n^{1/3} < \tilde{k} \leq k$, by Theorem 4.8,

$$t_{\text{cov}}^{(k)} \geq C \cdot \min\left(t_{\text{large-hit}}^{(\tilde{k},k)}, (n/\tilde{k}) \log(n)\right) \geq \min\left(t_{\text{large-hit}}^{(\tilde{k},k)}, n^{2/d}\right), \tag{43}$$

for some constant $C > 0$. For $t_{\text{large-hit}}^{(\tilde{k},k)}$, it follows from Lemma 4.11 where we fix u to be any vertex and S to be the complement of the ball of radius $n^{1/d}/10$ around u that

$$t_{\text{large-hit}}^{(\tilde{k},k)} = \Omega\left(\frac{n^{2/d}}{\log(k/\tilde{k})}\right) = \Omega\left(\frac{n^{2/d}}{\log \frac{k}{n^{1-2/d} \log n}}\right),$$

hence by (43) we obtain a matching lower bound. $\qquad\square$

6. Conclusion and open problems

In this work, we derived several new bounds on multiple stationary and worst-case cover times. We also introduced a new quantity called *partial mixing time*, which extends the definition of mixing time from single random walks to multiple random walks. By means of a min–max characterisation, we proved that the partial mixing time connects the stationary and worst-case cover times, leading to tight lower and upper bounds for many graph classes.

In terms of worst-case bounds, Theorem 3.1 implies that for any regular graph G and any $k \geq 1$, $t_{\text{cov}}^{(k)}(\pi) = \mathcal{O}\left(\left(\frac{n}{k}\right)^2 \log^2 n\right)$. This bound is tight for the cycle when k is polynomial in n but not for smaller k. We suspect that for any $k \geq 1$ the cycle is (asymptotically) the worst-case for $t_{\text{cov}}^{(k)}(\pi)$ amongst regular graphs, which suggests $t_{\text{cov}}^{(k)}(\pi) = \mathcal{O}\left(\left(\frac{n}{k}\right)^2 \log^2 k\right)$.

Some of our results have been only proven for the independent stationary case, but it seems plausible they extend to the case where the k random walks start from the *same* vertex. For example, extending the bound $t_{\text{cov}}^{(k)}(\pi) = \Omega((n/k)\log n)$ to this case would be very interesting.

Although our min–max characterisations involving partial mixing time yields tight bounds for many natural graph classes, it would be interesting to establish a general approximation guarantee (or find graph classes that serve as counter-examples). For the former, we believe techniques such as Gaussian Processes and Majorising Measures used in the seminal work of Ding, Lee and Peres [17] could be very useful.

Acknowledgements

References

[1] Aldous, D. and Diaconis, P. (1987) Strong uniform times and finite random walks. *Adv. Appl. Math.* **8**(1) 69–97.

[2] Aldous, D. and Fill, J. A. (2002) Reversible Markov chains and random walks on graphs. Unfinished monograph, recompiled 2014.

[3] Aldous, D. J. (1989) Lower bounds for covering times for reversible Markov chains and random walks on graphs. *J. Theoret. Probab.* **2**(1) 91–100.

[4] Aleliunas, R., Karp, R. M., Lipton, R. J., Lovász, L. and Rackoff, C. (1979) Random walks, universal traversal sequences, and the complexity of maze problems. In *20th Annual Symposium on Foundations of Computer Science, FOCS 1979*, IEEE Computer Society, pp. 218–223.

[5] Alon, N., Avin, C., Koucký, M., Kozma, G., Lotker, Z. and Tuttle, M. R. (2011) Many random walks are faster than one. *Combin. Probab. Comput.* **20**(4) 481–502.

[6] Andersen, R., Chung, F. R. K. and Lang, K. J. (2006) Local graph partitioning using pagerank vectors. In *Proceedings of the 47th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2006)*, IEEE Computer Society, pp. 475–486.

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

[8] Boczkowski, L., Guinard, B., Korman, A., Lotker, Z. and Renault, M. P. (2018) Random walks with multiple step lengths. In *LATIN 2018: Theoretical Informatics - 13th Latin American Symposium*, (Bender, M. A., Farach-Colton, M. and Mosteiro, M. A., eds), *Vol. 10807 of Lecture Notes in Computer Science*, Springer, pp. 174–186.

[9] Broder, A. Z., Karlin, A. R., Raghavan, P. and Upfal, E. (1994) Trading space for time in undirected s-t connectivity. *SIAM J. Comput.* **23**(2) 324–334.

[10] Clementi, A. E. F., D'Amore, F., Giakkoupis, G. and Natale, E. (2021) Search via parallel lévy walks on Z^2. In *PODC '21: ACM Symposium on Principles of Distributed Computing, 2021*, (Miller, A., Censor-Hillel, K. and Korhonen, J. H., eds), ACM, pp. 1–91.

[11] Cooper, C. (2011) Random walks, interacting particles, dynamic networks: Randomness can be helpful. In *Structural Information and Communication Complexity - 18th International Colloquium, SIROCCO 2011*, (Kosowski, A. and Yamashita, M., eds), *Vol. 6796 of Lecture Notes in Computer Science*, Springer, pp. 1–14.

[12] Cooper, C. and Frieze, A. (2007) The cover time of the preferential attachment graph. *J. Combin. Theory Ser. B* **97**(2) 269–290.

[13] Cooper, C. and Frieze, A. (2014) A note on the vacant set of random walks on the hypercube and other regular graphs of high degree. *Mosc. J. Comb. Number Theory* **4**(4) 21–44.

[14] Cooper, C., Radzik, T. and Siantos, Y. (2014) Estimating network parameters using random walks. *Social Netw. Analys. Mining* **4**(1) 168.

[15] Czumaj, A., Monemizadeh, M., Onak, K. and Sohler, C. (2019) Planar graphs: Random walks and bipartiteness testing. *Random Struct. Algorithms* **55**(1) 104–124.

[16] Czumaj, A. and Sohler, C. (2010) Testing expansion in bounded-degree graphs. *Comb. Probab. Comput.* **19**(5-6) 693–709.

[17] Ding, J., Lee, J. R. and Peres, Y. (2012) Cover times, blanket times, and majorizing measures. *Ann. Math.* **175**(3) 1409–1471.

[18] Dubhashi, D. P. and Panconesi, A. (2009) *Concentration of Measure for the Analysis of Randomized Algorithms*. Cambridge University Press, Cambridge.

[19] Efremenko, K. and Reingold, O. (2009) How well do random walks parallelize?. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, 12th International Workshop, APPROX 2009, and 13th International Workshop, RANDOM 2009*, (Dinur, I., Jansen, K., Naor, J. and Rolim, J. D. P., eds), *Vol. 5687 of Lecture Notes in Computer Science*, Springer, pp. 476–489.

[20] Elsässer, R. and Sauerwald, T. (2011) Tight bounds for the cover time of multiple random walks. *Theoret. Comput. Sci.* **412**(24) 2623–2641.

[21] Feige, U. (1995) A tight lower bound on the cover time for random walks on graphs. *Random Struct. Algorithms* **6**(4) 433–438.

[22] Feige, U. (1997) A spectrum of time-space trade-offs for undirected s-t connectivity. *J. Comput. Syst. Sci.* **54**(2) 305–316.

[23] Georgakopoulos, A., Haslegrave, J., Sauerwald, T. and Sylvester, J. (2022) The power of two choices for random walks. *Comb. Probab. Comput.* **31**(1) 73–100.

[24] Gharan, S. O. and Trevisan, L. (2012) Approximating the expansion profile and almost optimal local graph clustering. In *53rd Annual IEEE Symposium on Foundations of Computer Science, FOCS 2012*, IEEE Computer Society, pp. 187–196.

[25] Gkantsidis, C., Mihail, M. and Saberi, A. (2005) Hybrid search schemes for unstructured peer-to-peer networks. In *INFOCOM 2005. 24th Annual Joint Conference of the IEEE Computer and Communications Societies*, IEEE, pp. 1526–1537.

[26] Gordon, R. D. (1941) Values of Mills' ratio of area to bounding ordinate and of the normal probability integral for large values of the argument. *Ann. Math. Statist.* **12**(3) 364–366.

[27] Guinard, B. and Korman, A. (2020) Tight bounds for the cover times of random walks with heterogeneous step lengths. In *37th International Symposium on Theoretical Aspects of Computer Science, STACS 2020*, (Paul, C. and Bläser, M., eds), *Vol. 154 of LIPIcs*, Schloss Dagstuhl - Leibniz-Zentrum für Informatik, pp. 28:1–28:14.

[28] Hebisch, W. and Saloff-Coste, L. (1993) Gaussian estimates for Markov chains and random walks on groups. *Ann. Probab.* **21**(2) 673–709.

[29] Hermon, J. (2018) Frogs on trees? *Electron. J. Probab.* **23** Paper No. 17, 40.

[30] Hermon, J. and Sousi, P. (2021) Covering a graph with independent walks.

[31] Ivaskovic, A., Kosowski, A., Pajak, D. and Sauerwald, T. (2017) Multiple random walks on paths and grids. In *34th Symposium on Theoretical Aspects of Computer Science, STACS 2017*, (Vollmer, H. and Vallée, B., eds), Vol. 66 of LIPIcs, Schloss Dagstuhl - Leibniz-Zentrum für Informatik, pp. 44:1–44:14.

[32] Janson, S. (2018) Tail bounds for sums of geometric and exponential variables. *Statist. Probab. Lett.* **135** 1–6.

[33] Kahn, J., Kim, J. H., Lovász, L. and Vu, V. H. (2000) The cover time, the blanket time, and the Matthews bound. In *41st Annual Symposium on Foundations of Computer Science (Redondo Beach, CA, 2000)*, IEEE Computer Society Press, Los Alamitos, CA, pp. 467–475.

[34] Kahn, J. D., Linial, N., Nisan, N. and Saks, M. E. (1989) On the cover time of random walks on graphs. *J. Theoret. Probab.* **2**(1) 121–128.

[35] Karger, D. R. and Ruhl, M. (2004) Simple efficient load balancing algorithms for peer-to-peer systems. In *SPAA 2004: Proceedings of the Sixteenth Annual ACM Symposium on Parallelism in Algorithms and Architectures*, (Gibbons, P. B. and Adler, M., eds), ACM, pp. 36–43.

[36] Kempe, D., Kleinberg, J. M. and Demers, A. J. (2001) Spatial gossip and resource location protocols. In *Proceedings on 33rd Annual ACM Symposium on Theory of Computing, STOC 2001*, (Vitter, J. S., Spirakis, P. G. and Yannakakis, M., eds), ACM, pp. 163–172.

[37] Klasing, R., Kosowski, A., Pajak, D. and Sauerwald, T. (2013) The multi-agent rotor-router on the ring: a deterministic alternative to parallel random walks. In *ACM Symposium on Principles of Distributed Computing, PODC '13*, (Fatourou, P. and Taubenfeld, G., eds), ACM, pp. 365–374.

[38] Akash Kumar, C. S. and Stolman, A. (2018) Finding forbidden minors in sublinear time: An$^{1/2+o(1)}$-query one-sided tester for minor closed properties on bounded degree graphs. In *59th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2018*, (Thorup, M., ed), IEEE Computer Society, pp. 509–520.

[39] Akash Kumar, C. S. and Stolman, A. (2019) Random walks and forbidden minors II: apoly(d/ϵ)-query tester for minor-closed properties of bounded degree graphs. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing, STOC 2019*, (Charikar, M. and Cohen, E., eds), ACM, pp. 559–567.

[40] Lacki, J., Mitrovic, S., Onak, K. and Sankowski, P. (2020) Walking randomly, massively, and efficiently. In *Proccedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing, STOC 2020*, (Makarychev, K., Makarychev, Y., Tulsiani, M., Kamath, G. and Chuzhoy, J., eds), ACM, pp. 364–377.

[41] Lam, H., Liu, Z., Mitzenmacher, M., Sun, X. and Wang, Y. (2012) Information dissemination via random walks in d-dimensional space. In *Proceedings of the Twenty-Third Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2012*, (Rabani, Y., ed), SIAM, pp. 1612–1622.

[42] Levin, D. A., Peres, Y. and Wilmer, E. L. (2009) *Markov Chains and Mixing Times*. American Mathematical Society, Providence, RI, With a chapter by James G. Propp and David B. Wilson.

[43] Lezaud, P. (1998) Chernoff-type bound for finite Markov chains. *Ann. Appl. Probab.* **8**(3) 849–867.

[44] Lovász, L. (1996) Random walks on graphs: a survey. In *Combinatorics, Paul Erdős is Eighty, Vol. 2 (Keszthely, 1993)*, *Vol. 2 of Bolyai Society of Mathematical Studies*, Budapest: János Bolyai Mathematical Society, pp. 353–397.

[45] Lv, Q., Cao, P., Cohen, E., Li, K. and Shenker, S. (2002) Search and replication in unstructured peer-to-peer networks. In *Proceedings of the 16th international conference on Supercomputing, ICS 2002*, (Ebcioglu, K., Pingali, K. and Nicolau, A., eds), ACM, pp. 84–95.

[46] Lyons, R. and Gharan, S. O. (2018) Sharp bounds on random walk eigenvalues via spectral embedding. *Int. Math. Res. Not. IMRN* **2018**(24) 7555–7605.

[47] Lyons, R. and Peres, Y. (2016) *Probability on Trees and Networks*. Cambridge Series in Statistical and Probabilistic Mathematics, 42. Cambridge University Press, xv+699 pp.

[48] Mihail, M., Papadimitriou, C. H. and Saberi, A. (2006) On certain connectivity properties of the internet topology. *J. Comput. Syst. Sci.* **72**(2) 239–251.

[49] Mitzenmacher, M. and Upfal, E. (2017) *Probability and Computing: Randomization and Probabilistic Techniques in Algorithms and Data Analysis*. Second edition. Cambridge University Press, xx+467 pp.

[50] Motwani, R. and Raghavan, P. (1995) *Randomized Algorithms*. Cambridge University Press, Cambridge.

[51] Oliveira, R. I. and Peres, Y. (2019) Random walks on graphs: new bounds on hitting, meeting, coalescing and returning. In *Proceedings of the Sixteenth Workshop on Analytic Algorithmics and Combinatorics, ANALCO 2019*, (Mishna, M. and Munro, J. I., eds), SIAM, pp. 119–126.

[52] Oliveira, R. I. (2012) Mixing and hitting times for finite Markov chains. *Electron. J. Probab.* **17**(70) 12.

[53] Patel, R., Carron, A. and Bullo, F. (2016) The hitting time of multiple random walks. *SIAM J. Matrix Anal. Appl.* **37**(3) 933–954.

[54] Peres, Y. and Sousi, P. (2015) Mixing times are hitting times of large sets. *J. Theoret. Probab.* **28**(2) 488–519.

[55] Pettarin, A., Pietracaprina, A., Pucci, G. and Upfal, E. (2011) Tight bounds on information dissemination in sparse mobile networks. In *Proceedings of the 30th Annual ACM Symposium on Principles of Distributed Computing, PODC 2011*, (Gavoille, C. and Fraigniaud, P., eds), ACM, pp. 355–362.

[56] Rivera, N., Sauerwald, T. and Sylvester, J. (2021) Multiple random walks on graphs: Mixing few to cover many. In *48th International Colloquium on Automata, Languages, and Programming, ICALP 2021*, (Bansal, N., Merelli, E. and Worrell, J., eds), Vol. 198 of LIPIcs, Schloss Dagstuhl - Leibniz-Zentrum für Informatik, pp. 99:1–99:16.

[57] Sarma, A. D., Nanongkai, D., Pandurangan, G. and Tetali, P. (2013) Distributed random walks. *J. ACM* **60**(1) 31.

[58] Sauerwald, T. (2010) Expansion and the cover time of parallel random walks. In *Proceedings of the 29th Annual ACM Symposium on Principles of Distributed Computing, PODC 2010*, (Richa, A. W. and Guerraoui, R., eds), ACM, pp. 315–324.

[59] Sauerwald, T. and Sun, H. (2012) Tight bounds for randomized load balancing on arbitrary network topologies. In *53rd Annual IEEE Symposium on Foundations of Computer Science, FOCS 2012*, IEEE Computer Society, pp. 341–350.

[60] Spielman, D. A. and Teng, S.-H. (2013) A local clustering algorithm for massive graphs and its application to nearly linear time graph partitioning. *SIAM J. Comput.* **42**(1) 1–26.

A. Appendix: Elementary Results

Lemma A.1. *Let X be a non-negative integer random variable such that $\mathbb{E}[X] \geq b$ and there exists $c \geq 0$ such that $\mathbb{P}[X > \ell c] \leq \mathbb{P}[X > c]^{\ell}$ for all integers $\ell \geq 0$. Then for any $a < c$*

$$\mathbb{P}[X > a] \geq \frac{b-a}{b+2c}.$$

Proof. Let $p = \mathbb{P}[X > a] \geq \mathbb{P}[X > c]$ as $\mathbb{P}[X > x]$ is non-increasing in x. Now we have

$$b \leq a + \sum_{i=a+1}^{c-1} \mathbb{P}[x > i] + c \sum_{\ell=1}^{\infty} \mathbb{P}[X > c]^{\ell} \leq a + p(c-a) + cp/(1-p).$$

This implies $(1-p)b \leq a + 2pc$, rearranging gives the result. □

A.1. Returns in the torus

The following result is well-known, however we state it for completeness.

Lemma A.2. *For the d-dimensional torus $\mathbb{T}_d$ on n vertices and any $1 \leq t \leq t_{\text{rel}}$ we have*

$$\sum_{i=0}^{t} P_{v,v}^i = \begin{cases} \Theta(\sqrt{t}) & \text{if } d=1 \\ \Theta(1 + \log t) & \text{if } d=2 \\ \Theta(1) & \text{if } d \geq 3 \end{cases}.$$

Proof. We begin with the lower bounds. Let Q and P be the transition matrices of the lazy walk on the d-dimensional integer lattice $\mathbb{Z}^d$ and the d-dimensional torus $\mathbb{T}_d$, respectively. By [28, Theorem 5.1 (15)], for each $d \geq 1$ there exist constants $C > 0$ such that for any $t \geq 1$ and $v \in \mathbb{Z}^d$ we have $Q_{v,v}^t \geq (C/t)^{d/2}$. The lower bounds for the torus then follow by summation since for any $t \geq 0$ and $v \in V(\mathbb{T}_d)$ we have $P_{v,v}^t \geq Q_{v,v}^t$.

We now prove the three cases for the upper bounds separately.

Case (i) [$d = 1$, cycle]: For a lazy random walk in the cycle it holds for $t \geq 1$ that $P_{v,v}^t \leq \frac{1}{n} + \frac{c}{\sqrt{t}}$, for some constant $c > 0$ [46, Theorem 4.9]. Now, for any $C \geq 1$, and any $1 \leq t \leq Cn^2$, it holds for some $c' > 0$ that

$$\sum_{i=0}^{t} P_{v,v} \leq c'\sqrt{t} + C(t+1)/n^2 = O(\sqrt{t}).$$

Case (ii) [$d = 2$]: A random walk of length t in two dimension can be generated as follows; first sample a random integer x according to $B(t) \sim \text{Bin}(t, 1/2)$, where x is the number of lazy random walk steps the walk takes in the first dimension (so $t - x$ is the number of lazy random walk steps the walk takes in the second dimension). Hence, if Q denotes the law of a lazy random walk on a cycle with $\sqrt{n}$ vertices, we have $P_{v,v}^t = \mathbb{E}\left[Q_{v,v}^{B(t)} Q_{v,v}^{t-B(t)}\right]$. Observe that for real functions f and g, non-increasing and non-decreasing respectively, random variables B and B', where B' is an independent copy of B, we have

$$2\left(\mathbb{E}\left[f(B)g(B)\right] - \mathbb{E}\left[f(B)\right]\mathbb{E}\left[g(B)\right]\right) = \mathbb{E}\left[(f(B) - f(B'))(g(B) - g(B'))\right] \leq 0.$$

Note that $s \to Q_{v,v}^s$ is non-increasing [42, Exercise 12.5], so $s \to Q_{v,v}^{t-s}$ is non-decreasing, hence

$$P_{v,v}^t = \mathbb{E}\left[Q_{v,v}^{B(t)} Q_{v,v}^{t-B(t)}\right] \leq \mathbb{E}\left[Q_{v,v}^{B(t)}\right] \cdot \mathbb{E}\left[Q_{v,v}^{t-B(t)}\right] \leq \mathbb{E}\left[Q_{v,v}^{B(t)}\right]^2.$$

As $s \to Q_{v,v}^s$ is non-increasing, we have $Q_{v,v}^s \leq Q_{v,v}^{\lfloor t/4 \rfloor}$ for any $s \geq t/4$. Then, for $t \geq 1$ but $t = \mathcal{O}(n)$, it holds $Q_{v,v}^{\lfloor t/4 \rfloor} \leq c_1/\sqrt{t}$ for some constant $c_1 > 0$ by [46, Theorem 4.9]. Thus, by Hoeffding's bound,

$$\mathbb{E}\left[Q_{v,v}^{B(t)}\right]^2 \leq \frac{c_1^2}{t} + \mathbb{P}\left[\text{Bin}(t, 1/2) \leq t/4\right] \leq \frac{c_1^2}{t} + \exp(-t/2) \leq \frac{c_2}{t},$$

for some constant $c_2 > 0$. The statement of Case (ii) then follows by summation.

Case (iii) [$d \geq 3$]: By [42, Proposition 10.13] there is a constant $C > 0$ such that $\mathbb{E}_u[\tau_v] \leq C \cdot n$ for all $u, v \in V$. Hence by Markov's inequality,

$$\mathbb{P}_u[\tau_v \geq 2C \cdot n] \leq 1/2.$$

Therefore, for $t = 2C \cdot n$, and by averaging over the start vertex,

$$1/2 \leq \mathbb{P}_\pi\left[\tau_v \leq t\right] = \frac{\mathbb{E}_\pi\left[N_v^t\right]}{\mathbb{E}_\pi\left[N_v^t \mid N_v^t \geq 1\right]} \leq \frac{t \cdot 1/n}{1/2 \cdot \sum_{i=0}^{t/2} P_{v,v}^i},$$

and rearranging yields

$$\sum_{i=0}^{t} P_{v,v}^i \leq 2 \sum_{i=0}^{t/2} P_{v,v}^i \leq 16C,$$

where the first inequality holds by monotonicity of $P_{v,v}^i$ in $i \geq 0$. $\qquad \square$

A.2. Returns in the binary tree

To prove the results for the binary tree we must control the return probabilities of single random walks, we gather the results required for this task here.

Recall that if $R(x, y)$ is the effective resistance between x and y, (see [42, Section 9.4]), then for any $x, y \in V$ by [42, Proposition 9.5]:

$$\mathbb{P}_x\big[\tau_y < \tau_x^+\big] = \frac{1}{d(x)R(x, y)}. \qquad (44)$$

Lemma A.3. *Let r be the root of a binary tree of height $h \geq 4$, and $\mathcal{L}$ be the set of leaves. Then, for any $T \leq n$, $\sum_{t=0}^{T} P_{r,r}^t \leq 2 + 6T/n$. Additionally $\mathbb{P}_\pi[\tau_r \leq \tau_\mathcal{L}] \leq (8 \log n)/n$.*

Proof. Identify $\mathcal{L}$ as a single vertex and observe that $R(r, \mathcal{L}) = \sum_{i=1}^{h} (1/2)^i = 1 - 1/2^{h+1}$, thus $\mathbb{P}_r[\tau_\mathcal{L} < \tau_r^+] \geq 1/2$. Once the walk hits $\mathcal{L}$ equation (44) yields $\mathbb{P}_\mathcal{L}[\tau_r < \tau_\mathcal{L}^+] = (2^h(1 - 1/2^{h+1}))^{-1} \leq 3/n$. Let X_t be a random walk on the binary tree starting from vertex r, i.e. $X_0 = r$, and let $Z_t = \sum_{s=0}^{t} 1_{\{X_s=r\}}$ denotes the number of times X_t hits the root up to time t. Define the stopping time L_i as the i-th time the random walk hits some leaf, that is: $L_1 = \min\{t \geq 0 : X_t \in \mathcal{L}\}$, and for $i \geq 2$, $L_i = \min\{t > L_{i-1} : X_t \in \mathcal{L}\}$. For $i \geq 1$ let $C_i = \sum_{t=L_i+1}^{L_{i+1}} 1_{\{X_t=r\}}$ denote the number of visits to the root between times L_i and L_{i+1}, and also let $C_0 = \sum_{t=0}^{L_1} 1_{\{X_t=r\}}$. Since $L_i - L_{i-1} \geq 1$, then have that

$$\sum_{t=0}^{T} P_{r,r}^t = \mathbb{E}(Z_T) \leq \mathbb{E}(C_0) + \sum_{i=1}^{T} \mathbb{E}(C_i).$$

Now, $\mathbb{E}[C_0] \leq 2$ since $\mathbb{P}_r[\tau_\mathcal{L} < \tau_r^+] \geq 1/2$ and $\mathbb{E}[C_i | C_i \geq 1] = \mathbb{E}[C_0]$ as in the interval C_0 the walk starts from the root. Also for $i \geq 1$, $\mathbb{P}[C_i \geq 1] = \mathbb{P}_\mathcal{L}[\tau_r < \tau_\mathcal{L}^+] \leq 3/n$, and thus

$$\mathbb{E}[C_i] \leq (3/n)\mathbb{E}[C_i | C_i \geq 1] = (3/n)\mathbb{E}[C_0] \leq 6/n,$$

concluding that $\sum_{t=0}^{T} P_{r,r}^t \leq 2 + 6T/n$.

For the second result, let v_i be a vertex at distance $0 \leq i \leq h$ from the leaves. Since the walk moves up the tree with probability $1/3$ and down the tree w.p. $2/3$ we have $\mathbb{P}_{v_i}[\tau_r < \tau_\mathcal{L}] = \frac{2^i - 1}{2^h - 1}$, since this is the classical (biased) Gambler's ruin problem [42, Section 17.3.1]. It follows that

$$\mathbb{P}_\pi[\tau_r < \tau_\mathcal{L}] = \sum_{i=1}^{h} \mathbb{P}_{v_i}[\tau_r < \tau_\mathcal{L}] \cdot 2^{h-i} \pi(v_i) \leq \sum_{i=1}^{h} \frac{2^i}{2^h - 1} \cdot 2^{h-i} \cdot \frac{3}{2(n-1)} = \frac{3h}{n-1},$$

where the last equality holds as $2^{h+1} - 1 = n$ (and so $2^h = (n-1)/2$). Since $h \geq 4$, the number of vertices is at least 15, and then $h = \log_2(n+1) - 1 \leq 2 \log n$. Finally, since $n - 1 \geq (6/8)n$, for $n \geq 8$, it holds that $\mathbb{P}_\pi[\tau_r < \tau_\mathcal{L}] \leq \frac{3h}{(n-1)} \leq 8 \log n/n$. $\qquad \square$

Lemma A.4. *([29, eq 8.21]). Let u be any leaf in the binary tree. Then for any $t \geq 1$, $\sum_{i=0}^{t} P_{u,u}^t = \Theta(1 + \log(t) + t/n)$.*

Lemma A.5. *Let $\ell \in V(\mathcal{T}_n)$ be a leaf. Then for any vertex $u \in V(\mathcal{T}_n)$ and $t \geq 0$ we have $\sum_{i=0}^{t} P_{u,u}^i \leq 6 \cdot \sum_{i=0}^{t} P_{\ell,\ell}^i$.*

Proof. Clearly, if u is a leaf itself, then there is nothing to prove. Hence assume that u is an internal node. Note that $\sum_{s=0}^{t} P_{u,u}^s$ is the expected number of visits to u of a random walk of length t starting from u. Divide the random walk of length t into two epochs, where the second epochs starts as soon as a leaf in the subtree rooted at u is visited. We claim that in the first epoch the expected number of visits to u is constant. To show this we identify all the leaves of the subtree rooted at u as a single vertex $\mathcal{L}$, then by equation (44) we have

$$\mathbb{P}_u\big[\tau_{\mathcal{L}} < \tau_u^+\big] = \frac{1}{d(u)R(u, \mathcal{L})} \geq \frac{1}{3}$$

since $R(u, \mathcal{L}) = \sum_{i=1}^{h} (1/2)^i \leq 1$. Therefore, in expectation we need at most 3 excursions to reach $\mathcal{L}$, therefore the expected number of visits to u in the first epoch is at most 3. Then, for any leaf $\ell \in \mathcal{L}$, the expected number of visits to u in the second epoch satisfies

$$\sum_{s=1}^{t} P_{\ell,u}^s \leq 3 \sum_{s=1}^{t} P_{u,\ell}^s \leq 3 \sum_{s=0}^{t} P_{\ell,\ell}^s,$$

where the first inequality holds by reversibility and the second inequality holds since the expected number of visits to a vertex is maximised if a random walk starts from that vertex. Adding up the expected number of visits from the two epochs yields the claim. $\qquad \square$