Tight Bounds for the Cover Time of Multiple Random Walks*

Robert Elsässer[1] and Thomas Sauerwald[2]

[1] Department of Computer Science, University of Paderborn, Germany
[2] International Computer Science Institute, Berkeley, CA, USA

Abstract. We study the cover time of multiple random walks. Given a graph G of n vertices, assume that k independent random walks start from the same vertex. The parameter of interest is the *speed-up* defined as the ratio between the cover time of one and the cover time of k random walks. Recently Alon et al. developed several bounds that are based on the quotient between the cover time and maximum hitting times. Their technique gives a speed-up of $\Omega(k)$ on many graphs, however, for many graph classes, k has to be bounded by $\mathcal{O}(\log n)$. They also conjectured that, for any $1 \leqslant k \leqslant n$, the speed-up is at most $\mathcal{O}(k)$ on any graph. As our main results, we prove the following:

- We present a new lower bound on the speed-up that depends on the mixing-time. It gives a speed-up of $\Omega(k)$ on many graphs, even if k is as large as n.
- We prove that the speed-up is $\mathcal{O}(k \log n)$ on any graph. Under rather mild conditions, we can also improve this bound to $\mathcal{O}(k)$, matching exactly the conjecture of Alon et al.
- We find the correct order of the speed-up for any value of $1 \leqslant k \leqslant n$ on hypercubes, random graphs and expanders. For d-dimensional torus graphs ($d > 2$), our bounds are tight up to a factor of $\mathcal{O}(\log n)$.
- Our findings also reveal a surprisingly sharp dichotomy on several graphs (including d-dim. torus and hypercubes): up to a certain threshold the speed-up is k, while there is no additional speed-up above the threshold.

1 Introduction

Random walks come up and are studied in many sciences like mathematics, physics, computer science etc. While mathematicians have studied random walks on *infinite* graphs for a long time, computer scientists have spurred an interest on random walks on *finite* graphs during the last two decades. Roughly speaking,

* The first author was partially supported by the German Research Foundation under contract EL 399/2-1, and by the Integrated Project IST 15964 "Algorithmic Principles for Building Efficient Overlay Networks" of the EU. The second author was partially supported by a postdoctoral fellowship from the German Academic Exchange Service (DAAD).

S. Albers et al. (Eds.): ICALP 2009, Part I, LNCS 5555, pp. 415–426, 2009.

there have been two main lines of research. One is concerned with the development of rapidly mixing random walks, resulting in approximation schemes of #P hard problems (cf. [17] for more details and a survey on random walks). The second line of research deals with the time to explore a graph, formally known as cover time.

Random walks are an attractive tool for graph exploration due to their inherent simplicity, locality and robustness to dynamical changes. For example, Avin, Koucky, and Lotker [4] recently proved that a (slighly modified) random walk can still explore all vertices of a graph efficiently, even if the graph is dynamically changing during the covering procedure. Other algorithmic applications where random walks have been used are searching [13], routing [18], gossiping [16] and self-stabilization [12] etc.

Probably the first theoretical applications of the cover time traces back to Aleliunas, Karp, Lipton, Lovász, and Rackoff [2]. It was shown that by taking a random walk, it is possible to explore every undirected graph in polynomial time and logarithmic space. In response to their question about time-space tradeoffs, Broder, Karlin, Raghavan, and Upfal [7] studied the cover time of many, independent random walks, each of which starts from the stationary distribution.

Certainly, the situation becomes more challenging if all random walks start from the same vertex. Will they stick together and cover more or less the same set of vertices, or will they quickly disperse in different regions to ensure a fast covering? Alon, Avin, Koucky, Kozma, Lotker, and Tuttle [3] posed this question and studied the speed-up defined as the ratio between the cover time of a single random walk and the cover time of k random walks, where $1 \leqslant k \leqslant n$. As it turns out, the answer depends very much on the underlying graph: on complete graphs, a speed-up of k is always possible, while on the cycle the speed-up is only $\mathcal{O}(\log k)$. On certain graphs, there are even starting positions of the k walks such that the speed-up is $\Omega(2^k)$ (for small k).

Another reason why the cover time of random walks has been investigated is its intimate relation to other graph-theoretical parameters. For example, Broder and Karlin [6] gave a comprehensive collection of bounds relating the cover time to spectral properties of G. Chandra, Raghavan, Ruzzo, Smolensky, and Tiwari [9] established a tight connection between random walks and electrical networks and related the cover time to other properties such as the vertex-expansion.

1.1 Related Work

One slight drawback of the cover time of a single random walk is that it takes at least $\Omega(n \log n)$ steps on every graph, and may even increase to $\Omega(n^2)$ on regular and $\Omega(n^3)$ steps on non-regular graphs. This has led to several modified covering schemes. Adler, Halperin, Karp, and Vazirani [1] introduced a covering process where in each round one first chooses a vertex uniformly at random, and then chooses an uncovered neighbor of this vertex (if there is one). Later Dimitrov and Plaxton [11] proved that this process achieves a cover time of $\mathcal{O}(n + (n \log n)/d)$ on any d-regular graph. Note that in this scheme, one has to sample uniformly among *all* vertices (not just among visited neighbors) which is not completely inline with the scenario of a decentralized exploration process.

Another approach taken by Ikeda, Kubo, Okumoto, and Yamashita [14] and Avin and Krishnamachari [5] is to change the transition probabilities of the random walk. For example, Ikeda et al. [14] devised a way of locally computable transition probabilities which results in a cover time of $\mathcal{O}(n^2 \log n)$ on any graph. However, one limitation of all these approaches is that they can only explore a graph within $\Omega(n)$ steps.

Multiple random walks can break this barrier of $\Omega(n)$ and have been used by Broder et al. [7] to obtain tradeoffs between space and time for the s-t-connectivity problem. As mentioned before, they assumed that each random walk starts from an independent sample of the stationary distribution. While this indeed significantly speeds up the covering process, one has to sample again among *all* vertices. This could be one reason why researchers have recently studied multiple random walks which start all from the same vertex ([3, 10]). Alon et al. [3] derived several (asymptotic) lower and upper bounds on the speed-up on several graph classes, while Cooper et al. [10] focused on the class of random regular graphs and derived nearly exact bounds on the speed-up. Finally, multiple random walks starting from the same vertex are also a fundamental tool for property testing, cf. [15] for a recent analysis of a property tester of expanders. The basic idea is to count the collisions of random walks that start from the same vertex to estimate the expansion properties of a graph.

1.2 Our Contribution

Before describing our main results, we have to introduce a little bit of notation. Let G be any undirected, connected graph with n vertices. For any $1 \leqslant k \leqslant n$, let $\mathbf{E}\left[\,\mathsf{COV}_u^k\,\right]$ be the expected time for k random walks that start from u to cover all vertices. Let $\mathbf{E}\left[\,\mathsf{COV}_{\max}^k\,\right] = \max_{u \in V} \mathbf{E}\left[\,\mathsf{COV}_u^k\,\right]$ (we also use $\mathbf{E}\left[\,\mathsf{COV}_{\max}(G)\,\right] = \mathbf{E}\left[\,\mathsf{COV}_{\max}^1(G)\,\right]$ to stick to the common notation). For any undirected, connected graph G, we define the *speed-up* $S^k := \mathbf{E}\left[\,\mathsf{COV}_{\max}(G)\,\right] / \mathbf{E}\left[\,\mathsf{COV}_{\max}^k(G)\,\right]$. By $\mathsf{H}(u,v)$ we denote the expected time for the random walk to get from u to v; it is a well-known fact that $\max_{u,v} \mathsf{H}(u,v)$ approximates $\mathbf{E}\left[\,\mathsf{COV}_{\max}(G)\,\right]$ up to a factor of $\mathcal{O}(\log n)$ (see Theorem 2.2). The mixing time $\mathsf{MIX}_{1/2}(G)$ is the time required for the random walk to approach its stationary distribution (exact definition in Section 2).

We first present a general lower bound on the speed-up. It is based on the following upper bound on $\mathbf{E}\left[\,\mathsf{COV}_{\max}^k(G)\,\right]$:

Theorem 3.4 (from page 421)**.** *For any graph G and any k with $1 \leqslant k \leqslant n$,*

$$\mathbf{E}\left[\,\mathsf{COV}_{\max}^k(G)\,\right] = \mathcal{O}\left(\frac{\log n \cdot (\max_{u,v} \mathsf{H}(u,v) + \mathsf{MIX}_{1/2}(G))}{k} + \mathsf{MIX}_{1/2}(G)\right).$$

Graph	$\mathsf{COV}(G)$	$\mathsf{H}_{\max}$	$\mathsf{MIX}_{1/2}(G)$	Speed up $S^k(G)$ $k \in$	bounds
cycle	n^2	n^2	n^2	$[1, n]$	$= \log k$ [3, Thm. 6]
2-dim. torus	$n \log^2 n$	$n \log n$	n	$[1, \log n]$ $[1, n]$	$\geqslant k$ [3, Thm. 4] $\leqslant \log^2 n \log k$ [3, Cor. 25]
d-dim. torus, $d > 2$	$n \log n$	n	$n^{2/d}$	$[1, \log n]$ $[1, n^{1-2/d} \log n]$ $[1, n]$ $[1, n]$	$\geqslant k$ [3, Thm. 4] $\geqslant k$ [★, Cor. 1] $\leqslant n^{1-2/d} \log n \log k$ [3, Thm. 24] $\leqslant k$ [★, Cor. 4.6]
Hypercube	$n \log n$	n	$\log n \log \log n$	$[1, \log n]$ $[1, \frac{n}{\log \log n}]$ $[\frac{n}{\log \log n}, n]$	$\geqslant k$ [3, Thm. 4] $= k$ [★, Thm. 3.4 & Cor. 4.6] $= \frac{n}{\log \log n}$ [★, Thm. 3.4 & 5.3]
Complete	$n \log n$	n	1	$[1, n]$	$= k$ [3, Lem. 12]
Expander	$n \log n$	n	$\log n$	$[1, n]$ $[1, n]$	$\geqslant k$ [3, Thm. 18] $= k$ [★, Cor. 5.1]
Random	$n \log n$	n	$\log n$	$[1, \log n]$ $[1, n]$	$\geqslant k$ [3, Thm. 4] $= k$ [★, Cor. 5.1]

Fig. 1. Summary of the new and old results for the graphs mentioned by [3], where constant factors are neglected in all columns. $\mathsf{H}_{\max}$ stands for $\max_{u,v} \mathsf{H}(u, v)$. Our new results are marked with ★. For torus graphs, the bounds are tight up to a logarithmic factor and for all other graphs, the bounds are tight (for each $1 \leqslant k \leqslant n$).

This shows that $\mathbf{E}\left[\mathsf{COV}^k_{\max}(G)\right]$ is upper bounded by $\log n \cdot \max_{u,v} \mathsf{H}(u, v)/k$, as long as $\max_{u,v} \mathsf{H}(u, v)/k$ is not smaller than the mixing time (see Corollary 3.5 for a simpler, but slightly weaker statement than Theorem 3.4).

We point out that most previous general upper bounds on $\mathbf{E}\left[\mathsf{COV}^k_{\max}(G)\right]$ in Alon et al. [3] are at least $\Omega(n)$ [3, Thm. 4,5,13,14], and therefore only useful on most graphs when $k = \mathcal{O}(\log n)$. A similar bound to Theorem 3.4 from [3] is:

Theorem 1.1 ([3, Proof of Theorem 9]). *For any graph G and $1 \leqslant k \leqslant n$,*

$$\mathbf{E}\left[\mathsf{COV}^k_{\max}(G)\right] = \mathcal{O}\left(\frac{\mathsf{MIX}_{n^{-1}}(G)\, n\, (\log n)^2}{k}\right).$$

Note that the bound of Alon et al. [3] includes the mixing time as a factor, while in our bound (Theorem 3.4 above), for any k, the mixing time does not come into play at all, as long as $\max_{u,v} \mathsf{H}(u, v)/k$ is larger than $\mathsf{MIX}_{1/2}(G)$. Since for most graph classes (cmp. Figure 1) $\max_{u,v} \mathsf{H}(u, v) = \Theta(n)$ and $\mathsf{MIX}_{1/2}(G) = o(n)$, our theorem gives a lower bound on the speed-up of k for a wide range of k (cf. Figure 1, or Section 5 for more details).

The main idea to establish Theorem 3.4 is based on a coupling argument between one random walk and k random walks (see Theorem 3.3 for details). We believe that this technique might be very useful for deriving further bounds on the cover time of one or many random walks.

We continue to prove a general upper bound for any graph, namely that $S^k = \mathcal{O}(k \log n)$ for any $1 \leqslant k \leqslant n$. This already matches the conjecture of Alon et al. [3] up to a logarithmic factor. Under a rather mild condition on the mixing-time and cover time of one random walk, we improve this upper bound to $S^k = \mathcal{O}(k)$, establishing the conjecture of [3] for a large class of graphs

(Corollary 4.7). Finally, we also present an upper bound based on the diameter of the graph (Theorem 4.8).

Applications of our lower and upper bounds to concrete graphs are summarized in Figure 1, completing Table 1 of [3]. As an example, consider the hypercube with n vertices. We prove that $S^k = \Theta(k)$ as long as $k = \mathcal{O}(n/\log\log n)$. However, for $k = \Omega(n/\log\log n)$, $S^k = \Theta(n/\log\log n)$. The same dichotomy is established for d-dimensional torus graphs ($d > 2$), where also $n/\mathsf{MIX}_{1/2}(G)$ represents as a "sharp threshold" on the speed-up.

1.3 Road Map

In Section 2 we introduce our notation and some preliminary results. Section 3 contains the proof of our upper bound on S^k. This is followed by Section 4 consisting of several lower bounds on S^k. In Section 5 we show how to apply our general results to obtain tight bounds on S^k for concrete graph classes. We close in Section 6 with the conclusions. Several proofs are omitted due to space limitations.

2 Notations, Definitions and Preliminaries

Random Walk. A *random walk* (cf. [17] for a survey) on an undirected, connected graph $G = (V, E)$ starts at some specified vertex $u \in V$ and moves in each step along some adjacent edge chosen uniformly at random. To ensure convergence also on non-bipartite graphs, a common way is to add loop probabilities: at each step the random walk stays with probability 1/2 at the current vertex and otherwise it moves to a randomly chosen neighbor. It is a well-known fact that the loops only increase the cover time by a factor of 2.

There are two ways to represent the walk. The first and concrete one is to view the walk as an infinite sequence of vertices $X_0, X_1, \ldots$, where $X_0 = u$ is the starting vertex and X_t is the vertex visited at step t.

A more abstract way is to only consider the distribution of the walk. To this end, let $\mathbf{P}$ be the transition matrix of the walk, i.e., $p_{u,v} = \frac{1}{2\deg(u)}$ if $\{u, v\} \in E$, $p_{u,u} = \frac{1}{2}$ and $p_{u,v} = 0$ otherwise. Note that $\mathbf{P}$ is symmetric if and only if G is regular. Now define for each pair of vertices u, v, $p_{u,v}^t$ as the probability that a random walk starting at u visits the vertex v at step t. Hence the vector $p_u^t = (p_{u,v}^t)_{v \in V}$ represents the distribution of X_t, i.e., the visited vertex at step t. It is a well-known fact that under our assumptions on G, $\mathbf{p}_u(t)$ converges for $t \to \infty$ towards the *stationary distribution* π given by $\pi(v) = \deg(v)/(2|E|)$.

Mixing Time. To quantify the convergence speed, we define the *relative pointwise distance* ([20, p. 45]) as

$$\Delta(t) := \max_{u,v \in V} \frac{|p_{u,v}^t - \pi(v)|}{\pi(v)}.$$

Definition 1. *The* mixing time *of a random walk on G with transition matrix* $\mathbf{P}$ *is defined for any $0 < \varepsilon < 1$ by*

$$\mathsf{MIX}_{\varepsilon}^{\mathbf{P}}(G) := \min\{t \in \mathbb{N} \colon \Delta(t) \leqslant \varepsilon\}.$$

If the reference to $\mathbf{P}$ is obvious, we shall also just write $\mathsf{MIX}_{\varepsilon}(G)$. Our definition of mixing time should be compared with the one based on the variation distance used by Alon et al. [3], $\overline{\mathsf{MIX}}_{\varepsilon}(G) := \max_{u \in V} \min\{t \in \mathbb{N} \colon \|p_u^t - \pi\|_1 \leqslant \varepsilon\}$. The next lemma shows that $\mathsf{MIX}_{n^{-1}}(G)$ is not larger than $\overline{\mathsf{MIX}}_{n^{-1}}(G)$.

Lemma 2.1. *For any graph $G = (V, E)$, $\mathsf{MIX}_{n^{-1}}(G) = \mathcal{O}(\overline{\mathsf{MIX}}_{n^{-1}}(G))$.*

Hitting Time and Cover Time. For two vertices $u, v \in V(G)$, we define the *hitting time* from u to v as $\mathsf{H}(u, v) := \mathbf{E}[\min\{t \in \mathbb{N}\backslash\{0\} : X_t = v, X_0 = u\}]$, i. e., the expected number of steps to reach v from u. Denote by $\mathsf{COV}_s(G)$ the first time when a (single) random walk starting from s has visited all n vertices of G. Then the cover time is defined as $\mathbf{E}[\mathsf{COV}_{\max}(G)] := \max_{u \in V} \mathbf{E}[\mathsf{COV}_u(G)]$. (We point out that in several previous work the cover time is written without $\mathbf{E}[\cdot]$, however, in this work we also have to deal with the random variable $\mathsf{COV}_u(G)$). The following well-known result relates the maximum hitting time to the cover time.

Theorem 2.2 ([9, 18]). *For any graph $G = (V, E)$ we have* $\max_{u,v \in V} \mathsf{H}(u, v) \leqslant \mathbf{E}[\mathsf{COV}_{\max}(G)] \leqslant 2e^3 \cdot \max_{u,v \in V} \mathsf{H}(u, v) \ln n + n$

We shall consider the cover time when k random walks start at the same vertex, where $1 \leqslant k \leqslant n$. To this end, we study $\mathbf{E}\left[\mathsf{COV}_u^k(G)\right]$, defined as the expected time for k random walks starting from u to cover all n vertices of G. Set $\mathbf{E}\left[\mathsf{COV}_{\max}^k(G)\right] = \max_{u \in V} \mathbf{E}\left[\mathsf{COV}_u^k(G)\right]$. Clearly, $\mathbf{E}\left[\mathsf{COV}_{\max}^k(G)\right]$ decreases in k. Hence several of our lower bounds stated for $\mathbf{E}[\mathsf{COV}_{\max}^n(G)]$ directly imply the same bound on $\mathbf{E}\left[\mathsf{COV}_{\max}^k(G)\right]$ with $k \leqslant n$. Sometimes, we will also consider $\mathbf{E}\left[\mathsf{COV}_{\pi}^k(G)\right]$. In this case, each starting point of the k random walks is chosen independently from the stationary distribution π. We recall:

Theorem 2.3 ([7, Theorem 1]). *Let G be any graph with m edges. Then we have for any $1 \leqslant k \leqslant n$, $\mathbf{E}\left[\mathsf{COV}_{\pi}^k(G)\right] = \mathcal{O}\left(\frac{m^2}{k^2} \cdot \log^3 n\right)$.*

We continue with an auxiliary lemma.

Lemma 2.4. *Let X_1 and X_2 be two random variables taking values in a finite set S. Assume that there is a number $0 < C < 1$ such that for every $s \in S$, $\mathbf{Pr}[X_1 = s] \geqslant C\,\mathbf{Pr}[X_2 = s]$. Then there exists a coupling $\widehat{X} = (\widehat{X}_1, \widehat{X}_2)$ of X_1 and X_2 such that $\mathbf{Pr}\left[\widehat{X}_1 = \widehat{X}_2\right] \geqslant C$.*

3 Lower Bounds on the Speed-Up

A natural relation that has been also used by Alon et al. [3] is the following.

Lemma 3.1. *For any* $1 \leqslant k \leqslant n$, $\mathbf{E}\left[\mathsf{COV}^k_{\max}(G)\right] \leqslant \mathbf{E}\left[\mathsf{COV}^k_{\pi}(G)\right] + \mathsf{MIX}_{n^{-3}}(G)$.

We prove an extension where the threshold for the mixing time is much smaller. This apparently small difference will be crucial to obtain tight bounds for hypercubes (Section 5).

Lemma 3.2. *For any* $1 \leqslant k \leqslant n$, $\mathbf{E}\left[\mathsf{COV}^k_{\max}(G)\right] \leqslant 16\,\mathbf{E}\left[\mathsf{COV}^{k/2}_{\pi}(G)\right] + 4\,\mathsf{MIX}_{1/2}(G)$.

Proof Sketch. The basic idea is as follows. Let $X^1, X^2, \ldots, X^k$ be k random walks starting from the same vertex u. Moreover, let $Y^1, Y^2, \ldots, Y^k$ be k random walks, all starting from independent samples of π. Our goal is to relate the set of covered vertices by $X^1, \ldots, X^k$ to the covered ones by $Y^1, \ldots, Y^{k/2}$ at the cost of an additional $\mathsf{MIX}_{1/2}(G)$-term. In order to do so, we will prove that at least half of the random walks among $X^1, \ldots, X^k$ are located on a uniformly chosen vertex after $\mathsf{MIX}_{1/2}(G)$ steps.

Theorem 3.3. *For every graph* G *and* k *with* $1 \leqslant k \leqslant n$,

$$\mathbf{E}\left[\mathsf{COV}^k_{\pi}(G)\right] = \mathcal{O}\left(\frac{\log n \cdot (\max_{u,v} \mathsf{H}(u,v) + \mathsf{MIX}_{1/2}(G))}{k} + \mathsf{MIX}_{1/2}(G)\right).$$

Before we outline the proof of Theorem 3.3, let us point out that the result also improves over Theorem 2.3 for a wide range of k, provided that $\mathsf{MIX}_{1/2}(G)$ and $\max_{u,v} \mathsf{H}(u,v)$ are not too large.

Proof Sketch. We devise a coupling of a single random walk X to k random walks, each of which starts according to π. We shall divide the single random walk X into consecutive sections of length $\mathsf{MIX}_{1/2}(G)$. We then argue that a random walk starting from the stationary distribution has (almost) the same chance of visiting a vertex within $\mathsf{MIX}_{1/2}(G)$ steps as the single random walk has in one fixed section. This implies that the probability that the k random walks visit this vertex is (nearly) the probability that X visits the same vertex in one of the even sections. Here it is crucial to consider only the even (or odd) sections, so that the random walk X is located on a vertex according to π each time a new section begins.

Combining this result with Lemma 3.2 we get immediately:

Theorem 3.4. *For any graph* G *and any* k *with* $1 \leqslant k \leqslant n$,

$$\mathbf{E}\left[\mathsf{COV}^k_{\max}(G)\right] = \mathcal{O}\left(\frac{\log n \cdot (\max_{u,v} \mathsf{H}(u,v) + \mathsf{MIX}_{1/2}(G))}{k} + \mathsf{MIX}_{1/2}(G)\right).$$

Let us state a simpler, slightly weaker bound on the speed-up that follows directly from Theorem 3.4:

Corollary 3.5. *Let G be a graph that satisfies $\mathsf{MIX}_{1/2}(G) = \mathcal{O}(\max_{u,v} \mathsf{H}(u,v))$ and $\mathbf{E}\left[\,\mathsf{COV}_{\max}(G)\,\right] = \Theta(\max_{u,v} \mathsf{H}(u,v) \log n)$. Then for any $1 \leqslant k \leqslant n$,*

$$S^k(G) = \Omega\left(\frac{k}{1 + \frac{\mathsf{MIX}_{1/2}(G)}{\mathbf{E}[\,\mathsf{COV}_{\max}(G)\,]} \cdot k}\right).$$

Hence as long as $k = \mathcal{O}\left(\frac{\mathbf{E}[\,\mathsf{COV}_{\max}(G)\,]}{\mathsf{MIX}_{1/2}(G)}\right)$, Corollary 3.5 yields a speed-up of $\Omega(k)$. Note that all graphs (except cycles and 2-dim. torus) in Figure 1 satisfy the conditions of Corollary 3.5.

4 Upper Bounds on the Speed-Up

Alon et al. [3] gave a graph G and vertex u such that $\frac{\mathbf{E}[\,\mathsf{COV}_u(G)\,]}{\mathbf{E}\left[\,\mathsf{COV}_u^k(G)\,\right]} = \Omega(2^k)$ for $k = \Theta(\log n)$, so the speed-up is exponential in k. However, their example does not work when u is replaced by a worst-case starting vertex. This lead to their conjecture that the speed-up is always polynomial in k, if the starting vertex is worst-case. More precisely, they conjectured that for any graph and any $1 \leqslant k \leqslant n$, $S^k = \mathcal{O}(k)$.

We shall prove that $S^k = \mathcal{O}(k \log n)$ for any graph and k, matching the conjecture up to a factor of $\mathcal{O}(\log n)$. This also shows that while for an arbitary starting vertex an *exponential* speed-up is possible, the speed-up is always *polynomial*, if the starting vertex is worst-case.

Proposition 4.1. *For any graph G and any $1 \leqslant k \leqslant n$, $S^k = \mathcal{O}(k \log n)$.*

Proof. Fix a vertex w. Choose a vertex u such that

$$\mathbf{Pr}\left[\text{walk of length } \mathbf{E}\left[\,\mathsf{COV}_u^k(G)\,\right] \text{ starting at } u \text{ visits } v\right]$$

is minimized. We claim by way of contradiction that

$$\mathbf{Pr}\left[\text{walk of length } 2\,\mathbf{E}\left[\,\mathsf{COV}_u^k(G)\,\right] \text{ starting at } u \text{ visits } v\right] \geqslant \frac{1}{4k}.$$

Assuming the converse, the probability that all k random walks starting at u do not cover w would be at least $\prod_{i=1}^{k}\left(1 - \frac{1}{4k}\right) \geqslant 1 - \sum_{i=1}^{k} \frac{1}{4k} = \frac{3}{4}$, which in turn would imply $\mathbf{E}\left[\,\mathsf{COV}_u^k(G)\,\right] \geqslant \frac{3}{2}\,\mathbf{E}\left[\,\mathsf{COV}_u^k(G)\,\right]$, a contradiction.

Consider now a single random walk of length $16\,\mathbf{E}\left[\,\mathsf{COV}_u^k(G)\,\right] k \ln n$. Then,

$$\mathbf{Pr}\left[\text{walk of length } 16\,\mathbf{E}\left[\,\mathsf{COV}_u^k(G)\,\right] k \ln n \text{ starting at } u \text{ visits } v\right]$$
$$\geqslant 1 - \left(1 - \frac{1}{4k}\right)^{8k \ln n} \geqslant 1 - \frac{1}{n^2}.$$

Taking the union bound over all n vertices yields the claim.

4.1 Special Upper Bounds

Additionally, we shall derive three more specific lower bounds on $\mathbf{E}\left[\mathsf{COV}^k_{\max}(G)\right]$. As a consequence, they are most useful to upper bound the speed-up when the graph satisfies $\mathbf{E}\left[\mathsf{COV}_{\max}(G)\right] = \mathcal{O}(n \log n)$ (which is the case for most interesting graphs (cmp. Figure 1)).

We start by deriving a lower bound of $\Omega((n/k) \log n)$ for not too small k by using a relatively simple coupon-collecting argument. After that we present a lower bound of $\Omega((n/k) \log n)$ for not too large k, requiring that the mixing time is sublinear. Combining these bounds, we obtain that $\mathbf{E}\left[\mathsf{COV}^k_{\max}(G)\right] = \Omega((n/k) \log n)$ for any $1 \leqslant k \leqslant n$ (if the mixing time is sublinear).

We start with a bound based on a coupon-collecting argument. We view each random walk as an independent string of n letters (corresponding to n vertices). Then we bound the probability that all letters occur in a sample of k random strings.

Theorem 4.2. *Let k be an arbitrary integer satisfying $k \geqslant n^{\varepsilon}$ for an arbitrary constant $0 < \varepsilon < 1$. Then,* $\mathbf{E}\left[\mathsf{COV}^k_{\max}(G)\right] = \Omega\left(\frac{n}{k} \log n\right)$.

For $k < n^{\varepsilon}$, we devise a lower bound on $\mathbf{E}\left[\mathsf{COV}^k_{\max}(G)\right]$ that requires a sublinear mixing time. We use the following result from Broder and Karlin [6].

Lemma 4.3 ([6, Lemma 12]). *Consider a single random walk $X_1, X_2, \ldots$ with a symmetric transition matrix $\mathbf{P}$. Let T_s be the first time when s different vertices are covered. Then for any $m \in \mathbb{N}$,*

$$\mathbf{E}\left[T_{\lfloor (m+1)n/(m+2) \rfloor} - T_{\lfloor (m)n/(m+1) \rfloor}\right] \geqslant \frac{1}{2}\frac{n}{m+2} - \mathcal{O}(\mathsf{MIX}_{n^{-1}}(G) \cdot m).$$

Using the lemma above, we can show the following corollary:

Corollary 4.4. *Let $X = (X_1, X_2, \ldots)$ be a random walk on regular graph. Let T_s be the first time that s different vertices are covered. Let $1 \leqslant m \leqslant n$ be any positive integer. Define $\gamma_m := \frac{1}{2}\frac{n}{m+2} - \mathcal{O}(\mathsf{MIX}_{n^{-1}}(G) \cdot m)$. Then,*

$$\mathbf{Pr}\left[T_{\lfloor (m+1)n/(m+2) \rfloor} - T_{\lfloor (m)n/(m+1) \rfloor} \geqslant \frac{1}{4}\gamma_m\right] \geqslant \frac{1}{16}.$$

Theorem 4.5. *Assume that $\mathsf{MIX}_{n^{-1}}(G) = \mathcal{O}(n^{1-\varepsilon})$ for a constant $\varepsilon > 0$. Then for any regular graph G and $k \leqslant \sqrt[4]{n/\mathsf{MIX}_{n^{-1}}(G)}$,* $\mathbf{E}\left[\mathsf{COV}^k_{\max}\right] = \Omega\left(\frac{n}{k} \log n\right)$.

Proof Sketch. As in [6] our goal is to divide the random walks viewing one after another into a certain number of epochs, where a new epoch starts if a certain number of new vertices has been covered. Then we can bound the remaining time in each epoch by Corollary 4.4. The technical difficulty arises when the lower bound by Corollary 4.4 is larger than the remaining time of the walk. In this case we assume (quite pessimistically) that the random walk has finished one epoch, but this suffices, since k is rather small.

Combining Theorem 4.2 and Theorem 4.5 we obtain immediately:

Corollary 4.6. *For any regular graph G with* $\mathsf{MIX}_{n^{-1}}(G) = \mathcal{O}(n^{1-\varepsilon})$ *for a constant* $\varepsilon > 0$ *and any* $1 \leqslant k \leqslant n$*, we have* $\mathbf{E}\left[\mathsf{COV}^k_{\max}(G)\right] = \Omega\left(\frac{n}{k}\log n\right)$.

Turning back to the original question on upper bounding S^k we get:

Corollary 4.7. *For any regular graph G that satisfies* $\mathsf{MIX}_{n^{-1}}(G) = \mathcal{O}(n^{1-\varepsilon})$ *and* $\mathbf{E}\left[\mathsf{COV}_{\max}(G)\right] = \Theta(n\log n)$*, we have for any* $1 \leqslant k \leqslant n$*,* $S^k = \mathcal{O}(k)$*.*

This establishes the conjecture of Alon et al. [3] for a large class of graphs including most graphs of Figure 1.

Obviously, $\mathrm{diam}(G)$ is a lower bound on $\mathbf{E}\left[\mathsf{COV}^k_{\max}(G)\right]$ for each k. Using a result of [8], we can prove the following improvement (if $\mathrm{diam}(G) \geqslant \log n$):

Theorem 4.8. *For any graph G with diameter* $\mathrm{diam}(G)$*,* $\mathbf{E}\left[\mathsf{COV}^n_{\max}(G)\right] = \Omega\left(\frac{\mathrm{diam}(G)^2}{\log n}\right)$.

5 Applications to Concrete Graphs

Expanders and Random Graphs. There are several (mostly equivalent) definitions of expanders. Here, we call a regular graph an expander if $\mathsf{MIX}_{n^{-1}}(G) = \mathcal{O}(\log n)$ (this is a more general definition than [3], where additionally the degree has to be constant). It is also a well-known fact that any expander graph satisfies $\max_{u,v} \mathsf{H}(u,v) = \mathcal{O}(n)$ (cf. [6, 17]). Hence Corollary 3.5 implies a speed-up of $\Omega(k)$ for any $1 \leqslant k \leqslant n$. Moreover, Corollary 4.6 establishes tightness.

For any given $(1+\varepsilon)\log(n)/n < p < 1, \varepsilon > 0$, an Erdős-Rényi random graph is constructed by placing an edge between each pair of vertices independently with probability p. Similar to regular expanders, we can prove the same result for random graphs leading to the following corollary.

Corollary 5.1. *For any regular expander graph and almost all Erdős-Rényi random graphs with* $p \geqslant (1+\varepsilon)\log(n)/n$*, we have for any* $1 \leqslant k \leqslant n$ *that* $S^k = \Theta(k)$*.*

Hypercubes. Let us consider the speed-up on the $\log n$-dimensional hypercube H_n with n vertices. It is known that $\max_{u,v\in V} \mathsf{H}(u,v) = \mathcal{O}(n)$ (cf. [17]) which readily implies that $\mathbf{E}\left[\mathsf{COV}_{\max}(H_n)\right] = \Theta(n\log n)$.

Lemma 5.2. *For the hypercube H_n,* $\mathsf{MIX}_{1/2}(G) = \mathcal{O}(\log n \log\log n)$*.*

We remark that $\overline{\mathsf{MIX}}_{n^{-1}}(H_n) = \Omega(\log^2 n)$, so it is crucial to use $\mathsf{MIX}_{1/2}(H_n)$. Hence, as long as $k \leqslant C_1 n/(\log n \log\log n)$ for a sufficiently small constant C_1, Corollary 3.5 and Corollary 4.7 imply that the speed up is $\Theta(k)$. (We point out that using the techniques of [10], a more precise bound on the speed could be obtained). Let us now consider the case when k is large.

Theorem 5.3. *For the hypercube H_n,* $\mathbf{E}\left[\mathsf{COV}^n_{\max}(H_n)\right] = \Omega(\log n \log\log n)$*.*

Hence the speed-up on hypercubes undergoes a surprisingly sharp transition: it equals k if $k = \mathcal{O}(n/\log\log n)$, but as soon as $k = \Theta(n/\log\log n)$ the speed-up does not increase further.

Cayley Graphs with Small Degree (including Torus Graphs). Let us now consider torus graphs. For cycles, Alon et al. [3, Theorem 6] prove that $S^k = \Theta(\log k)$ for any $1 \leqslant k \leqslant n$. For the two-dimensional torus graph, they proved [3, Theorem 4 & 8, Corollary 25] that $S^k(G) = \Omega(k)$ for $k \leqslant \log n$, but $S^k(G) = \mathcal{O}(\log^2 n \log k)$ for any $1 \leqslant k \leqslant n$. Therefore, we only have to consider the d-dimensional torus with $d \geqslant 3$ in the following. In fact, we shall look at Cayley graphs more generally. Recall that an undirected Cayley graph is a graph whose vertex set is equal to the elements of a finite group and the edge set is given by a set of self-inverse group generators (cf. [19]). We recall the following lemma.

Lemma 5.4 ([19]). *For any Cayley graph G, $\mathsf{MIX}_{1/2}(G) = \mathcal{O}(\Delta \operatorname{diam}(G)^2 \log n)$.*

Now applying Corollary 3.5 and Theorem 4.8 we obtain the following.

Theorem 5.5. *Let G be a Δ-regular Cayley graph such that $\mathbf{E}[\,\mathsf{COV}_{\max}(G)\,] = \Theta(\max_{u,v} \mathsf{H}(u,v) \log n)$. Then, for any $k \leqslant \frac{\mathbf{E}[\,\mathsf{COV}_{\max}(G)\,]}{\Delta \operatorname{diam}(G)^2 \log n}$, $S^k(G) = \Omega(k)$. Moreover for any $1 \leqslant k \leqslant n$, $S^k(G) = \mathcal{O}\left(\frac{\mathbf{E}[\,\mathsf{COV}_{\max}(G)\,]}{\operatorname{diam}(G)^2} \log n\right)$.*

Hence for any Cayley graph with small degree Δ, there is a sharp threshold point near $\frac{\mathbf{E}[\,\mathsf{COV}_{\max}(G)\,]}{\operatorname{diam}(G)^2}$. For d-dimensional torus with $d > 2$ we can prove a slightly stronger result, since it is known that $\max_{u,v} \mathsf{H}(u,v) = \Theta(n)$ and $\mathsf{MIX}_{1/2}(G) = \Theta(\operatorname{diam}(G)^2) = \Theta(n^{2/d})$ (cf. [3, 17]). Applying Corollary 3.5 for the lower bound, and, Theorem 4.8 and Corollary 4.7 for the upper bound gives:

Corollary 1. *Let G be a d-dimensional torus with $d > 2$. Then for any $1 \leqslant k \leqslant n^{1-2/d} \log n$, $S^k(G) = \Omega(k)$. Moreover for any $1 \leqslant k \leqslant n$, $S^k(G) = \mathcal{O}\left(\min\{k, n^{1-2/d} \log^2 n\}\right)$.*

6 Conclusions

We presented several lower and upper bounds on the speed-up defined as the ratio between the cover time of one and the cover time of k random walks. On a concrete level, our results fill several gaps left open in the previous work of Alon et al. [3] (cmp. Figure 1). From a higher perspective, our findings also provide an answer to the question raised by [3] about a good characterization of a best-possible speed-up. For a large class of graphs, a speed-up of $\Omega(k)$ is possible up to a certain threshold (roughly $n \log n$ divided by the mixing time), while above the threshold the speed-up does not increase further.

Bibliography

[1] Adler, M., Halperin, E., Karp, R., Vazirani, V.: A Stochastic Process on the Hypercube with Applications to Peer-to-Peer Networks. In: 35th Annual ACM Symposium on Theory of Computing (STOC 2003), pp. 575–584 (2003)

[2] Aleliunas, R., Karp, R., Lipton, R., Lovász, L., Rackoff, C.: Random Walks, Universal Traversal Sequences, and the Complexity of Maze Problems. In: 20th Annual IEEE Symposium on Foundations of Computer Science (FOCS 1979), pp. 218–223 (1979)

[3] Alon, N., Avin, C., Koucky, M., Kozma, G., Lotker, Z., Tuttle, M.: Many Random Walks are faster than one. In: 20th Annual ACM Symposium on Parallel Algorithms and Architectures (SPAA 2008), pp. 119–128 (2008)
[4] Avin, C., Koucky, M., Lotker, Z.: How to Explore a Fast-Changing World (Cover Time of a Simple Random Walk on Evolving Graphs). In: Aceto, L., Damgård, I., Goldberg, L.A., Halldórsson, M.M., Ingólfsdóttir, A., Walukiewicz, I. (eds.) ICALP 2008, Part I. LNCS, vol. 5125, pp. 121–132. Springer, Heidelberg (2008)
[5] Avin, C., Krishnamachari, B.: The power of choice in random walks: an empirical study. In: 9th International Symposium on Modeling Analysis and Simulation of Wireless and Mobile Systems, pp. 219–228 (2006)
[6] Broder, A., Karlin, A.: Bounds on the cover time. In: 29th Annual IEEE Symposium on Foundations of Computer Science (FOCS 1988), pp. 479–487 (1988)
[7] Broder, A., Karlin, A., Raghavan, P., Upfal, E.: Trading space for time in undirected s–t-connectivity. SIAM Journal on Computing 23(2), 324–334 (1994)
[8] Carne, T.: A Transmutation Formula for Markov Chains. Bulletin des Sciences Mathematiques 109, 399–405 (1985)
[9] Chandra, A., Raghavan, P., Ruzzo, W., Smolensky, R., Tiwari, P.: The Electrical Resistance of a Graph Captures its Commute and Cover Times. Computational Complexity 6(4), 312–340 (1997)
[10] Cooper, C., Frieze, A., Radzik, T.: Multiple random walks and interacting particle systems. In: 36th International Colloquium on Automata, Languages, and Programming (ICALP 2009) (2009) (to appear)
[11] Dimitrov, N.B., Plaxton, C.G.: Optimal Cover Time for a Graph-Based Coupon Collector Process. In: Caires, L., Italiano, G.F., Monteiro, L., Palamidessi, C., Yung, M. (eds.) ICALP 2005, vol. 3580, pp. 702–716. Springer, Heidelberg (2005)
[12] Dolev, S., Schiller, E., Welch, J.L.: Random walk for self-stabilizing group communication in ad hoc networks. IEEE Transactions on Mobile Computing 5(7), 893–905 (2006)
[13] Gkantsidis, C., Mihail, M., Saberi, A.: Random walks in peer-to-peer networks. Performance Evaluation 63(3), 241–263 (2006)
[14] Ikeda, S., Kubo, I., Okumoto, N., Yamashita, M.: Impact of Local Topological Information on Random Walks on Finite Graphs. In: Baeten, J.C.M., Lenstra, J.K., Parrow, J., Woeginger, G.J. (eds.) ICALP 2003. LNCS, vol. 2719, pp. 1054–1067. Springer, Heidelberg (2003)
[15] Kale, S., Seshadhri, C.: An expansion tester for bounded degree graphs. In: Aceto, L., Damgård, I., Goldberg, L.A., Halldórsson, M.M., Ingólfsdóttir, A., Walukiewicz, I. (eds.) ICALP 2008, Part I. LNCS, vol. 5125, pp. 527–538. Springer, Heidelberg (2008)
[16] Kempe, D., Dobra, A., Gehrke, J.: Gossip-based computation of aggregate information. In: 44th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2003), pp. 482–491 (2003)
[17] Lovász, L.: Random walks on graphs: A survey. Combinatorics, Paul Erdös is Eighty 2, 1–46 (1993)
[18] Motwani, R., Raghavan, P.: Randomized Algorithms, 7th edn. Cambridge University Press, Cambridge (1995)
[19] Pak, I.: Mixing time and long paths in graphs. In: 13th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA 2002), pp. 321–328 (2002)
[20] Sinclair, A.: Algorithms for Random Generation and Counting. Birkhäuser, Basel (1993)