# How Well Do Random Walks Parallelize?

Klim Efremenko [*] Omer Reingold [†]

August 24, 2009

**Abstract**

A random walk on a graph is a process that explores the graph in a random way: at each step the walk is at a vertex of the graph, and at each step it moves to a uniformly selected neighbor of this vertex. Random walks are extremely useful in computer science and in other fields. A very natural problem that was recently raised by Alon, Avin, Koucky, Kozma, Lotker, and Tuttle (though it was implicit in several previous papers) is to analyze the behavior of $k$ independent walks in comparison with the behavior of a single walk. In particular, Alon et al. showed that in various settings (e.g., for expander graphs), $k$ random walks cover the graph (i.e., visit all its nodes), $\Omega(k)$-times faster (in expectation) than a single walk. In other words, in such cases $k$ random walks efficiently "parallelize" a single random walk. Alon et al. also demonstrated that, depending on the specific setting, this "speedup" can vary from logarithmic to exponential in $k$.

In this paper we initiate a more systematic study of multiple random walks. We give lower and upper bounds both on the cover time *and on the hitting time* (the time it takes to hit one specific node) of multiple random walks. Our study revolves over three alternatives for the starting vertices of the random walks: the worst starting vertices (those who maximize the hitting/cover time), the best starting vertices, and starting vertices selected from the stationary distribution. Among our results, we show that the speedup when starting the walks at the worst vertices cannot be too large - the hitting time cannot improve by more than an $O(k)$ factor and the cover time cannot improve by more than $\min\{k \log n, k^2\}$ (where $n$ is the number of vertices). These results should be contrasted with the fact that there was no previously known upper-bound on the speedup and that the speedup can even be *exponential* in $k$ for random starting vertices. We further show that for $k$ that is not too large (as a function of various parameters of the graph), the speedup in cover time is $O(k)$ *even for walks that start from the best vertices* (those that minimize the cover time). As a rather surprising corollary of our theorems, we obtain a new bound which relates the cover time $C$ and the mixing time $\mathrm{mix}$ of a graph. Specifically, we show that $C = O(m\sqrt{\mathrm{mix}}\log^2 n)$ (where $m$ is the number of edges).

[*]Department of Computer Science and Applied Mathematics, The Weizmann Institute of Science, Rehovot, 76100 Israel. E-mail: `klimefrem@gmail.com`.

[†]Department of Computer Science and Applied Mathematics, The Weizmann Institute of Science, Rehovot, 76100 Israel. E-mail: `omer.reingold@weizmann.ac.il`. Research supported by grant 1300/05 from the Israel Science Foundation.

# 1 Introduction

A random walk on a graph is a process of exploring the graph in a random way. A simple random walk starts at some node of a graph and at each step moves to a random neighbor. Random walks are fundamental in computer science. They are the basis of MCMC (Markov-Chain Monte-Carlo) algorithms, and have additional important applications such as randomness-efficient sampling (via random walks on expanders) [AKS87], and space-efficient graph connectivity algorithms [AKL$^+$79]. Random walks became a common notion in many fields, such as computational physics, computational biology, economics, electrical engineering, social networks, and machine learning.

Assume that we have some network (e.g. a communication or a social network), and some node $u$ sends a message. Assume that at each step this message is sent to a random neighbor of the last recipient. The message will travel through the network as a random walk on a graph. The expected time until the message will arrive to some other node $v$ is called the hitting time $h(u, v)$. The expected time until the message will visit all the nodes is called the cover time $C_u$. The hitting time and the cover time of a random walk are thoroughly studied parameters (see surveys [AF99, LWP, Lov96]).

In this paper we consider the following natural question: What happens if we take multiple random walks instead of a single walk? Assume that instead of one copy, $k$ copies of the same message were sent. How long would it take for one of these copies to reach some node $v$? How long would it take until each node receives at least one of the $k$ copies? What are the speedups in the hitting and cover times of multiple walks compared with a single walk?

Multiple random walks were studied in a series of papers [BKRU89, Fei97, BF93] on time-space tradeoffs for solving undirected $s$-$t$ connectivity. These papers considered upper bounds for the cover time of multiple random walks, each paper giving a different answer for different distributions of the starting vertices of the random walks. In randomized parallel algorithms, multiple random walks are a very natural way of exploring a graph since they can be easily distributed between different processes. For example, multiple random walks were used in [HZ96, KNP99] for designing efficient parallel algorithms for finding the connected components of an undirected graph.

Multiple random walks were suggested as a topic of independent interest by Alon, Avin, Koucky, Kozma, Lotker, and Tuttle [AAK$^+$07]. Alon et al. [AAK$^+$07] studied lower bounds on the relation between the cover time of a simple random walk and of multiple random walks when the walks start from the same node. The paper proves that if the number of random walks $k$ is small enough (i.e., asymptotically less than $\frac{C}{h_{\max}}$, where $C$ and $h_{\max}$ are the maximal cover time and hitting time respectively) then the relation between the cover time of a single random walk and of multiple random walks is at least $k - o(k)$. In such a case, we can argue that multiple random walks "parallelize" a single walk efficiently (as they don't increase the total amount of work by much). [AAK$^+$07] also showed that there are graphs with logarithmic speedup (e.g., the cycle), and there are graphs with an exponential speedup (e.g., the so called barbell graph; we will shortly discuss a related example). [AAK$^+$07] leaves open the question of upper bounds for the speedup.

The goal of this paper is to systematically study multiple random walks. In addition to the cover time of multiple random walks we will also discuss the hitting time, proving both lower and upper bounds on the speedup. We will extend the discussion to the case where not all the walks start from the same node.

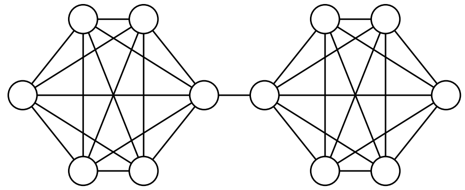

Figure 1: Two cliques graph - how the speedup changes according to the starting vertices

Before getting into the details of our results, let us consider an example which illustrates how multiple random walks behave differently according to the choice of their starting vertices. Consider a graph $G$ which is composed of two cliques of size $n$ connected by a single edge (see Figure1). While the cover time of a single random walk will not depend on the starting vertex and is $\Theta(n^2)$, the cover time of multiple random walks will be very different for different starting vertices of the random walks. When the walks start from the worst vertices (all walks start from the same clique) the cover time is $\Theta(\frac{n^2}{k})$. Even for $k = 2$, when the random walks start from the best vertices (one walk starts at one clique and the other from another clique) the cover time is $\Theta(n \log n)$. When the starting vertices of $k$ random walks are drawn independently from the stationary distribution, then the probability that all starting vertices will fall into the same clique is $2^{-k}$. Therefore, for $k \leq \log n - \log \log n$, the cover time in this case is $\Theta(2^{-k} n^2)$. When considering the hitting times, we get the same behavior for the worst starting vertices and for randomly-chosen starting vertices. The case of the best starting vertices is uninteresting when discussing the hitting time as the hitting time in such a case is zero (even for a single walk).

As we can see from the aforementioned example, both the cover and the hitting times heavily depend on the starting vertices. Therefore, we study these three scenarios separately:

1. The case when the random walks start from the nodes which maximize the cover/hitting time (worst starting vertices).
2. The case when the random walks start from the nodes which minimize the cover time (best starting vertices).
3. The case when the starting vertices are drawn independently according to the stationary distribution (random starting vertices).

## Our contribution

In this paper we systematically study multiple random walks and their speedup both in terms of the cover time and in terms of the hitting time. We give various lower and upper bounds for different ways of choosing the starting vertices. Our main bounds on the speedup of multiple random walks are summarized in Table 1.

**Upper bounds on the speedup:** [AAK+07] left open the question of upper bounding the speedup of multiple random walks. In this work we show that the answer depends on how the starting vertices

| | Worst case | Average Case | Best Case |
|---|---|---|---|
| Hitting time<br>Upper bounds | $O(k)$<br>*for any* $k$, Theorem 3.3 | $k + o(k)$<br>*for* $k \log n = o(\frac{h_{\max}}{\mathrm{mix}})$<br>Theorem 4.10 | Not applicable |
| Hitting time<br>Lower bounds | $\Omega(k)$<br>*for* $k \log n = O(\frac{h_{\max}}{\mathrm{mix}})$<br>Theorem 3.10 | $\Omega(k)$<br>*for any* $k$, 3.9<br>Theorems 3.8, & 3.9 | Not applicable |
| Cover time<br>Upper bounds | $O(\min\{k^2, k \log n\})$<br>Theorems 4.1 & 4.2 | $k + o(k)$<br>*for* $k \log k = o(\frac{C}{\mathrm{mix}})$<br>Theorem 4.8 | $k + o(k)$<br>*for* $k = o(\frac{C}{h_{\max}})$<br>Theorem 4.4 |
| Cover time<br>Lower Bounds | $\Omega(\frac{k}{\log n})$ for $k \log n = O(\frac{h_{\max}}{\mathrm{mix}})$<br>Theorem 4.3<br>$k - o(k)$ for $k = o(\frac{C}{h_{\max}})$<br>Theorem 5 in [AAK+07] | $\Longrightarrow$ | $\Longrightarrow$ |

Table 1: Summary of main bounds on the speedup. Notation: $n$ - number of vertices; $k$ - the number of walks; $C$ - maximal cover time; $h_{\max}$ - maximal hitting time; $\mathrm{mix}$ - mixing time

are selected. In Theorem 3.3 we show that for the worst starting vertices, the speedup on *hitting time* is at most $O(k)$. In Section 4, we use this theorem to show that the speedup on the *cover time* is at most $O(\min(k^2, k \log n))$. As we can see from the example above, the speedup for the best or even for random starting vertices may be very large (e.g., exponential in $k$). Still, we are able to show in Section 4 that even in these cases, if the number of walks is small enough then the speedup will be at most $k + o(k)$. In Theorem 4.4 (see also Corollary 4.6 ) we show that for $k \ll \frac{C}{h_{\max}}$ the speedup for the best starting vertices is at most $k + o(k)$. This result is interesting for graphs with a large gap between the cover time and the hitting time. For random starting vertices, Theorem4.8 (see also Corollary 4.11) shows that if $k \log k \ll \frac{C}{\mathrm{mix}}$, then the speedup is at most $k + o(k)$. The mixing time, $\mathrm{mix}$, of a graph is the number of steps a random walk has to make until its position is distributed almost according to the stationary distribution.

**Lower bounds on the speedup:** In Theorem 3.8 we show that the speedup for the hitting times is at least $k$ when all the starting vertices are drawn from the stationary distribution. This theorem also allows us to prove lower bounds for the case of worst starting vertices for graphs with small mixing time. Using this theorem we prove in Theorem 3.10 that when the number of walks is less than $\tilde{O}(\frac{h_{\max}}{\mathrm{mix}})$ the speedup for the hitting times is at least $\Omega(k)$. We get similar results for the cover time (Theorem 4.3). Namely, we show that the speedup for the cover time is at least $\Omega(\frac{k}{\log n})$, when $k$ is less than $\tilde{O}(\frac{h_{\max}}{\mathrm{mix}})$. This result improves the lower bound of $\Omega(\frac{k}{\log n \cdot \mathrm{mix}})$ from [AAK+07].

**A new relation between the cover time and the mixing time:** Finally, our study of multiple random walks gives a rather surprising implication on the study of a single random walk. Our results, together with the results of [BKRU89] about multiple random walks, imply a new relation between the cover time and the mixing time of a graph. Specifically, we prove that $C = O(m\sqrt{\mathrm{mix}} \log^2 n)$. The best previous result we are aware of is due to Broder and Karlin [BK88]. In [BK88] it was proven that $C = O(\frac{m \log n}{1-\lambda(G)})$, where $\lambda(G)$ is the second eigenvalue of the normalized adjacency matrix. A known relation between $\lambda(G)$ and $\mathrm{mix}$ is that $\Omega(\frac{1}{1-\lambda(G)}) \leq \mathrm{mix} \leq O(\frac{\log n}{1-\lambda(G)})$ (cf. [Sin92], Proposition 1). Therefore a corollary of [BK88] is that $C = O(\mathrm{mix} m \log n)$. Our result improves this bound whenever $\mathrm{mix} = \omega(\log^2 n)$.

Our new relation also has an application in electrical engineering. View a graph $G$ as an electrical network with unit resistors as edges. Let $R_{st}$ be the effective resistance between nodes $s$ and $t$. Then it was shown in [CRRS89] that for any two nodes $s$ and $t$ it holds that $mR_{st} \leq C$. Therefore, together with our result it implies that $R_{st} \leq \sqrt{\mathrm{mix}} \log^2 n$. The best previous upper bound on the electrical resistance in terms of the mixing time was also obtained by Chandra et al. [CRRS89] and was $R_{st} \leq \frac{2}{1-\lambda(G)} = O(\mathrm{mix})$.

**Related Work** Independently of our work, Elsässer and Sauerwald [ES09] recently studied multiple random walks. Their most related results are upper bounds and lower bounds on the speed-up of cover time for worst case starting points. In fact, [ES09] gives an upper bound of $O(k \log n)$ on the speed-up of any graph (similarly to our Theorem 4.1) and a lower bound of $\Omega(\frac{k}{\log n})$ under some conditions on mixing time (similarly to our Theorem 4.3). Under some mild conditions, they are also able to prove an upper bound of $O(k)$. Another recent work on multiple random walks is due to [CCR09]. This work studies multiple random walks in random graphs, and among other result show that for random $d$-regular graph the speed-up is $O(k)$.

## 2 Preliminaries

Let $G(V, E)$ be a graph with $V$ the set of nodes and $E$ the set of edges. In this paper we only consider graphs that are strongly connected i.e. for any two nodes $u, v \in V$ there exists a directed path from $u$ to $v$ and a path from $v$ to $u$. A cycle in the graph $G$ is a sequence of nodes $u_1, u_2 \ldots u_n$ such that $u_n = u_1$ and $(u_{i-1}, u_i) \in E$. The cycle length is $n$.

**Definition 2.1.** *A graph $G(V, E)$ is called aperiodic if the $\gcd$ of the lengths of all the cycles in the graph is $1$.*

In particular, an undirected graph is aperiodic iff it is non-bipartite. In this paper we will consider only aperiodic graphs.

For any node $v \in V$ we will denote the set of neighbors of $v$ by $N(v) = \{u \in V \mid (v, u) \in E\}$, and let $\delta(v) = |N(v)|$ be the degree of $v$. Let $X_u(t)$ be a simple random walk on $G$ starting at $u$. This is a Markov chain with $X_u(0) = u$ and

$$\Pr[X_u(t) = w \mid X_u(t-1) = v] = \frac{\text{\# edges } (v, w) \in E}{\delta(v)}$$

In particular, if $G$ does not contain parallel edges then $\Pr[X_u(t) = w \mid X_u(t-1) = v] = \frac{1}{\delta(v)}$ if $(v, w) \in E$ and zero otherwise.

For any distribution $\nu$ on $V$, denote by $X_\nu(t)$ the distribution of $X_u(t)$ where $u$ is drawn from the distribution $\nu$. We denote by $M$ the matrix of transition probabilities i.e. $M(i, j) = \Pr[X_u(t) = j \mid X_u(t-1) = i]$. For any undirected graphs it holds that $\delta(i)M(i, j) = \delta(j)M(j, i)$. The distribution of $X_\nu(1)$ is $\nu M$ and the distribution of $X_\nu(t)$ is $\nu M^t$.

**Definition 2.2.** *A distribution $\pi$ on the vertices of $G$ is a stationary distribution if for any $t > 0$ the distribution $X_\pi(t)$ is identical to $\pi$. Equivalently, it means that $\pi M = \pi$.*

For any strongly connected graph there exists a unique stationary distribution. For example, for an undirected graph $G(V, E)$ the stationary distribution is $\pi(v) = \frac{\delta(v)}{2|E|}$. An important property of the stationary distribution is that if $G$ is aperiodic then, for any starting vertex $v$, the distribution $X_v(t)$ converges to the stationary distribution (as $t$ goes to infinity). An important parameter of a graph is how quick the convergence to the stationary distribution is:

**Definition 2.3.** *The mixing time* $\text{mix}$ *of a graph $G$ is the smallest integer $t$ such that for any starting vertex $v$ we have that the random walk $X_v(t)$ is close to the stationary distribution. Formally,*

$$\text{mix} = \min\left\{t \mid \forall v, \left\|X_v(t) - \pi\right\|_1 \leq (2\boldsymbol{e})^{-1}\right\}.$$

Note that a constant $(2\mathbf{e})^{-1}$ is an arbitrary constant. The following fact shows that if we replace $(2\mathbf{e})^{-1}$ by any other constant it will not change $\text{mix}$ too much.

**Fact 2.4** (cf. [LWP] chapter 4)**.** *For any starting vertex $v$ and for any non-negative integer $s$, $X_v(s \cdot \text{mix})$ is $e^{-s}$ close to a stationary distribution i.e.*

$$\left\|X_v(s \cdot \text{mix}) - \pi\right\|_1 \leq \boldsymbol{e}^{-s}$$

**Definition 2.5.** *Let $X$ be any discrete random variable. Then the expectation of $X$ is:*

$$\boldsymbol{E}(X) = \sum_{i \in \textit{Image(X)}} i \cdot \Pr[X = i]$$

**Fact 2.6.** *Let $X$ be any non-negative random variable that takes integral values. Then:*

$$\boldsymbol{E}(X) = \sum_{t=0}^{\infty} \Pr[X > t]$$

**Fact 2.7.** *Let $a > 0$ be any positive number and $X$ be any non-negative random variable that takes integral values. Then:*

$$a\sum_{t=1}^{\infty} \Pr[X > at] \leq \boldsymbol{E}(X) = \sum_{t=0}^{\infty} \Pr[X > t] \leq a\sum_{t=0}^{\infty} \Pr[X > at]$$

**Fact 2.8.** *Let $a > 0, b > 0$ be any two positive numbers and $X$ be any non-negative random variable that takes integral values. Then:*

$$\sum_{t=0}^{\infty} \Pr[X \geq at + b] \geq \frac{\boldsymbol{E}X - b}{a} - 1$$

Let $\varsigma(u, v)$ be the time it takes for a random walk that starts at $u$ to reach $v$ i.e. $\varsigma(u, v) = \min\{t \mid X_u(t) = v\}$. Note that $\varsigma(u, v)$ is a random variable. Let the hitting time $h(u, v) = \mathbf{E}(\varsigma(u, v))$ be the expected time for the random walk to traverse from $u$ to $v$. Let $h_{\max} = \max_{u,v \in V} h(u, v)$ and $h_{\min} = \min_{u,v \in V} h(u, v)$ be the maximal and minimal hitting times. Similarly let $\tau_u$ be the time for the simple random walk to visit all the nodes of the graph. More formally $\tau_u = \min\{t \mid \{X_u(1), X_u(2), \ldots, X_u(t)\} = V\}$ . Let $C_u = \mathbf{E}(\tau_u)$ be the cover time for a simple walk starting at $u$. The cover time $C = \max_u(C_u)$ is the maximal (over the starting vertex $u$) expected time it takes for a single walk to cover the graph. The following theorem gives lower and upper bounds on the cover time (the bounds are known to be tight up to lower order terms).

**Theorem 2.9** ([Fei95b, Fei95a]). *For any (connected) undirected graph $G = (V, E)$ with $n$ vertices it holds that:*

$$n \ln n \le C \le \frac{4}{27} n^3.$$

The following theorem provides fundamental bounds on the cover time in terms of the hitting time (for more details see [LWP] Chapter 11 or [Mat88]):

**Theorem 2.10** (cf. [Mat88]). *For every graph G with $n$ vertices*

$$h_{min} \cdot \log n \le C \le h_{max} \cdot \log n.$$

Note that there also exists a trivial bound of $h_{max} \le C$. It will be convenient for us to define the following parameter of a graph: $H(G) = \frac{C(G)}{h_{max}(G)}$. Note that $1 \le H(G) \le \log n$. Also note that there exist graphs where $H(G) = O(1)$ (for example the cycle), and there exist graphs with $H(G) = \Omega(\log n)$ (for example the complete graph).

The *lazy* random walk on a graph $G$ is a walk that at each time step walks on a random outgoing edge with probability $1/2$ and with probability $1/2$ stays in place. Alternatively, one can see it as a simple random walk where the time between every two consecutive steps is distributed according to a geometric distribution. We will also refer to the notion of a continuous-time random walk, which is a simple random walk on a graph $G$ except that the times between every two consecutive steps of the walk is exponentially distributed (in particular, the walk is not limited to move in integral times). The continuous-time random walk is a continuous version of the lazy random walk. We will denote by $h_{\text{lazy}}(u, v)$ and $C_{\text{lazy}}$ the hitting and cover times of a lazy random walk.

**Remark 2.11.** *We want to mention here that the hitting time and the cover time of a lazy random walk is exactly twice the hitting time and the cover time of a non-lazy random walk. (cf. [AF99] chapter 2) i.e. $h_{lazy}(u, v) = 2h(u, v)$ and $C_{lazy} = 2C$.*

The next theorem gives a relation between hitting and mixing times of the graph.

**Theorem 2.12** (cf. [LWP] chapter 10). *For every lazy random walk we have that* $\text{mix} \le 2h_{\text{max}}$.

**Remark 2.13.** *In order to simplify the notation we will omit $G$ from the above notation, when the graph $G$ will be clear from the context.*

Let us state a useful lemma about probability measures. Though this lemma is well known, we give here its proof for completeness.

**Lemma 2.14.** *Let $\mu = (\mu_1, \mu_2), \nu = (\nu_1, \nu_2)$ be two probability measures on $V \quad V$. For every $x_1$ in the support of $\mu_1$, let $\mu_{x_1}$ be the conditional probability measure on $V$ (i.e. $\mu_{x_1}(x_2) = \frac{\mu(x_1,x_2)}{\mu_1(x_1)}$). Similarly, for every $x_1$ in the support of $\nu_1$, let $\nu_{x_1}$ be the conditional probability measure on $V$. Then if $\| \mu_1 - \nu_1 \|_1 \le \varepsilon_1$ and for every $x_1$ which is in the support of both $\mu_1$ and $\nu_1$ we have that $\| \mu_{x_1} - \nu_{x_1} \|_1 \le \varepsilon_2$ then*

$$\| \mu - \nu \|_1 \le \varepsilon_1 + \varepsilon_2$$

*Proof.* To simplify notation we will assume that the support of both $\mu_1$ and $\nu_1$ is $V$.

$$\| \mu - \nu \|_1 = \int_x \int_y |\mu(x,y) - \nu(x,y)| \,\mathbf{d}x\mathbf{d}y$$

By definition of $\mu_x, \nu_x$ it is equal to:

$$\| \mu - \nu \|_1 = \int_x \int_y |\mu_1(x)\mu_x(y) - \nu_1(x)\nu_x(y)| \,\mathbf{d}x\mathbf{d}y$$

Let us use the following equality:

$$\mu_1(x)\mu_x(y) - \nu_1(x)\nu_x(y) = \mu_1(x)(\mu_x(y) - \nu_x(y)) + \nu_x(y)(\mu_1(x) - \nu_1(x))$$

and we will get

$$\begin{aligned}
&\int_x \int_y |\mu_1(x)(\mu_x(y) - \nu_x(y)) + \nu_x(y)(\mu_1(x) - \nu_1(x))| \,\mathbf{d}x\mathbf{d}y \leq \\
&\int_x \int_y |\mu_1(x)(\mu_x(y) - \nu_x(y))| \,\mathbf{d}x\mathbf{d}y + \int_x \int_y |\nu_x(y)(\mu_1(x) - \nu_1(x))| \,\mathbf{d}x\mathbf{d}y = \\
&\int_x |\mu_1(x)| \int_y |(\mu_x(y) - \nu_x(y))| \,\mathbf{d}x\mathbf{d}y + \int_y |\nu_x(y)| \int_x |(\mu_1(x) - \nu_1(x))| \,\mathbf{d}x\mathbf{d}y \leq \\
&\varepsilon_1 + \varepsilon_2.
\end{aligned}$$

□

As a corollary from this lemma we get:

**Corollary 2.15.** *Let $\mu_1, \mu_2 \ldots \mu_k, \nu_1, \nu_2 \ldots, \nu_k$ be independent probability measures on $V$ such that for all $i$ it holds that $\| \nu_i - \mu_i \|_1 \leq \varepsilon$. Let $\mu = (\mu_1, \mu_2 \ldots \mu_k)$ and $\nu = (\nu_1, \nu_2 \ldots, \nu_k)$ be the two corresponding probability measures on $V^k$. Then:*

$$\| \nu - \mu \|_1 \leq k\varepsilon.$$

## 2.1 Multiple random walks

We will now like to consider the behavior of $k$ parallel independent random walks on a graph. Here we are interested in how much time it will take until at least one of these walks will hit some specific vertex and in the time it takes for all the nodes to be covered. For illustration, imagine $k$ runners, carrying the same news and following independent random paths (assume they begin at time $0$ from nodes $u_1, u_2, \ldots u_k$). We are interested in two main questions: the time until one of these runners will bring the news to some specific node and the time until each node will get the news. Defining it more formally $\varsigma(\{u_1, u_2, \ldots u_k\}, v) = \min_{i=1}^{k} \varsigma(u_i, v)$ is the random variable corresponding to the hitting time of $k$ random walks, where some of the $u_i$'s may be equal. Let $h(\{u_1, u_2, \ldots u_k\}, v) = \mathbf{E}(\varsigma(\{u_1, u_2, \ldots u_k\}, v))$ be the hitting time of $k$ random walks starting at vertices $u_i$. If all the walks start at the same vertex $u$ we will write it as $h^k(u, v)$. Let $h^k_{\max} = \max_{u_i, v} h(\{u_1, u_2, \ldots u_k\}, v)$ be the maximal hitting time of $k$ random walks. Similarly, for the cover time we define $\tau_{u_1,u_2,\ldots u_k} = \min\{t \mid \bigcup_{i=1}^{k} \{X_{u_i}(1), X_{u_i}(2), \ldots X_{u_i}(t)\} = V\}$ and define

$C_{u_1,u_2,\dots u_k} = \mathbf{E}\tau_{u_1,u_2,\dots u_k}$ to be the expected cover time. Let $C^k = \max_{u_1,u_2,\dots u_k} C_{u_1,u_2,\dots u_k}$. In this paper we analyze the relationship between the hitting time of a single random walk and multiple random walks as well as the relationship between their cover times.

The proof of Theorem 2.10 (see [LWP] Chapter 11) easily extends to multiple walks implying the following theorem:

**Theorem 2.16.** *For every (strongly connected) graph $G$ with $n$ vertices, and for every $k$*

$$\frac{C_k}{h^k_{\max}} \leq \log n.$$

# 3 Hitting time of multiple random walks

In this section we study the behavior of the hitting time of $k$ random walks. The first question we will consider is: what are the starting vertices of multiple random walks which maximize the hitting time? Later, we will give a lower bound on the maximal hitting time of multiple random walks. We will prove that $\frac{h_{\max}}{h^k_{\max}} = O(k)$. Then we will consider the case where the walks' starting vertices are chosen independently according to the stationary distribution. Note that in this setting the ratio between hitting times is not upper bounded by $O(k)$; in fact it may even be exponential in $k$. We will prove that in this setting the ratio between the hitting time of the single walk and the hitting time of $k$ walks is at least $k$. Next we will use this theorem in order to prove that for graphs with small mixing time the ratio $\frac{h_{\max}}{h^k_{\max}} = \Omega(k)$. Finally, we consider the evaluation of the hitting time and prove some other lemmas about the hitting time, which will be useful in our analysis of the cover time.

## 3.1 Worst to start in a single vertex

Let us prove that the maximal hitting time is achieved when all the walks start from the same node.

**Theorem 3.1.** *For every graph $G = (V, E)$, for every $v \in V$ it holds that*

$$\max_{u_1,u_2,\dots u_k} h(\{u_1, u_2, \dots u_k\}, v) = \max_u h^k(u, v).$$

Before we prove this theorem let us state a generalization of Hölder's Inequality:

**Lemma 3.2** (cf. [BB65], pp 20)**.** *[Hölder] Let $a_i : \mathbb{N} \mapsto R^+$ for $i = 1, \dots, n$ be $n$ sequences such that for all $i$ we have $\sum_{j=0}^{\infty} a_i(j)^n < \infty$. Then*

$$\sum_{j=0}^{\infty}\left(\prod_{i=1}^{n} a_i(j)\right) \leq \left(\prod_{i=1}^{n}(\sum_{j=0}^{\infty} a_i(j)^n)\right)^{\frac{1}{n}}.$$

*Proof of Theorem 3.1.* Let us prove that for every set of starting vertices $\{u_1, u_2, \dots u_k\} \subset V$ and for every end vertex $v \in V$ there exists some starting vertex $u_j$ such that:

$$h(\{u_1, u_2, \dots u_k\}, v) \leq h^k(u_j, v). \tag{1}$$

Recall that $\varsigma(u,v) \triangleq \min\{t \mid X_u(t) = v\}$ is the first time the random walk hits $v$ and $\varsigma(\{u_1, u_2, \dots u_k\}, v) \triangleq \min_i \varsigma(u_i, v)$ is the first time $k$ random walks hit $v$. Let us denote by $a_i(t) = \Pr[\varsigma(u_i, v) > t]$ the probability that the $i^{\text{th}}$ random walk *does not* hit $v$ before time $t$. Then,

$$\Pr[\varsigma(\{u_1, u_2, \dots u_k\}, v) > t] = \Pr[\min_i \varsigma(u_i, v) > t] = \Pr[\forall i\ \varsigma(u_i, v) > t] = \prod_{i=1}^{k} a_i(t).$$

Therefore from Fact (2.6) it follows that:

$$h(\{u_1, u_2, \dots u_k\}, v) = \mathbf{E}(\varsigma(\{u_1, u_2, \dots u_k\}, v)) = \sum_{j=0}^{\infty} \left( \prod_{i=1}^{k} a_i(j) \right). \tag{2}$$

By Lemma 3.2 (Hölder's Inequality) we have that

$$h(\{u_1, u_2, \dots u_k\}, v) \leq \left( \prod_{i=1}^{k} (\sum_{j=0}^{\infty} a_i(j)^k) \right)^{\frac{1}{k}}. \tag{3}$$

From Equation (2) it follows that:

$$h^k(u_i, v) = \sum_{j=0}^{\infty} a_i(j)^k.$$

So by Equation (3) we have that:

$$h(\{u_1, u_2, \dots u_k\}, v) \leq \left( \prod_{i=1}^{k} (h^k(u_i, v)) \right)^{\frac{1}{k}},$$

and in particular $h(\{u_1, u_2, \dots u_k\}, v) \leq \max_{i=1}^{k}(h^k(u_i, v))$. □

## 3.2 Upper bound on the speedup of the hitting time of multiple random walks

We will now prove that the ratio between the hitting time of a single random walk and the hitting time of $k$ random walks is at most $O(k)$.

**Theorem 3.3.** *For any graph $G$*

$$h_{\text{max}} \leq 4kh^k_{\text{max}}.$$

*Proof.* Recall that $\varsigma^k(u,v)$ is the time it takes for $k$ random walks starting at $u$ to hit $v$. Fix any two vertices $u, v \in V$,we need to prove that $h(u,v) \leq 4kh^k_{\text{max}}$.

By Markov Inequality we have that:

$$\Pr[\varsigma^k(u,v) > 2h^k_{\text{max}}] \leq \frac{1}{2}.$$

Equivalently,

$$\Pr[\varsigma^k(u,v) \leq 2h^k_{\text{max}}] \geq \frac{1}{2}.$$

By definition,

$$\Pr[\varsigma^k(u,v) \le 2h^k_{\max}] = \Pr[\min_{i=1,\dots k} \varsigma_i(u,v) \le 2h^k_{\max}].$$

So we have that after $2h^k_{\max}$ steps one out of $k$ random walks will hit the vertex $v$ with probability at least $1/2$. By a union bound we get that $\Pr[\varsigma(u,v) \le 2h^k_{max}] \ge \frac{1}{2k}$. So the probability that a single random walk will hit $v$ in $2h^k_{\max}$ steps is at least $\frac{1}{2k}$ and this is true for any two nodes $u, v \in V$. Thus for any integer $t$ the probability that a random walk of length $2th^k_{\max}$ does not visit $v$ is at most $(1 - \frac{1}{2k})^t$. (One can view the walk as $t$ sequential attempts to visit $v$.) Namely,

$$\Pr[\varsigma(u,v) \ge t(2h^k_{\max})] \le \left(1 - \frac{1}{2k}\right)^t. \tag{4}$$

Let us now prove that $h(u,v) \le 4kh^k_{\max}$. Recall that $h(u,v)$ is the expected value of $\varsigma(u,v)$. From Fact 2.7 it follows that:

$$h(u,v) = \sum_{t=0}^{\infty} \Pr[\varsigma(u,v) > t] \le 2h^k_{\max} \sum_{t=0}^{\infty} \Pr[\varsigma(u,v) > t(2h^k_{\max})].$$

Using Equation (4) we get

$$h(u,v) \le 2h^k_{max} \sum_{t=0}^{\infty} \left(1 - \frac{1}{2k}\right)^t = 4kh^k_{max}.$$

$\square$

By a slightly more complicated argument we can replace the constant 4 in Theorem3.3 by $\mathbf{e} + o(1)$. However it seems plausible that the right constant is $1$.

**Open Problem 3.4.** *Prove or disprove that for any graph $G$*

$$h_{\max} \le kh^k_{\max}.$$

### 3.3 Lower bounds on the speedup of the hitting time of multiple random walks

In this section, we consider the case where the starting vertices of the random walks are selected according to the stationary distribution. Theorem3.3 shows that for worst-case starting vertices the ratio between the hitting times of a single walk and multiple walks is at most $O(k)$. But as we will soon show, when the starting vertices of all walks are drawn independently from the stationary distribution then, loosely speaking, this ratio becomes at least $k$. Note that in some graphs the ratio of hitting times, when the starting vertices are selected according to the stationary distribution, may even become exponential in $k$. Indeed, such an example is given in Figure1 and is discussed in the introduction (the discussion there is for the cover time but the analysis for the hitting time is very similar)

The main technical tool of this section is the following lemma:

**Lemma 3.5** (cf. [AF99], Chapter 3, pp 29 )**.** *Let $G(V,E)$ be a (connected) undirected graph and let $u$ be chosen according to the stationary distribution. Then:*

$$\Pr[\varsigma(u,v) > s + t] \ge \Pr[\varsigma(u,v) > s] \Pr[\varsigma(u,v) > t].$$

We want to mention here that in [AF99] this lemma is stated for a continuous-time Markov chain. We want to give here proof due to Ofer Zeitouni [Zei09] of this lemma for the case when $s = t$ for two reasons: this proof is works also for discreet random walk and this proof is very simple and provides an intuition.

**Lemma 3.6.** *Let $G(V, E)$ be a (connected) undirected graph and let $u$ be chosen according to the stationary distribution. Then:*

$$\Pr[\varsigma(u, v) > 2t] \geq \Pr[\varsigma(u, v) > t]^2 \tag{5}$$

*Proof.* The following simple claim is the heart of the proof of this lemma:

**Claim 3.7.** *Let $X(t)$ be a random walk on undirected graph which starts from the stationary distribution(i.e. $\Pr(X(0) = u) = \pi(u)$) Then for any sequence of nodes $x_0, x_1, \ldots x_t$ it holds that:*

$$\Pr(X(0) = x_0, X(1) = x_1, \ldots, X(t) = x_t) = \Pr(X(0) = x_t, X(1) = x_{t-1}, \ldots, X(t) = x_0).$$

*Proof.* Then for any sequence of nodes $x_0, x_1, \ldots x_t$ it holds that:

$$\begin{gathered}\Pr(X(0) = x_0, X(1) = x_1, \ldots, X(t) = x_t) = \\ \frac{d(x_0)}{2m} \frac{1}{d(x_0)} \frac{1}{d(x_1)} \cdots \frac{1}{d(x_{t-1})} = \\ \frac{1}{2md(x_1)d(x_2)\ldots d(x_{t-1})},\end{gathered}$$

therefore it holds that,

$$\Pr(X(0) = x_0, X(1) = x_1, \ldots, X(t) = x_t) = \Pr(X(0) = x_t, X(1) = x_{t-1}, \ldots, X(t) = x_0).$$

□

On intuitive level this claim states that reversible random walk which starts from stationary distribution could be reveresed. The intuition for proof is let us take random walk of size $2t$ then this walk could be seen as two walks one walk is walk from time $t$ to time $2t$ and the second walk is the reverse walk starting at time $t$ and going down to time $0$ then from claim follows that reverse walk is equvalent to strait walk. Therefore the probability that random walk walk will not hit $v$ at time $2t$ equal to product of probabilties that two random walks of length $t$ will not hit node $v$. Let $S$ be the set of all sequences of nodes of size $2t + 1$ which does not contain $v$. Then it holds that:

$$\begin{gathered}\sum_{u \in V} \pi(u) \Pr(\varsigma(u, v) > 2t) = \\ \sum_{\vec{x} \in S} \Pr(X(0) = x_0, X(1) = x_1, \ldots, X(t) = x_{2t})\end{gathered}$$

Using conditional probability we get:

$$\begin{gathered}\sum_{\vec{x} \in S} \Pr(X(0) = x_0, X(1) = x_1, \ldots, X(t) = x_{2t}) = \\ \sum_{\vec{x} \in S} \Pr(X(0) = x_0, \ldots, X(t) = x_t) \Pr(X(t+1) = x_{t+1}, \ldots, X(t) = x_{2t} | X(t) = x_t) = \\ \sum_{\vec{x} \in S} \Pr(X(0) = x_0, \ldots, X(t) = x_t) \frac{\Pr(X(t) = x_t, \ldots, X(t) = x_{2t})}{\Pr(X(t) = x_t)}\end{gathered}$$

From claim above it follows that:

$$\sum_{\vec{x}\in S} \Pr(X(0) = x_0, \ldots, X(t) = x_t) \frac{\Pr(X(t)=x_t,\ldots,X(t)=x_{2t})}{\Pr(X(t)=x_t)} =$$
$$\sum_{\vec{x}\in S} \Pr(X(0) = x_t, \ldots, X(t) = x_0) \frac{\Pr(X(t)=x_t,\ldots,X(t)=x_{2t})}{\Pr(X(t)=x_t)} =$$

By definition this expression is equal to

$$\sum_{x_t\in V} (\pi(x_t)^2) \frac{\Pr(\varsigma(x_t,v)>t)^2}{\Pr(X(t)=x_t)} =$$
$$\sum_{u\in V} \pi(u) \Pr(\varsigma(x_t, v) > t)^2$$

Using Cauchy-Schwarz inequality we get the desired result:

$$\sum_{u\in V} \pi(u) \Pr(\varsigma(u, v) > t)^2 \geq \left( \sum_{u\in V} \pi(u) \Pr(\varsigma(u, v) > t) \right)^2$$

□

The next theorem gives a lower bound on the ratio between hitting times for random starting vertices.

**Theorem 3.8.** *Let $G(V, E)$ be a (connected) undirected graph. Let $X$ be a random walk on $G$. Let $u, u_1, \ldots u_k \in V$ be independently chosen according to the stationary distribution of $G$. Then for any $k = 2^i$ it holds that:*

$$\boldsymbol{E}_u(h(u, v)) \geq k\, (\boldsymbol{E}_{u_i} h(\{u_1, u_2, \ldots u_k\}, v) - 1)\,.$$

Since any $k$ can be approximate up to factor $2$ by $2^i$ we get the following corollary from this theorem.

**Corollary 3.9.** *Let $G(V, E)$ be a (connected) undirected graph. Let $u, u_1, \ldots u_k \in V$ be independently chosen according to the stationary distribution of $G$. Then:*

$$\boldsymbol{E}_u(h(u, v)) \geq k/2\, (\boldsymbol{E}_{u_i} h(\{u_1, u_2, \ldots u_k\}, v) - 1)\,.$$

*Proof of Theorem 3.8.* Recall that $h(u, v) = \mathbf{E}\varsigma(u, v)$. From Facts 2.6 and 2.7 it follows that:

$$h(u, v) = \sum_{t=0}^{\infty} \Pr[\varsigma(u, v) > t] \geq k \sum_{t=1}^{\infty} \Pr[\varsigma(u, v) > kt].$$

By Lemma 3.5 we have

$$k \sum_{t=1}^{\infty} \Pr[\varsigma(u, v) \ > kt] \geq k \sum_{t=1}^{\infty} \Pr[\varsigma(u, v) > t]^k.$$

Let us note that

$$\Pr[\varsigma(u, v) \geq t]^k = \Pr[\varsigma(\{u_1, u_2, \ldots u_k\}, v) \geq t],$$

where $u_1, u_2, \ldots u_k$ are independently distributed according to the stationary distribution. Therefore,

$$k \sum_{t=1}^{\infty} \Pr[\varsigma(u, v) > t]^k = k\, (\mathbf{E}_{u_i} h(\{u_1, u_2, \ldots u_k\}, v) - 1)\,,$$

□

**Lower bound on the speedup for worst starting vertices** The lower bound on the speedup for walks starting at the stationary distribution translates into a lower bound that also applies to walks starting at the worst vertices: First let the walks converge to the stationary distribution and then apply the previous lower bound. The bounds that we obtain are especially meaningful when the mixing time of the graph is sufficiently smaller than the hitting time.

**Theorem 3.10.** *Let $G(V,E)$ be a (connected) undirected graph. Then*

$$h^k_{\max} \leq \frac{2h_{\max}}{k} + O(\mathrm{mix}(\log n + \log k)).$$

*Proof of Theorem 3.10.* Let $X^1(t), X^2(t), \dots X^k(t)$ be $k$ independent random walks beginning at some arbitrary node $u$. By ignoring the first $\mathrm{mix}(3\log n + \log k)$ steps of each walk we get $k$ related random walks: $\tilde{X}^1(t), \tilde{X}^2(t), \dots \tilde{X}^k(t)$ where

$$\tilde{X}^i(t) = X^i(t + \mathrm{mix}(3\log n + \log k).$$

Let $\nu = (\nu_1, \nu_2, \dots \nu_k)$ be the distribution of the starting vertex of the new walks $\tilde{X}^i$. Let $\pi_k = \underbrace{(\pi, \pi, \dots \pi)}_{k \text{ times}}$ be the distribution of $k$ independent vertices drawn from the stationary distribution. From Fact 2.4 we know that the distribution of $X_u(\mathrm{mix}(3\log n + \log k))$ is $k^{-1}n^{-3}$ close to the stationary distribution i.e. $\| \nu_i - \pi \|_1 \leq k^{-1}n^{-3}$. Since the $k$ distributions $\nu_i$ are independent then from Corollary 2.15 it follows that

$$\| \nu - \pi_k \|_1 \leq k(k^{-1}n^{-3}) = n^{-3}. \tag{6}$$

We therefore have that the expected time it will take the $k$ walks $\tilde{X}^i$ to hit $v$ is close to the expected hitting time for $k$ walks which start at the stationary distribution:

$$\begin{aligned} & \left(\mathbf{E}_{\{u_i \sim \nu\}} h(< u_1, u_2, \dots u_k >, v) - \mathbf{E}_{\{u_i \sim \pi_k\}} h(< u_1, u_2, \dots u_k >, v)\right) \\ \leq\ & \| \nu - \pi_k \|_1 \ h^k_{\max} \\ \leq\ & n^{-3} h^k_{\max} \end{aligned}$$

Where the first inequality is obtained by viewing $\mathbf{E}h(< u_1, u_2, \dots u_k >, v)$ as a function of $< u_1, u_2, \dots u_k >$ which takes positive values bounded by $h^k_{\max}$. By the definition of the distributions $\tilde{X}^i$ it follows that:

$$\begin{aligned} & h^k(u,v) \\ \leq\ & \mathrm{mix}(3\log n + \log k) + \mathbf{E}_{\{u_i \sim \nu\}} h(< u_1, u_2, \dots u_k >, v) \\ \leq\ & \mathrm{mix}(3\log n + \log k) + \mathbf{E}_{\{u_i \sim \pi_k\}} h(< u_1, u_2, \dots u_k >, v) + n^{-3} h^k_{\max}. \end{aligned}$$

From Theorem 2.9 it follows that $h^k_{\max} \leq O(n^3)$ thus $n^{-3}h^k_{\max} \leq O(1)$. So:

$$h^k(u,v) \leq \mathrm{mix}(3\log n + \log k) + \mathbf{E}_{\{u_i \sim \pi_k\}} h(< u_1, u_2, \dots u_k >, v) + O(1). \tag{7}$$

From Theorem 3.9 we have

$$\mathbf{E}_{\{u_i \sim \pi_k\}} h(< u_1, u_2, \dots u_k >, v) \leq \frac{2h_{\max}}{k} + 1.$$

Thus from Equation (7) it follows that:

$$h^k_{\max} \leq \mathrm{mix}(3\log n + \log k) + \frac{2h_{\max}}{k} + O(1).$$

$\square$

As a corollary we get:

**Corollary 3.11.** *Let $G(V,E)$ be a (connected) undirected graph such that $k\mathrm{mix}(\log n + \log k) = o(h_{\max})$. Then:*

$$\frac{h_{\max}}{h^k_{\max}} \geq k/2(1 - o(1)).$$

## 3.4 Calculating the hitting time of multiple random walks

We would like to address a question which is somewhat orthogonal to the main part of this paper. Namely, we would like to discuss how the hitting time of multiple walks can be calculated. Let us observe that multiple random walks on graph $G$ can be presented as a single random walk on another graph $G^k$.

**Definition 3.12.** *Let $G = (V,E)$ be some graph. Then the graph $G^k = (V',E')$ is defined as follows: The vertices of $G^k$ are $k$-tuples of vertices of $G$ i.e.*

$$V' = \underbrace{V \oplus V \ldots \oplus V}_{k \text{ times}} = V^k.$$

*For every $k$ edges of $G$, $(u_i, v_i)$ for $i = 1, \ldots, k$ we have an edge between $u' = (u_1, u_2, \ldots u_k)$ and $v' = (v_1, v_2 \ldots v_k)$ in $G^k$*

One can view $k$ random walks on $G$ as a single random walk on $G^k$ where the first coordinate of $G^k$ corresponds to the first random walk, the second coordinate corresponds to the second random walk, and so on.

Let $A \subset V^k$ be the set of all nodes of $G^k$ which contain the node $v \in V$. Assume that we have $k$ random walks beginning at $u_1, u_2, \ldots u_k$. Then the time it will take to hit $v$ is equal to the time for a single random walk on $G^k$ beginning at node $(u_1, u_2, \ldots u_k)$ to hit the set $A$. Thus instead of analyzing multiple random walks we can study a single random walk on $G^k$. There is a polynomial time algorithm for calculating hitting times of a single random walk (cf. [Lov96]). This gives us an algorithm, which is polynomial in $n^k$, for calculating $h(\{u_1, u_2, \ldots u_k\}, v)$. A natural question is whether there exist more efficient algorithms.

**Open Problem 3.13.** *Find a more efficient algorithm for calculating $h(\{u_1, u_2, \ldots u_k\}, v)$.*

## 3.5 Useful facts about hitting times

Let us give some basic facts about the hitting times of random walks which we will use later.

**Lemma 3.14** (cf. [AF99] Chapter 2, Section 4.3)**.** *Let $X(t)$ be a random walk starting at any node $u$. Then for any non-negative integer $a$ and for any node $v$ the time $\varsigma(u,v)$ until the walk will hit $v$ is bounded as follows:*

$$\Pr[\varsigma(u,v) > \boldsymbol{e} a h_{\max}] \le \boldsymbol{e}^{-a},$$

Using this lemma we can bound the sum of $k$-independent hitting times:

**Lemma 3.15.** *Let $X_1, X_2, \dots X_k$ be $k$ independent random walks starting at any $k$ vertices $u_i$. Let $\varsigma_i(u_i, v_i)$ be the time until walk $i$ will hit some node $v_i$. Then:*

$$\Pr[\sum_{i=1}^{k} \varsigma_i(u_i, v_i) > ksh_{\max}] \le \boldsymbol{e}^{-(s/2e-2)k}.$$

*Proof.* It holds that:

$$\Pr\left[\sum \varsigma_i(u_i,v_i) > ksh_{\max}\right] = \Pr\left[\mathbf{e}^{\sum \frac{\varsigma_i(u_i,v_i)}{2\mathbf{e}h_{\max}}} > \mathbf{e}^{\frac{ks}{2\mathbf{e}}}\right].$$

In order to simplify the notation let us set $r = \frac{1}{2\mathbf{e}h_{\max}}$. By Markov Inequality it follows that:

$$\Pr\left[\sum \varsigma_i(u_i,v_i) > ksh_{\max}\right] \le \frac{\mathbf{E}\left(\mathbf{e}^{\sum r\varsigma_i(u_i,v_i)}\right)}{\mathbf{e}^{\frac{ks}{2\mathbf{e}}}}.$$

From the independence of $\varsigma_i(u_i, v_i)$ it follows that:

$$\Pr[\sum \varsigma_i(u_i,v_i) > ksh_{\max}] \le \prod_{i=1}^{k} \frac{\mathbf{E}\left(\mathbf{e}^{r\varsigma_i(u_i,v_i)}\right)}{\mathbf{e}^{\frac{s}{2\mathbf{e}}}}. \tag{8}$$

Thus in order to prove the lemma it is enough to prove that $\mathbf{E}\left(\mathbf{e}^{r\varsigma_i(u_i,v_i)}\right) \le \mathbf{e}^2$. From Fact 2.6 it follows that:

$$\mathbf{E}(\mathbf{e}^{r\varsigma_i(u_i,v_i)}) = \int_{x=0}^{\infty} \Pr[\mathbf{e}^{r\varsigma_i(u_i,v_i)} > x]dx.$$

By substituting $x = \mathbf{e}^y$ we get that:

$$\mathbf{E}(\mathbf{e}^{r\varsigma_i(u_i,v_i)}) = \int_{y=-\infty}^{\infty} \Pr[\mathbf{e}^{r\varsigma_i(u_i,v_i)} > \mathbf{e}^y]\mathbf{e}^y dy$$

Since probabilities are at most one it follows that:

$$\begin{aligned} & \mathbf{E}(\mathbf{e}^{r\varsigma_i(u_i,v_i)}) \\ \le \; & \int_{-\infty}^{1/2} \mathbf{e}^y \mathbf{d}y + \int_{y=1/2}^{\infty} \Pr[\mathbf{e}^{r\varsigma_i(u_i,v_i)} > \mathbf{e}^y]\mathbf{e}^y dy \\ \le \; & \sqrt{\mathbf{e}} + \int_{y=1/2}^{\infty} \Pr[\mathbf{e}^{r\varsigma_i(u_i,v_i)} > \mathbf{e}^y]\mathbf{e}^y dy \end{aligned}$$

Equivalently, (recalling that $r = \frac{1}{2\mathbf{e}h_{\max}}$) we can rewrite:

$$\mathbf{E}(\mathbf{e}^{r\varsigma_i(u_i,v_i)}) \le \sqrt{\mathbf{e}} + \int_{y=1/2}^{\infty} \Pr[\varsigma_i(u_i,v_i) > 2\mathbf{e}h_{\max}y]\mathbf{e}^y dy.$$

From Lemma 3.14 it follows that $\Pr[\varsigma_i(u_i, v_i) > 2\mathbf{e}h_{\max}y] \leq \mathbf{e}^{-\lfloor 2y \rfloor} \leq \mathbf{e}\mathbf{e}^{-2y}$. Thus

$$\mathbf{E}(\mathbf{e}^{r\varsigma_i(u_i,v_i)}) \leq \sqrt{\mathbf{e}} + \mathbf{e}\int_{y=1/2}^{\infty} \mathbf{e}^{-2y}\mathbf{e}^{y}dy = \sqrt{\mathbf{e}} + \frac{\mathbf{e}}{\sqrt{\mathbf{e}}} \leq \mathbf{e}^2.$$

Therefore, from Equation (8) it follows that:

$$\Pr[\sum \varsigma_i(u_i, v_i) > ksh_{\max}] \leq \mathbf{e}^{-\left(\frac{s}{2e}-2\right)k}.$$

$\square$

# 4 Cover time of multiple random walks

Let us turn our attention from the hitting time to the cover time. As in the case of the hitting time, the cover time heavily depends on the starting vertices of the random walks. The graph given by Figure 1 and discussed in the introduction gives an example where the speedup in cover time of $k$ random walks is linear in $k$ for worst-case starting vertices, it is exponential in $k$ for random starting vertices, and even for $k = 2$ it is $\Omega(n/\log n)$ for the best starting vertices.

Theorem 2.10 gives a relation between hitting times and cover times. Thus, our results on hitting times from the previous section also give us results on the cover times. In Subsection4.1 we will give these results and will analyze the speedup, $\frac{C}{C_k}$, for worst starting vertices. We show that it is bounded by $\min\{k^2, k\log n\}$ for any $k$. We will also show that for $k$ such that $k \log n\mathrm{mix} = O(h_{\max})$ the speedup is $\Omega(\frac{k}{\log n})$.

We will show in Subsection 4.2 that when $k$ random walks begin from the best starting vertices for $k = o(H(G))$ the speedup is roughly $k$ and is therefore essentially equal to the speedup for the worst case. In Subsection 4.3 we will show that when the starting vertices are drawn from the stationary distribution for $k$ such that $\mathrm{mix}k\log k = o(C)$, the speedup is at most $k$.

## 4.1 The worst starting vertices

As a simple corollary of Theorem 3.3 we obtain the following relation:

**Theorem 4.1.** *The speedup* $\frac{C}{C^k}$ *is at most* $4kH(G) \leq 4k\log n$

*Proof.* Recall that $C^k \geq h_{\max}^k$ so:

$$\frac{C}{C^k} \leq \frac{C}{h_{\max}^k} = \frac{h_{\max}}{h_{\max}^k}H(G).$$

From Theorem 3.3 follows:

$$\frac{h_{\max}}{h_{\max}^k}H(G) \leq 4kH(G).$$

And finally from Theorem 2.10 we have that $4kH(G) \leq 4k\log n$. $\square$

From this theorem it follows that for $k = \Omega(H(G))$ the speedup is $O(k^2)$. Theorem 4.4 implies that if $k < 0.01H(G)$ then the speedup $\frac{C}{C^k}$ is at most $2k$. Therefore, we can conclude a bound for every $k$:

**Theorem 4.2.** *For every (strongly connected) graph $G$ and every $k$, it holds that* $\frac{C}{C^k} = O(k^2)$.

From Theorem 3.10 we can also deduce a lower bound on the speedup for rapidly-mixing graphs:

**Theorem 4.3.** *Let $G(V, E)$ be an undirected graph and let $k$ be such that $k \log n \mathrm{mix} = O(h_{\max})$ then*

$$\frac{C}{C_k} \geq \Omega\left(\frac{k}{\log n}\right).$$

*Proof.* From Theorem 2.16 it follows that

$$\frac{C}{C_k} \geq \frac{h_{\max}}{h_{\max}^k \log n}.$$

Since $k \log n \mathrm{mix} = O(h_{\max})$, Theorem 3.10 implies that $\frac{h_{\max}}{h_{\max}^k} = \Omega(k)$. Thus:

$$\frac{C}{C_k} \geq \Omega\left(\frac{k}{\log n}\right).$$

□

## 4.2 The best starting vertices

As we discussed earlier, multiple random walks can be dramatically more efficient than a single random walk if their starting vertices are the best nodes (rather than the worst nodes). In fact, we have seen an example where taking two walks instead of one reduces the cover time by a factor of $\Omega(n/\log n)$. In this section we show that in graphs where the cover time is significantly larger than the hitting time, a few random walks cannot give such a dramatic speedup in the cover time, *even when starting at the best nodes*: If $k = o(H(G))$ (recall that $H(G) = \frac{C}{h_{max}}$), then the speedup $\frac{C}{C_{u_1,u_2,\ldots u_k}}$ (where $u_1, u_2, \ldots u_k$ are best possible) is not much bigger than $k$. Note that in the case where $k = o(H(G))$ it has been shown in [AAK+07] that the speedup $\frac{C}{C_{u_1,u_2,\ldots u_k}}$ is at least $k - o(k)$, even if $u_1, u_2, \ldots u_k$'s are worst possible. Combining the two results we get that the speedup is roughly $k$ regardless of where the $k$ walks start.

We want to show that the cover time of a single random walk is not much larger than $k$ times the cover time of $k$ random walks. For that we will let the single walk simulate $k$ random walks (starting from vertices $u_i$) as follows: The single walk runs until it hits $u_1$, then it simulates the first random walk. Then it runs until it hits $u_2$ and simulates the second random walk and so on until hitting $u_k$ and simulating the $k$'th random walk. The expected time to hit any vertex from any other vertex is bounded by $h_{\max}$. Thus intuitively the above argument should imply the following bound: $C \leq kC_{u_1,u_2,\ldots u_k} + kh_{\max}$. Unfortunately, we do not know how to formally prove such a strong bound. The difficulty is that the above argument only shows how a single walk can simulate $k$ walks for $t$ steps, where $t$ is fixed ahead of time. However, what we really need is for the single walk to simulate $k$ walks until the walks cover the graph. In other words, $t$ here is not fixed ahead of time but rather a random variable which depends on the $k$ walks. Nevertheless, we are still able to prove the following bound which is weaker by at most a constant factor:

**Theorem 4.4.** *For every graph $G$ and for **any** $k$ nodes $u_1, u_2, \dots u_k$ in $G$, it holds that:*

$$C \leq kC_{u_1,u_2,\dots u_k} + O(kh_{\max}) + O\left(\sqrt{kC_{u_1,u_2,\dots u_k} h_{\max}}\right).$$

*Proof.* Consider a single random walk, $\{X(i)\}_{i=1}^{\infty}$, starting at an arbitrary node $u$. Fix some integer $t$ and define $k$ sub-walks of $X(\cdot)$ of length $t$ each starting at the nodes $u_1, u_2, \dots u_k$, as follows. The starting vertex of the $i$th walk is denoted by $s_i$ and its end vertex is denoted by $e_i$, thus the $i$th sub-walk is $Y_i \triangleq \{X(s_i), X(s_i+1) \dots X(e_i)\}$, where both $s_i$ and $e_i$ are random variables defined as follows. $s_i$ is the minimal such that $X(s_i) = u_i$ under the condition that $s_1 > 0$ and for $i > 1$ it holds that $s_i > e_{i-1}$. The end vertices are simply defined by $e_i \triangleq s_i + t - 1$. Each $Y_i$ simulates the $i$th out of $k$ random walk. Note that since the start vertex of each $Y_i$ is fixed (to $u_i$) then by the Markov property of $X$ the sub-walks are all independent. In other words, the distribution $\{Y_1, Y_2, \dots, Y_k\}$ is indeed identical to the distribution of $k$ independent random walks of length $t$ starting at $u_1, u_2, \dots u_k$. Thus the probability that $\bigcup Y_i \neq V$ is equal to the probability that $k$ random walks starting at $\{u_i\}$ will not cover the graph by time $t$. Recall that $\tau_{u_1,u_2\dots,u_k}$ denotes the time it takes for the $k$ random walks to cover the graph, and $\tau = \tau_u$ is the time for the single walk. It follows that

$$\Pr[\bigcup Y_i \neq V] = \Pr[\tau_{u_1,u_2\dots,u_k} > t]. \tag{9}$$

For $i = 1, \dots, k$ define the random variable $\varsigma_i = s_i - e_{i-1}$ (when we define $e_0 = 0$). This random variable represents the time it takes the single random walk to hit node $u_i$ (after completing the sub-walk $Y_{i-1}$). Note that here too the $\varsigma_i$'s are independent random variables. Thus for any $s > 0$

$$\Pr[\tau > kt + ksh_{\max}] \leq \Pr\left[\left(\bigcup Y_i \neq V\right) \vee \left(\sum_{i=1}^{k} \varsigma_i > skh_{\max}\right)\right].$$

By the union bound we get

$$\Pr[\tau > kt + ksh_{\max}] \leq \Pr\left[\bigcup Y_i \neq V\right] + \Pr\left[\sum_{i=1}^{k} \varsigma_i > skh_{\max}\right].$$

From Equation (9) it follows that

$$\Pr[\tau > kt + ksh_{\max}] \leq \Pr[\tau_{u_1,u_2\dots,u_k} > t] + \Pr\left[\sum_{i=1}^{k} \varsigma_i > \ skh_{\max}\right]. \tag{10}$$

Lemma 3.15 gives us the following bound:

$$\Pr\left[\sum_{i=1}^{k} \varsigma_i > \ skh_{\max}\right] \leq \mathbf{e}^{-(\frac{s}{2e}-2)k}.$$

Using this bound in Equation (10) we get:

$$\Pr[\tau > kt + ksh_{\max}] \leq \Pr[\tau_{u_1,u_2\dots,u_k} > t] + \mathbf{e}^{-(\frac{s}{2e}-2)k}.$$

This equation is true for any value of $s$. We will set $s = (\frac{\delta t}{h_{\max}} + 2)2\mathbf{e}$, which implies

$$\Pr[\tau > kt + 2\mathbf{e}k\delta t + 4\mathbf{e}kh_{\max}] \le \Pr[\tau_{u_1,u_2\ldots,u_k} > t] + \mathbf{e}^{-\frac{\delta tk}{h_{\max}}}.$$

After summation on $t$ we get:

$$\sum_{t=0}^{\infty} \Pr[\tau > kt(1+2\mathbf{e}\delta) + 4\mathbf{e}kh_{\max}] \le \sum_{t=0}^{\infty} \Pr[\tau_{u_1,u_2\ldots,u_k} > t] + \sum_{t=0}^{\infty} \mathbf{e}^{-\frac{\delta tk}{h_{\max}}}$$

From Fact 2.6 it follows that $\sum_{t=0}^{\infty} \Pr[\tau_{u_1,u_2\ldots,u_k} > t] = C_{u_1,u_2,\ldots u_k}$. Thus

$$\sum_{t=0}^{\infty} \Pr[\tau > kt(1+2\mathbf{e}\delta) + 4\mathbf{e}kh_{\max}] \le C_{u_1,u_2,\ldots u_k} + \frac{h_{\max}}{\delta k} + 1. \tag{11}$$

On the other hand from Fact 2.8 it follows that:

$$\sum_{t=0}^{\infty} \Pr[\tau > kt(1+2\mathbf{e}\delta) + 4\mathbf{e}kh_{\max}] \ge \frac{C_u - 4\mathbf{e}kh_{\max}}{(1+2\mathbf{e}\delta)k} - 1. \tag{12}$$

From Equations (11) and (12) it follows that

$$\frac{C_u - 4\mathbf{e}kh_{\max}}{(1+2\mathbf{e}\delta)k} - 1 \le C_{u_1,u_2,\ldots u_k} + \frac{h_{\max}}{\delta k} + 1.$$

Equivalently, we can write it as

$$C_u \le kC_{u_1,u_2,\ldots u_k} + 4e(k+1)h_{\max} + 2ek\delta(C_{u_1,u_2,\ldots u_k} + O(1)) + \frac{h_{\max}}{\delta} + O(1).$$

Optimizing over $\delta$ we get:

$$\delta = \frac{1}{\sqrt{2\mathbf{e}kC_{u_1,u_2,\ldots u_k}h_{\max}}}.$$

So:

$$C_u \le kC_{u_1,u_2,\ldots u_k} + 4e(k+1)h_{\max} + 2\sqrt{2\mathbf{e}kC_{u_1,u_2,\ldots u_k}h_{\max}} + O(1).$$

□

In [AAK+07] the following theorem was proved:

**Theorem 4.5** (Theorem 5 from [AAK+07])**.** *Let $G$ be a strongly connected graph and $k = o(H(G))$ then $\frac{C}{C_k} \ge k - o(k)$.*

In the case where $k = o(H(G))$ then $O(kh_{\max}) + O(\sqrt{C_{u_1,u_2,\ldots u_k}kh_{\max}}) = o(C)$ and therefore $C \le kC_{u_1,u_2,\ldots u_k} + o(C)$. As a corollary we get:

**Corollary 4.6.** *Let $G$ be a strongly connected graph and $k = o(H(G))$ then for any starting vertices $u_1, u_2, \ldots u_k$ it holds that: $\frac{C}{C_{u_1,u_2,\ldots u_k}} = k \pm o(k)$*

It seems plausible that the speedup is at most $k$ for *any* starting vertices, also when $k$ is significantly larger than $H(G)$. When $k \geq \mathbf{e}^{H(G)}$ we can give an example where $kC_{u_1,u_2,\ldots u_k} << C$. Consider a graph $G$ which is composed of a clique of size $n$ and $t$ vertices where each vertex is connected by one edge to some node of a clique. We will assume that $n >> t$. The maximal hitting time for this graph is $O(n^2)$. The cover time of this graph is $O(n^2 \log t)$ and $H(G) = \log t$. If $k = t$ then when $k$ multiple random walks start from the $t$ vertices which are not in the clique, then $C_{u_1,u_2,\ldots u_k} = \frac{n \log n}{k} + O(1)$. Therefore, a natural open problem is the following:

**Open Problem 4.7.** *Prove or disprove that for some constant $\alpha > 0$, for any graph $G$, if $k \leq \mathbf{e}^{\alpha H(G)}$ then*

$$C \leq O(k) C_{u_1,u_2,\ldots u_k}.$$

## 4.3 Random starting vertices

Finally we consider the cover time of $k$ walks that start from vertices drawn from the stationary distribution. In this case, Theorem 3.8 loosely states that the ratio between the *hitting times* is at least $k$. Now let us show an upper bound on the ratio between the *cover time* of a single random walk and multiple random walks.

The intuition for the bound is quite similar to the intuition behind the proof of Theorem 4.4 (the proofs are quite a bit different). We will simulate $k$ random walks by a single walk. The single random walk will first run $\ln(k)\mathrm{mix}$ steps, getting to a vertex that is distributed almost according to the stationary distribution. The walk then simulates the first of the $k$ random walks. Next, the walk takes $\ln(k)\mathrm{mix}$ steps again and simulates the second random walk and so on until simulating the $k$th random walk. Since the start vertex of the $k$ simulated walks are *jointly* distributed almost as if they were independently sampled from the stationary distribution it seems that we should obtain the following upper bound: $C \leq k\mathbf{E}_{u_i} C_{u_1,u_2,\ldots u_k} + k\ln(k)\mathrm{mix}$, where $u_1, u_2, \ldots u_k$ are independently drawn from the stationary distribution. But as before we can not make this intuition formal, mainly because we do not know ahead of time how long the $k$ random walks will take until they cover the graph. We will instead prove the following bound which again may be weaker by at most a constant factor:

**Theorem 4.8.** *Let $G = (V, E)$ be any (strongly connected) graph. Let $u_1, u_2, \ldots u_k$ be drawn from the stationary distribution of $G$. Then:*

$$C \leq k\boldsymbol{E}_{u_i} C_{u_1,u_2,\ldots u_k} + O(k\ln(k)\mathrm{mix}) + O\left(k\sqrt{\boldsymbol{E} C_{u_1,u_2,\ldots u_k}\mathrm{mix}}\right).$$

Under some restrictions, the mixing time cannot be much larger than the maximal hitting time (see Theorem 2.12) and often will be much smaller. In such cases, Theorem 4.8 may be more informative than Theorem 4.4 in the sense that it implies a bound of roughly $k$ on the speedup as long as $k = \tilde{O}(\frac{C}{\mathrm{mix}})$ (rather than $k = O(\frac{C}{h_{\max}})$ as implied by Theorem 4.4). On the other hand, the starting vertices in Theorem 4.8 are according to the stationary distribution rather than arbitrary starting vertices as in Theorem 4.4.

*Proof of Theorem 4.8.* Recall that $\tau$ is the time it takes for a single walk to cover $G$ and $\tau_{u_1,u_2,\ldots u_k}$ is the time it takes for multiple walks to cover $G$. The following lemma will be the key in the proof of the theorem:

**Lemma 4.9.** *Let $s, t$ be any two non-negative integers and $u_1, u_2, \ldots u_k$ are drawn from the stationary distribution. Then:*

$$\Pr[\tau > kt + k \cdot s \cdot \mathrm{mix}] \leq \Pr[\tau_{u_1,u_2,\ldots u_k} > t] + k\boldsymbol{e}^{-s}.$$

*Proof.* Let $X$ be a single random walk and $\{X_\pi^i\}_{i=1}^k$ be $k$ independent random walks of length $t$ starting from the stationary distribution $\pi$. Let us define $\{Y_i\}_{i=1}^k$ to be $k$ sub-walks of $X$ as follows: The first walk $Y_1$ is a sub-walk of $X$ of length $t$ starting at time $s \cdot \mathrm{mix}$. The second sub-walk $Y_2$ starts at time $s \cdot \mathrm{mix}$ after the end of the walk $Y_1$ (and is of length $t$ as well). The last sub-walk $Y_k$ is of length $t$ and starts at time $s \cdot \mathrm{mix}$ after the end of the walk $Y_{k-1}$. Formally:

$$Y_i \triangleq \{X(i \cdot s \cdot \mathrm{mix} + (i-1)t + 1), X(i \cdot s \cdot \mathrm{mix} + (i-1)t + 2) \ldots X(i \cdot s \cdot \mathrm{mix} + (i-1)t + t)\}.$$

Note that $\{Y_i\}_{i=1}^k$ are included in a single random walk of length $kt + k \cdot s \cdot \mathrm{mix}$. Therefore:

$$\Pr[\tau > kt + k \cdot s \cdot \mathrm{mix}] \leq \Pr[\bigcup_{i=1}^k Y_i \neq V]. \tag{13}$$

We also know from the definition that:

$$\Pr[\tau_{u_1,u_2,\ldots u_k} > t] = \Pr[\bigcup_{i=1}^k X_\pi^i \neq V]. \tag{14}$$

Let $\mu, \nu$ be the distributions of $\{Y_i\}_{i=1}^k$ and $\{X_\pi^i\}_{i=1}^k$. Then from Equations (13) and (14), it follows that:

$$\Pr[\tau > kt + k \cdot s \cdot \mathrm{mix}] \leq \Pr[\tau_{u_1,u_2,\ldots u_k} > t] + \parallel \mu - \nu \parallel_1 .$$

Thus in order to prove this lemma it is enough to prove that the distribution of $\{Y_i\}_{i=1}^k$ is $k\mathbf{e}^{-s}$ close to the distribution of $\{X_\pi^i\}_{i=1}^k$. The proof will go by induction on $k$.

$\underline{k=1}$: The starting vertex of $Y_1$ is $X(s \cdot \mathrm{mix})$. From Fact 2.4 it follows that the distribution of $X(s \cdot \mathrm{mix})$ is $\mathbf{e}^{-s}$ close to the stationary distribution. The distributions of the walks $Y_1$ and $X_\pi^1$ are determined by their starting vertices. Therefore, since the distribution on the starting vertices are $\mathbf{e}^{-s}$ close, the distributions of $Y_1$ and $X_\pi^1$ are $\mathbf{e}^{-s}$ close as well.

$\underline{k \Rightarrow k+1}$: Let $\mu = (\mu_1, \mu_2), \nu = (\nu_1, \nu_2)$ be the distributions of $\{Y_i\}_{i=1}^{k+1}$ and $\{X_\pi^i\}_{i=1}^{k+1}$, where $\mu_1, \nu_1$ are the distributions on the first $k$ walks and $\mu_2, \nu_2$ are distributions on the last walk. From the induction assumption we know that $\parallel \mu_1 - \nu_1 \parallel_1 \leq k\mathbf{e}^{-s}$. Fix any value $x^{(k)}$ of the first $k$ walks. The walks $\{X_\pi^i\}_{i=1}^{k+1}$ are independent and so the distribution $\nu_{x^{(k)}} = \nu_2$ is simply the distribution $X_\pi$ (i.e., the distribution of a $t$-long walk starting at the stationary distribution). The distribution $\mu_{x^{(k)}}$ is the conditional distribution on $Y_{k+1}$ given $\{Y_i\}_{i=1}^k = x^{(k)}$. Since the walk $Y_{k+1}$ starts $s \cdot \mathrm{mix}$ steps after the walks $\{Y_i\}_{i=1}^k$ end, the distribution of $Y_{k+1}$ is $\mathbf{e}^{-s}$ close to the distribution of $X_\pi^{k+1}$ even conditioned on $\{Y_i\}_{i=1}^k = x^{(k)}$. Therefore $\parallel \mu_{(x^{(k)}} - \nu_{x^{(k)}} \parallel_1 \leq \mathbf{e}^{-s}$. From Lemma 2.14 it follows that $\parallel \mu - \nu \parallel_1 \leq k\mathbf{e}^{-s} + \mathbf{e}^{-s} = (k+1)\mathbf{e}^{-s}$, and the lemma follows. $\square$

From Lemma 4.9 we get the following inequality:

$$\Pr[\tau > kt + k \cdot s \cdot \mathrm{mix}] \leq \Pr[\tau_{u_1,u_2,\ldots u_k} > t] + k\mathbf{e}^{-\lfloor s \rfloor}$$

Since $-\lfloor s \rfloor \leq -s + 1$ it follows that:

$$\Pr[\tau > kt + k \cdot s \cdot \mathrm{mix}] \leq \Pr[\tau_{u_1,u_2,\ldots u_k} > t] + k\mathbf{e}^{-s+1}$$

Setting $s = \delta t + \log k + 1$ (so that $k\mathbf{e}^{-s+1} = \mathbf{e}^{-\delta t}$) and summing over $t$ we get:

$$\sum_{t=0}^{\infty} \Pr[\tau > tk(1+\delta\text{mix}) + (\log k + 1)\text{mix}k] \leq \sum_{t=0}^{\infty} \Pr[\tau_{u_1,u_2,\ldots u_k} > t] + \mathbf{e}^{-t\delta}$$

From Facts 2.6 it follows that:

$$\sum_{t=0}^{\infty} \Pr[\tau_{u_1,u_2,\ldots u_k} > t] = \mathbf{E}(C_{u_1,u_2,\ldots u_k})$$

Therefore,

$$\sum_{t=0}^{\infty} \Pr[\tau > tk(1+\delta\text{mix}) + (\log k + 1)\text{mix}k] \leq \mathbf{E}(C_{u_1,u_2,\ldots u_k}) + \mathbf{e}^{-t\delta}$$

In addition

$$\sum_{t=0}^{\infty} \mathbf{e}^{-t\delta} \leq \int_{t=0}^{\infty} \mathbf{e}^{-\delta t}\mathbf{d}t + 1 = \frac{1}{\delta} + 1,$$

Which implies

$$\sum_{t=0}^{\infty} \Pr[\tau > tk(1+\delta\text{mix}) + (\log k + 1)\text{mix}k] \leq \mathbf{E}(C_{u_1,u_2,\ldots u_k}) + \frac{1}{\delta} + 1. \tag{15}$$

From Fact 2.8 it follows that:

$$\sum_{t=0}^{\infty} \Pr[\tau > tk(1+\delta\text{mix}) + (\log k + 1)\text{mix}k] \geq \frac{C - (\log k + 1)\text{mix}k}{k(1+\delta\text{mix})} - 1.$$

Thus from Equation 15 it follows that:

$$\frac{C - (\log k + 1)\text{mix}k}{k(1+\delta\text{mix})} \leq \mathbf{E}(C_{u_1,u_2,\ldots u_k}) + \frac{1}{\delta} + 2$$

Equivalently, we can write it as:

$$C \leq k\mathbf{E}(C_{u_1,u_2,\ldots u_k}) + \delta\text{mix}k(\mathbf{E}(C_{u_1,u_2,\ldots u_k}) + 2) + \frac{k}{\delta} + 2k + (\log k + 2)\text{mix}k$$

Optimizing over $\delta$ we get $\delta = \frac{1}{\sqrt{\text{mix}\mathbf{E}C_{u_1,u_2,\ldots u_k}}}$. So:

$$C \leq k\mathbf{E}(C_{u_1,u_2,\ldots u_k}) + O\left(k\sqrt{\text{mix}\mathbf{E}C_{u_1,u_2,\ldots u_k}}\right) + O(\text{mix}k \log k)$$

□

We note that the proof of Theorem 4.8 also works if we consider the hitting times (rather than the cover times), implying the following theorem:

**Theorem 4.10.** *Let $G = (V, E)$ be any (strongly connected) graph. Let $u, v$ be any nodes of the graph and let $u_1, u_2, \ldots u_k$ be drawn from the stationary distribution of $G$. Then:*

$$h(u,v) \leq k\boldsymbol{E}_{u_i} h(\{u_1, u_2, \ldots u_k\}, v) + O(k \ln(k)\mathrm{mix}) + O\left(k\sqrt{\boldsymbol{E}_{u_i} h(\{u_1, u_2, \ldots u_k\}, v)\mathrm{mix}}\right).$$

As a corollary of Theorems 4.8 it follows that if $k \log k\mathrm{mix}$ is negligible relative to the cover time then the speedup of the cover time is at most $k$:

**Corollary 4.11.** *Let $G = (V, E)$ be any (strongly connected) graph. Let $u_1, u_2, \ldots u_k$ be drawn from the stationary distribution of $G$. Then if $k \log k = o(C/\mathrm{mix})$ then*

$$\frac{C}{\boldsymbol{E}_{u_i} C_{u_1,u_2,\ldots u_k}} \leq k + o(k)$$

.

*Proof.* From Theorem 4.8 we have that:

$$C \leq k\mathbf{E}(C_{u_1,u_2,\ldots u_k}) + O\left(k\sqrt{\mathrm{mix}\mathbf{E}C_{u_1,u_2,\ldots u_k}}\right) + O(\mathrm{mix}k \log k).$$

By the assumption $\mathrm{mix}k \log k = o(C)$. Thus we get:

$$C \leq k\mathbf{E}(C_{u_1,u_2,\ldots u_k}) + O\left(\sqrt{(k\mathrm{mix})(k\mathbf{E}C_{u_1,u_2,\ldots u_k})}\right) + o(C).$$

If $C \leq kC_{u_1,u_2,\ldots u_k}$ then we are done. So let us assume that $kC_{u_1,u_2,\ldots u_k} \leq C$. Then

$$C \leq k\mathbf{E}(C_{u_1,u_2,\ldots u_k}) + O(\sqrt{o(C)C}) + o(C).$$

Thus:

$$C \leq k\mathbf{E}_{u_i}(C_{u_1,u_2,\ldots u_k}) + o(C).$$

□

Similarly, from Theorem 4.10 we obtain the following corollary:

**Corollary 4.12.** *Let $G = (V, E)$ be any (strongly connected) graph. Let $u_1, u_2, \ldots u_k$ be drawn from the stationary distribution of $G$ and $u, v$ any nodes. Then if $k \log k = o(h(u,v)/\mathrm{mix})$ then*

$$\frac{h(u,v)}{\boldsymbol{E}_{u_i} h(\{u_1, u_2, \ldots u_k\}, v)} \leq k + o(k)$$

.

## 5 A New Relation Between Cover and Mixing Time

In this section we will show how we can use the results proven above in order to prove a new upper bound on the cover time in terms of mixing time. In order to do this we will need the following bound from [BKRU89].

**Theorem 5.1** (cf. [BKRU89] Theorem 1)**.** *Let $G$ be a connected undirected graph with $n$ vertices and $m$ edges. Let $u_1, u_2, \ldots u_k$ be drawn from the stationary distribution of $G$. Then:*

$$\boldsymbol{E}_{u_i}(C_{u_1,u_2,\ldots u_k}) \leq O(\frac{m^2 \log^3 n}{k^2}).$$

As a rather intriguing corollary of Theorem 5.1 and Theorem 4.8 we get the following bound on the cover time.

**Theorem 5.2.** *Let $G$ be a connected undirected graph with $n$ vertices and $m$ edges. Then:*

$$C \leq O(m\sqrt{\mathrm{mix}}\log^2 n).$$

*Proof.* From Theorem 4.8 it follows that:

$$C \leq k\mathbf{E}_{u_i} C_{u_1,u_2,\ldots u_k} + O(k\ln(k)\mathrm{mix}) + O\left(k\sqrt{\mathbf{E}C_{u_1,u_2,\ldots u_k}\mathrm{mix}}\right).$$

Thus from Theorem 5.1 we get the following bound on $C$:

$$C \leq O(\frac{m^2 \log^3 n}{k}) + O(k\ln(k)\mathrm{mix}) + O(m\log^{1.5} n\sqrt{\mathrm{mix}}).$$

As long as $k$ is at most polynomial in $n$ it follows that $\log k = O(\log n)$. Thus:

$$C \leq O(\frac{m^2 \log^3 n}{k}) + O(k\ln(n)\mathrm{mix}) + o(m\log^2 n\sqrt{\mathrm{mix}}).$$

Setting $k = \frac{m\log n}{\sqrt{\mathrm{mix}}}$ implies the theorem. □

# 6 Future research

This paper systematically studies the behavior of multiple random walks. While we have given various upper and lower bounds for the speedup of multiple random walks, there is still much more that we do not know on this topic, with a few examples being Open Problems 3.4, 3.13 and 4.7. In this section, we will discuss a few additional directions for further research.

Our knowledge on the hitting time of multiple random walks is more complete than our knowledge on their cover time. Indeed, analyzing the hitting time seems easier than analyzing the cover time. Designing new tools for analyzing the cover time of multiple random walks is an important challenge. For example, we have proved that the maximal hitting time of multiple random walks is obtained when all the walks start from the same vertex (see Theorem 3.3), but we don't know if the same is also true for the cover times:

**Open Problem 6.1.** *Prove or disprove that for any graph $G$*

$$\max_{u_1,u_2,\ldots u_k} C^k_{u_1,u_2,\ldots u_k} = \max_u C^k_{u,u,\ldots u}.$$

We have proved that in the case of worst starting vertices the speedup of the hitting time is at most $4k$, and we raised the question of whether the correct constant is one (see Open Problem 3.4). It seems however, that for the cover time the speedup may be larger than $k$ (though it is still possible that it is $O(k)$). Consider a walk on a "weighted" path $a - b - c$ with self loops such that the probability of staying in place is $1 - \frac{1}{x}$. In other words, consider a Markov chain $X(t)$ with the following transition probabilities:

$$\Pr[X(t) = b | X(t-1) = a] = \Pr[X(t) = b | X(t-1) = c] = \frac{1}{x}$$
$$\Pr[X(t) = c | X(t-1) = b] = \Pr[X(t) = a | X(t-1) = b] = \frac{1}{2x}$$

Calculating the cover times gives the following: The worst starting vertex of a single random walk is $b$ and the cover time is $5x + o(x)$. The worst starting vertices of $2$ random walks is when both walks start at $a$ and the cover time in such a case is $2.25x + o(x)$. Thus, in this case the speedup for $2$ walks is $2.222$. It is an interesting question to find stronger examples (where the speedup is larger than $k$), and of course it would be interesting to find a matching upper bound on the speedup.

A technical issue that comes up in our analysis is that in order to understand the behavior of multiple random walks it may be helpful to understand the behavior of short random walks. For example, what kind of bound can be obtained on $\Pr[\varsigma(u, v) \geq h_{\max}/2]$ (for an undirected and connected graph).

Finally, it will be interesting to explore additional applications of multiple random walks, either in computer science or in other fields.

## Acknowledgements

We thank Chen Avin, Uriel Feige, Gady Kozma, Zvi Lotker and Ofer Zeitouni for useful discussions. In particular, we are grateful to Gady for pointing out a flaw in one of our earlier arguments. We thank Rivka Efremenko for reviewing drafts of this paper.

## References

[AAK+07] N. Alon, C. Avin, M. Koucky, G. Kozma, Z. Lotker, and M. R. Tuttle. Many Random Walks Are Faster Than One. *ArXiv e-prints*, 705, May 2007. Manuscript available at http://arxiv.org/abs/0705.0467.

[AF99] David Aldous and Jim Fill. *Reversible Markov Chains and Random Walks on Graphs*. 1999. Manuscript available at http://www.stat.berkeley.edu/ aldous/RWG/book.html.

[AKL+79] Romas Aleliunas, Richard M. Karp, Richard J. Lipton, László Lovász, and Charles Rackoff. Random walks, universal traversal sequences, and the complexity of maze problems. In *20th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 218–223, San Juan, Puerto Rico, 29–31 October 1979. IEEE.

[AKS87] Miklós Ajtai, János Komlós, and E. Szemerédi. Deterministic simulation in LOGSPACE. In *Proceedings of the Nineteenth Annual ACM Symposium on Theory of Computing (STOC)*, pages 132–140, New York City, 25–27 May 1987.

[BB65] Edwin Beckenbach and Richard Bellman. *Inequalities*. Springer-Verlag, 1965.

[BF93] Greg Barnes and Uriel Feige. Short random walks on graphs. In *STOC*, pages 728–737, 1993.

[BK88] A.Z. Broder and A.R. Karlin. Bounds on the cover time. *Symposium on Foundations of Computer Science*, 0:479–487, 1988.

[BKRU89] Andrei Z. Broder, Anna R. Karlin, Prabhakar Raghavan, and Eli Upfal. Trading space for time in undirected s-t connectivity. In *STOC*, pages 543–549, 1989.

[CCR09] Alan Frieze Colin Cooper and Tomasz Radzik. Multiple random walks in random regular graphs. In *ICALP*, 2009.

[CRRS89] A. K. Chandra, P. Raghavan, W. L. Ruzzo, and R. Smolensky. The electrical resistance of a graph captures its commute and cover times. In *STOC '89: Proceedings of the twenty-first annual ACM symposium on Theory of computing*, pages 574–586, New York, NY, USA, 1989. ACM.

[ES09] Robert Elsässer and Thomas Sauerwald. Tight bounds for the cover time of multiple random walks. In *ICALP*, 2009.

[Fei95a] Uriel Feige. A tight lower bound on the cover time for random walks on graphs. *Random Struct. Algorithms*, 6(4):433–438, 1995.

[Fei95b] Uriel Feige. A tight upper bound on the cover time for random walks on graphs. *Random Struct. Algorithms*, 6(1):51–54, 1995.

[Fei97] Uriel Feige. A spectrum of time-space trade-offs for undirected s-t connectivity. *J. Comput. Syst. Sci.*, 54(2):305–316, 1997.

[HZ96] Shay Halperin and Uri Zwick. An optimal randomised logarithmic time connectivity algorithm for the erew pram. *J. Comput. Syst. Sci.*, 53(3):395–416, 1996.

[KNP99] David R. Karger, Noam Nisan, and Michal Parnas. Fast connected components algorithms for the EREW PRAM. *SIAM J. Comput.*, 28(3):1021–1034, 1999.

[Lov96] László Lovász. Random walks on graphs: A survey. *in: Combinatorics, Paul Erdos is Eighty*, 2:353–398, 1996.

[LWP] David A Levin, E. Wilmer, and Y. Peres. *Markov Chains and Mixing Times*. Oxford University Press.

[Mat88] Peter Matthews. Covering problems for markov chains. *The Annals of Probability*, 16(3):1215–1228, 1988.

[Sin92] Alistair Sinclair. Improved bounds for mixing rates of markov chains and multicommodity flow. *Combinatorics, Probability and Computing*, 1:351–370, 1992.

[Zei09] Ofer Zeitouni. private communication. 2009.