Bounds on the Mixing Time and Partial Cover of Ad-Hoc and Sensor Networks

UCLA CSD Technical Report 040028

Chen Avin and Gunes Ercal
Computer Science Department
University of California, Los Angeles
Los Angeles, CA 90095-1596, USA.
Email: {avin, ercal}@cs.ucla.edu

Spring 2004

Abstract

In [1], the authors proposed the partial cover of a random walk on a broadcast network to be used to gather information and supported their proposal with experimental results. In this paper, we demonstrate analytically that for sufficiently large broadcast radius, the partial cover of a random walk on a broadcast network is in fact efficient and generates a good distribution of the visited nodes. Our result is based on bounding the conductance, which intuitively measures the amount of bottlenecks in a graph. We show that the conductance of a random broadcast network is $\Theta(R)$, and this bound allows us to analyze properties of the random walk such as rapid mixing and load balancing. We find that for the random walk to be both efficient and have a high quality cover and partial cover, radius $R = O(1/poly(logN))$ is sufficient. Experimental results on the random unit disk graphs that resemble the conductance of the 3D grid indicate that the analytical bounds on efficiency, namely cover time and partial cover time, are not tight. In particular, $R = O(1/N^{1/3})$ is sufficient radius to obtain optimal cover time and partial cover time, if one is not concerned about the quality of the distribution of the visited nodes (for example in a query based on majority vote).

1 Introduction

The task of information collection and processing over wireless sensor networks is one of the main challenges in such domains [2] [3] [4]. The strict energy constraints [5] and the dynamics of the network (caused by node mobility, node failure, unreliable communication, etc.) prevent in many cases adopting traditional solutions from related areas like ad-hoc networks and distributed databases. Many such systems rely on state information stored at the nodes and global

data structures for proper operation (for example, pointers to cluster heads, routing tables, and spanning trees). As such, those types of systems have critical points of failure and in dynamic environments must adopt failure recovery mechanisms which impact significantly on the overall performance.

In recent years emphasis has been shifting from deterministic algorithms to randomized algorithms which are often simpler and fast. In particular, we can consider the simple random walk for information collection in a sensor network due to the simplicity of the process, robustness to failures (no critical points of failure), minimal overhead, and locality of computation. This approach is gaining credence and popularity, as recently several authors [1, 6, 7] have considered random walk approaches in various networking settings.

The authors [1] proposed the *partial cover* of a random walk on a broadcast network to be used to gather and process information. As the *cover time* [8] is the expected number of steps taken by a random walk to visit every node in the network, the *partial cover time* (PCT) is the expected time required to cover only a constant fraction of the network. In sensor network applications, for most tasks, it is not necessary to consult every node in the network. In fact, the authors had shown, derived from the well known Matthews bound [9], that the *partial cover time* is a factor of $O(\log N)$ times more efficient than the *cover time*. Substantiated by experimental results we showed that the partial cover of 80% of the network is in fact efficient in the number of messages in comparison to other systems and still maintains all the other nice properties of random walks mention above.

Kempe *et al.* [6] proposed and analyzed parallel random walk techniques for gossiping in networks. Their approach differs from [1] approach in the nature of what is expected of the data collection task. The former is concerned with disseminating global information throughout the network to be stored in every node (all-to-all), whereas the latter is concerned with answering a query regarding the global status of the network by the random walk of a single token (all-to-one). While the [6] approach is much more time-efficient in comparison to the [1] approach, it is energy-inefficient for use in a sensor network or broadcast network due to a number of factors: the total number of messages sent, the correspondingly large associated interference, and the girth of information stored at each node. Therefore, it is necessary to consider the particular nature of the network and the data collection task.

Since then, Gkantsidis *et al.* [7] have proposed random walk techniques for peer-to-peer networks based on that random walk are especially efficient when the underlying topology of the network is an expander [10]. As such, they have advocated expander construction in peer-to-peer network design and use random walks to perform search. One nice property about expanders is that they give optimal rate of convergence to stationary distribution, namely mixing rate, due to the smallness of the second largest eigenvalue λ_2 of the weighted transition matrix associated with the underlying network topology. As shall be discussed, λ_2 is essentially related to other relevant and desirable properties of random walks including partial cover time and load balancing.

In this paper, as we are concerned with all-to-one data collection in a sensor

or broadcast network, we expand upon our previous work [1]. In particular we investigate the efficiency and quality of the random walk as a function of the broadcast radius and size of the network. Delineating such a relationship is beneficial for network designers to either set broadcast radius accordingly or to otherwise know what random walk properties are to be expected given their network's broadcast radius and size.

2 Overview of the Models, Methods, and Results

A well-known model representing sensor networks (or in general broadcast networks) and their underlying Markov chain is the class of *unit disk graphs* (UDGs) [11]. Let $G_R(N)$ be a random UDG constructed as follows: Place $N = |V|$ nodes uniformly at random in unit a squared area and then connect every pair of nodes at Euclidean distance less than or equal to R. (Although we can think of UDGs in any number of dimensions, in this work we assume only the two dimensional case as that has direct relevance to ad-hoc networks.)

Note that, on the one hand, a $2D$ grid network is a UDG of borderline radius to ensure connectivity. And, on the other hand, the complete graph K_N is a UDG with maximal density and connectivity. As we shall see, the $2D$ grid lacks many desirable random walk properties (such as rapid mixing and optimal cover time), whereas the complete graph behaves optimally with respect to such random walk properties. It is easy to see that by increasing R of a UDG we are increasing the connectivity and the degree of the nodes and can shift from the grid to the complete graph. Then, intuitively, a question to ask is, do we need to increase R to *maximum* (i.e $\sqrt{2}$) in order get these properties, or, rather, is there a more continuous relationship between R and these properties? This question was a primary motivation in investigating the relationship between the radius and mixing time of a UDG.

As interference and energy grow with increased radius, in particular, we wonder what are small radii for which we may yet obtain good, even optimal, random walk properties. Of course, such bounds on the radii depend on the particular properties desired, and the particular properties desired are related to the following important and relevant questions to consider for a random walk approach in the context of global data collection over broadcast networks:

1. How long should we wait for the random walk token to collect information from all of the nodes?

2. How long should we wait for the random walk token to collect information from a constant fraction c of the nodes?

3. Does the random walk have good load balance properties?

4. What is the quality of a random walk (measured by the distribution of visited nodes) after some t number of steps?

5. What is the quality of a the partial cover after having collected information from a constant fraction c of the nodes?

Question 3 exactly concerns the *mixing rate*, the rate of convergence of a random walk to stationary distribution, of the random walk over $G_R(N)$. For regular graphs, the more rapidly mixing the random walk, the more uniformly the nodes are sampled, and thus the more balanced is the load. Intuitively, the fewer the bottlenecks of the UDG, the more rapidly mixing the random walk is. *Conductance* is a measure of bottlenecks that is particularly appealing geometrically and which is directly related to the second eigenvalue λ_2 of the weighted transition matrix [12]. We analytically demonstrate that for sufficiently large N as a function of R, the conductance of $G_R(N)$ is $\Phi(G_R(N)) = \Theta(R)$ with high probability[1]. We also give a useful continuous approximation to $\Phi(G_R(N))$. Based on the conductance results, we show that for a $G_R(N)$ the radius sufficient for rapid mixing is $R = O(1/poly(logN))$. The idea of using conductance to bound the mixing time of broadcast networks was recently independently proposed in yet unpublished manuscript [13].

Questions 1 and 2 exactly concern the cover time C and, respectively, partial cover time PCT of $G_R(N)$. Since the random walk should be run C steps or PCT steps to be expected to have collected information from all or almost all of the nodes, we investigate for which broadcast radii R we obtain small C or PCT. Unfortunately, from analytical results based on the conductance we are only guaranteed optimal C and PCT for constant radius R. However, by observing that the k dimensional mesh for $k > 2$ also has optimal C and PCT though it is not even rapid mixing, we hypothesized that perhaps the $G_R(N)$ with conductance similar to that of the 3D mesh may also has optimal C and PCT. Since conductance alone is an insufficient measure to analytically tightly bound C and PCT, our hypothesis was also based on the observation that, in the case of UDGs, increase in conductance directly implies increase in average degree and decrease in the hop-diameter of the network as well. Therefore, we formalize the notion of *resemblance* of UDGs to other graphs based on the similarity in conductance, which immediately implies some other comparisons as well. In fact, our experimental results show that the $G_R(N)$ that resembles the 3D mesh has C and PCT even lower than that of the 3D mesh. Therefore, by experimental results we have that radius $R = O(1/N^{1/3})$ is sufficient for optimal C and PCT. It should also be noted that experimental results indicate that the $G_R(N)$ that resembles the hypercube has slightly better mixing rate and PCT than that of the hypercube. This further demonstrates the usefulness of the notion of resemblance as much is already known about certain graphs such as the grids and hypercube.

For the last two questions, we measure the quality of our random walk in terms of how small a *hole*, namely a contiguous unvisited area, remains after a given number of steps. Such a characterization is eminently sensible due to the geometric characterization of UDGs. We approximate the hole size by its upper bound as the maximum minimum Euclidean distance from unvisited nodes to

[1] meaning with probability $1 - O(\frac{1}{N})$

visited nodes after some number of steps. And, so the last two questions exactly concern how small a hole remains after some amount of time. Our experimental results indicate that the quality of the random walk is significantly improved by rapid mixing. This intuitively makes sense as the quality of a random walk is related to load-balancing properties. Thus, although for efficiency radius $R = O(1/N^{1/3})$ suffices, to achieve both efficiency and high quality geometric distribution of visited nodes, radius $R = O(1/poly(logN)$ is recommended.

Note that a high quality partial cover may not be necessary depending upon the nature of the data collection task. For example, to collect a majority vote, with very high probability after 90% PCT, the majority vote will have been obtained regardless of the distribution quality. However, for a temperature sensing system, for example, the quality of the random walk may be eminently important.

3 Preliminaries

3.1 Markov chains and the Simple Random Walk

The probabilistic rules by which a random walk operates is defined by the corresponding *Markov chain*. Let $\mathfrak{M}$ be a Markov chain over state space Ω and probability transition matrix K (i.e $K(x, y)$ is the probability to move from x at time t to y at time $t + 1$). Let P_t denote the probability distribution of $\mathfrak{M}$ at time t. Given some starting distribution P_0, one obtains the next step distribution P_1 by applying K to P_0, then obtains the subsequent distribution P_2 by applying K to P_1 and so forth. Therefore, for all time $t > 0$,

$$P_t = P_0 K^{t-1} \tag{1}$$

In such terms, the stationary distribution of $\mathfrak{M}$, if such exists, is then defined as the unique probability vector π such that

$$\pi K = \pi \tag{2}$$

A primary motivation in considering a random walk approach as opposed to a deterministic protocol is simplicity and locality of computation. So, if the random walk is currently at node q, then the simplest probabilistic rule by which to choose the next node is simply to choose a node uniformly at random from among the set of neighbors of q. We call the Markov chain $\mathfrak{N} = (\Omega, K)$ corresponding to such a random walk the *simple random walk*. Note that we may just as well define such $\mathfrak{N}$ by its underlying graph $G = (V, E)$. For such G, for any node $v \in V$, let $\delta(v)$ denote the degree of v, that is the number of neighbors of v in G and let $K(v, u) = \frac{1}{\delta(v)}$.

The simple random walk $\mathfrak{N} = (\Omega, K)$ over some connected graph $G = (V, E)$ has a stationary distribution π such that, for any node $q \in V$ [14],

$$\pi(q) = \frac{\delta(q)}{2m} \tag{3}$$

where $m = |E|$. Further, when the underlying graph G is regular, that is when there is d such that for all q in $\mathfrak{N}$, $\delta(q) = d$, the stationary distribution is the uniform distribution [14]

$$\pi(q) = \frac{d}{2m} = \frac{1}{N} \ \forall q \in \Omega \tag{4}$$

where $N = |\Omega| = |V|$.

At stationary distribution, it is clear that the random walk has optimal load-balancing qualities for regular graphs G. Similarly, it is clear that the faster the random walk on a regular graph converges to stationarity, the greater its load-balancing qualities and better the quality of the partial cover as mentioned regarding question (4).

3.2 Mixing Time and the Spectral Gap $(1 - \lambda_2)$

The efficiency with which a random walk of $\mathfrak{R}$ may be used to sample over state space Ω with respect to stationary distribution π is precisely given by the rate at which P_t converges to π. In order to speak of convergence of probabilities, one must have a notion of distance over time. The *variation distance* at time t with respect to the initial state x is defined to be [15]

$$\Delta_x(t) = \max_{S \subseteq \Omega} |K^t(x, S) - \pi(S)| = \frac{1}{2} \sum_{y \in \Omega} |K^t(x, y) - \pi(y)| \tag{5}$$

The rate of convergence of $\mathfrak{R}$ to stationarity may be measured by the *mixing time*, the function [15]

$$\tau_x(\epsilon) = min\{t \mid \Delta_x(t') \leq \epsilon, \forall t' \geq t\} \tag{6}$$

A chain $\mathfrak{M}$ is considered *rapid mixing* iff $\tau_x(\epsilon)$ is $O(poly(log(N/\epsilon)))$. For $\mathfrak{M}$ to be used for efficient sampling, we want that $\mathfrak{M}$ is rapid mixing.

As every step t of a random walk on $\mathfrak{M}$ is represented by the matrix product of P_0 and K^t, and from (2) above the stationary distribution π is an eigenvector corresponding to largest eigenvalue 1 of K, plausibly spectral theory is very useful in analyzing $\tau_x(\epsilon)$. Let $1 = \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N > -1$ denote the eigenvalues of K in decreasing order for ergodic [2], reversible [3] Markov chain $\mathfrak{M}$. It is well-known that a large value of the *spectral gap* $(1 - \lambda_2)$ exactly captures rapid convergence to stationarity [15]:

Proposition 1 *For an ergodic Markov chain, the quantity $\tau_x(\epsilon)$ satisfies*

(i) $\tau_x(\epsilon) \leq (1 - \lambda_2)^{-1}(\ln \pi(x)^{-1} + \ln \epsilon^{-1})$

(ii) $\max_{x \in \Omega}(\epsilon) \geq \frac{1}{2}\lambda_2(1 - \lambda_2)^{-1} \ln(2\epsilon)^{-1}$

[2] meaning that a stationary distribution exists

[3] The simple random walk is always reversible (i.e $\pi(x)K(x, y) = \pi(y)K(y, x)$)

As we would like the starting state of a random walk to be arbitrary, the above implies that a large spectral gap $(1-\lambda_2)$ is both a necessary and sufficient condition for rapid mixing. A well-known method for bounding λ_2 to prove rapid mixing when the underlying graph has a geometric interpretation is a *Conductance argument* [12]. This is the method we shall use, as UDGs have a strong geometric interpretation.

3.3 Conductance

Intuitively, one would expect that when the Markov chain $\mathfrak{M}$ doesn't have *bottlenecks*, the lower the probability of getting stuck in any particular set of states, and thus the more rapidly mixing $\mathfrak{M}$ is. The property of "no bottlenecks" is formalized in a continuous manner with the notion of *conductance*.

The *conductance* of a reversible Markov chain $\mathfrak{M}$ is defined by

$$\Phi = \Phi(\mathfrak{M}) = \min_{S \subset \Omega, 0 < \pi(S) \leq 1/2} \frac{Q(S, \overline{S})}{\pi(S)} \tag{7}$$

where $\bar{S} = \Omega - S$, $Q(x,y) = \pi(x)K(x,y) = \pi(y)K(y,x)$, $\pi(S)$ is the probability density of S under the stationary distribution π, and $Q(S, \overline{S})$ is the sum of $Q(x,y)$ over all $(x,y) \in S \times (\Omega - S)$.

In graph-theoretic terms, the conductance of $\mathfrak{M}$ is the minimum over all subsets $S \subset \Omega$ of the ratio of the weighted flow across the cut $Cut(S, \overline{S})$ to the weighted capacity of S. Intuitively, the higher the conductance of $\mathfrak{M}$, there are fewer bottlenecks in $\mathfrak{M}$, and the more rapidly mixing $\mathfrak{M}$ is. This intuition is confirmed by the following theorem [15]:

Theorem 1 *The second eigenvalue λ_2 of a reversible Markov chain $\mathfrak{M}$ satisfies*

$$1 - 2\Phi \leq \lambda_2 \leq 1 - \frac{\Phi^2}{2} \tag{8}$$

The above Theorem along with Proposition (1) yield the following powerful corollary bounding the mixing time via conductance [12]:

Corollary 1 *Let $\mathfrak{M}$ be a finite, reversible, ergodic and aperiodic* [4] *Markov chain, and let Φ be the conductance of $\mathfrak{M}$. Then, for any initial state x, the mixing time of $\mathfrak{M}$ satisfies*

$$\tau_x(\epsilon) \leq 2\Phi^{-2}(\ln \pi(x)^{-1} + \ln \epsilon^{-1}) \tag{9}$$

[4] Any chain can be made aperiodic simply by adding loop probabilities of $\frac{1}{2}$ at each node, this clearly does not affect the stationary diistribution. We assume that our random UDGs are aperiodic.

4 Mixing Time of UDGs

4.1 Bounding the Conductance of the k dimensional Grid

To begin with a simple example of a conductance argument with similarities to the conductance argument for general UDGs, we consider the case of the 2 dimensional grid which is a sub-class of the class of regular UDGs.

Let $M(2, N)$ denote the two dimensional grid of N nodes. It is easy to see that, since the graph has a regular geometric structure, the minimum conductance occurs when we consider $min(|S|, |\overline{S}|)$ of maximum capacity, that is when $\pi(S) = \pi(\overline{S}) = \frac{1}{2}$ so that S has half of the nodes of $M(2, N)$. Furthermore, as there are many possible ways of separating the nodes of $M(2, N)$ into two halves S and $\overline{S}$, we need to consider the separation that gives the minimum flow across $Cut(S, \overline{S})$, which occurs when the length of the boundary between S and $\overline{S}$ is minimized (since every edge has the same weight due to regularity). It is easy to check that the separation satisfying this is with a separating line l parallel to one of the axis. Since there are $N^{\frac{1}{2}}$ edges crossing such a cut and each edge has weight $w = \frac{1}{4}$ [5], the conductance of the two dimensional grid of N nodes is

$$\begin{aligned} \Phi(M(2,N)) &= 2\sum_{\substack{x\in S\\ y\in \bar{S}}} \frac{1}{N}\frac{1}{4} = 2N^{\frac{1}{2}}\frac{1}{4N} \\ &= (2N^{\frac{1}{2}})^{-1} \end{aligned}$$

This argument easily generalizes to the k dimensional grid $M(k, N)$, and we obtain the following by Theorem 1 and Corollary 1 above:

Lemma 2 *For the k dimensional grid $M(k, N)$ of N nodes we have the following:*

1. $\Phi(M(k, N)) = (kN^{\frac{1}{k}})^{-1}$
2. $\frac{1}{2}(kN^{\frac{1}{k}})^{-2} \leq 1 - \lambda_2 \leq 2(kN^{\frac{1}{k}})^{-1}$
3. $\tau_x(\epsilon) \leq 2k^2 N^{\frac{2}{k}}(\ln N + \ln \epsilon^{-1})$

4.2 Bounding the Conductance of $G_R(N)$

Let $G_R(N)$ be a UDG constructed as mention earlier. Our main analytical results are the following:

Theorem 2 *If $N \geq 4R^{-2}\ln(4R^{-2})$, then with high probability*

$$\Phi(G_R(N)) = \Theta(R) \tag{10}$$

[5] We ignore the two nodes on the borders which have only 3 neighbors.

From Theorem 2, Theorem 1 and Corollary 1 we obtain these bounds:

Corollary 3 *If $N \geq 4R^{-2}\ln(4R^{-2})$, then with high probability we have for $G_R(N)$ the following :*

1. $1 - \lambda_2 = \Omega(R^2)$

2. $1 - \lambda_2 = O(R)$

3. $\tau_x(\epsilon) = O(R^{-2}(\ln N + \ln \epsilon^{-1}))$

Note immediately from above that $G_R(N)$ is rapidly mixing if $R^{-1} = O(poly(\log N))$.

Before we begin the proof of Theorem 2, note the following well-known and very useful result from literature [10]:

Lemma 4 (Balls in Bins) *If one throws $N \geq B\log B$ balls randomly into B bins, then both the minimum and the maximum number of balls in any bin is $\Theta(\frac{N}{B})$ with high probability (meaning with probability $1 - O(\frac{1}{N})$).*

From Lemma 4, we may prove the following about $G_R(N)$:

Lemma 5 *If $N \geq 2R^{-2}\ln(2R^{-2})$, then with high probability $\delta(v) = \Theta(NR^2), \forall v \in V$.*

Proof 1 (of Lemma 5) *First note that $N \geq 2R^{-2}\ln(2R^{-2})$, guarantees from Lemma 4 that if we divide the area into $2R^{-2}$ bins of size $\frac{R}{\sqrt{2}} \times \frac{R}{\sqrt{2}}$, the number of nodes in every bin will be $\Theta(N/R^{-2})$ with high probability. (Increasing N tightens the bound.) Since, for every bin, the set of nodes in the bin forms a clique, and every node $v \in V$ is in some bin, we have that $\delta(v) = \Omega(NR^2), \forall v \in V$. Furthermore, as every node may be connected to the nodes of at most nine bins (that is its own bin and the bordering bins), we have that $\delta(v) = \Theta(NR^2), \forall v \in V$.*

Since $N = 2R^{-2}\ln(2R^{-2}))$ satisfies the connectivity requirements from [16], and obviously increasing N will not disconnect the network, it can be shown that $G_R(N)$ is connected with high probability. In other words, with high probability, $G_R(N)$ is connected, almost regular, and there are no large areas in the graph with low density. Now we may begin the proof of our main result:

Proof 2 (of Theorem 2) *Let $Cut(S, \bar{S})$ denote the cut size between S and $\bar{S}$ (the total number of edges crossing from S to $\bar{S}$). Due to regularity and uniformity with high probability by Lemma 5, we can adopt the reasoning from the conductance argument for the grid $M(2, N)$ given in the previous subsection,*

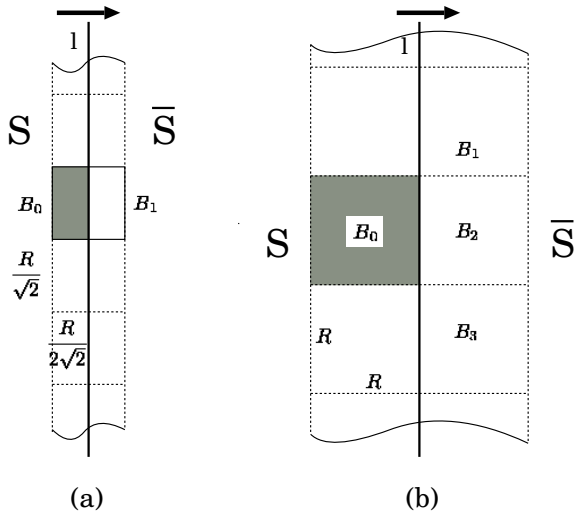

Figure 1: (a) Lower bound for the Conductance in UDG. (b) upper bound for the Conductance in UDG

and we can clearly observe that the minimum conductance is when we divide the area into two halves S and $\bar{S}$ as in Fig. 6.

First we prove the lower bound. Partition the area into bins of size $\frac{R}{2\sqrt{2}} \times \frac{R}{\sqrt{2}}$ as in Figure 1 (a). Since $N \geq 4R^{-2}\ln(4R^{-2})$, from Lemma 4, with high probability the number of nodes in any bin is $N\frac{R^2}{4}$. Notice that the set of nodes in any two horizontally adjacent bins forms a clique. Therefore, to lower bound $Cut(S,\bar{S})$, we are only considering the crossing edges within each separate such clique along the dividing line l. Since there are at least $\frac{\sqrt{2}}{R}$ cliques along the dividing line l, and for each bin on the left side of l we have $\Omega(NR^2)$ such edges crossing to the right of l, we obtain the desired lower bound $\Omega(R^3N^2)$ with high probability.

And, now we prove the upper bound. Partition the area into bins of size $R \times R$ as in Figure 1 (b). Note that for each edge (u,v) crossing l, v must be in some left bin B_0 adjacent to l, and so u must be in one of three possible bins B_1, B_2, B_3 that are on the right of l and touching B_0 as shown in the picture. To upper bound $Cut(S,\bar{S})$, we consider the maximum number of crossing edges from any $R \times R$ sized bin B_0 in S to three $R \times R$ sized bins $B_1, B_2, and B_3$ in $\bar{S}$. As there are $\frac{1}{R}$ such bins as B_0, and from Lemma 4, with high probability the number of nodes in any bin is NR^2, we get the desired upper bound as follows:

$$Cut(S,\bar{S}) = O(\frac{1}{R} \cdot NR^2 \cdot 3NR^2) = O(R^3N^2) \tag{11}$$

with high probability.

So, combining the upper and lower bounds, we have that with high probability,

$$Cut(S, \bar{S}) = \Theta(R^3 N^2) \tag{12}$$

And, thus, by $\delta(v) = \Theta(NR^2), \forall v \in V$ *we complete the proof:*

$$\begin{aligned}
\Phi(G_R(N)) &= 2Q(x,y) = \sum_{\substack{x \in S \\ y \in \bar{S}}} \pi(x) K(x,y) \\
&= 2 \sum_{\substack{x \in S \\ y \in \bar{S}}} \Theta(\frac{1}{N}) \Theta(\frac{1}{R^2 N}) \\
&= 2\Theta(R^3 N^2) / \Theta(R^2 N^2) \\
&= \Theta(R)
\end{aligned}$$

5 Cover Time and Partial Cover Time of UDGs

Known results for the cover time of specific graphs vary from the best case of $O(N \log N)$ to the worst case of $O(N^3)$. The best cases correspond to dense, highly connected graphs, for example, the complete graph, d-regular graphs with $d > \frac{N}{2}$, and the hypercube. When connectivity decreases and bottlenecks exist in the graph, the cover time increases, as exemplified by the bar-bell graph and, to a lesser degree, by the line graph which has cover time $\Theta(N^2)$.

Therefore, intuitively, one would anticipate a relationship between the spectral gap $(1-\lambda_2)$ and small cover time. In confirmation of this intuition, a bound for the cover time for regular graphs G that is based on the spectral gap $(1-\lambda_2)$ is given by [17] and [18]:

Theorem 3 *For regular graph* $G = (V, E)$ *with* $N = |V|$ *and second largest eigenvalue* λ_2 *the cover time of* G *is bounded as follows:* $C(G) = O(N \log N/(1-\lambda_2))$

From the same analysis of [18] and [17] one may directly obtain the partial cover time of a regular graph via a partial balls in bins argument:

Lemma 6 *For regular graph* $G = (V, E)$ *with* $N = |V|$ *and second largest eigenvalue* λ_2 *the partial cover time of* G *is bounded as follows for any constant* c *such that* $0 < c < 1$*:* $PCT_c(G) = O(N/(1-\lambda_2))$

Therefore, it follows from Theorem 3, Lemma 6, and Theorem 2 that we can bound the cover time and PCT of $G_R(N)$ as follows:

Corollary 7 *If* $N \geq 2R^{-2} \ln(2R^{-2})$*, then with high probability*

$$C(G_R(N)) = O(R^{-2} N \log N) \tag{13}$$

and, for any constant c*,* $0 \leq c < 1$*,*

$$PCT_c(G_R(N)) = O(R^{-2} N) \tag{14}$$

If these bounds on cover time and PCT were tight, then the only way to achieve optimal cover time for UDGs would be by choosing a radius R that is constant irrespective of the network size N. Recalling that our definition of $G_R(N)$ is normalized to a unit area, this would mean that only broadcast networks of constant hop diameter may have optimal cover and partial cover. Even the radius required for rapid mixing, which is $R = O(1/poly(\log N))$, is several orders lower than such a radius.

However, fortunately, the bounds given by Theorem 3, Lemma 6, and correspondingly by Corollary 7 are not tight. An especially demonstrative case of this is for k dimensional grid $M(k, N)$ of N nodes for any $k > 2$. The cover time and PCT for $M(k, N)$ are known to be optimal, that is $\Theta(N \log N)$ and $\Theta(N)$ respectively [19]. Yet, recalling Lemma 2, Theorem 3 and Lemma 6 yield only that $C(M(k, N)) = O(k^2 N^{1+\frac{2}{k}} \log N)$ and that $PCT_c(M(k, N)) = O(k^2 N^{1+\frac{2}{k}})$.

The following theorem provides, in many cases, tighter bounds on the cover time C in term of the maximum hitting time h_{max}. For arbitrary nodes $i, j \in V$, let h_{ij} be the expected number of steps for the random walk to move from i to j, namely the *hitting time* between i and j. Then h_{max} (h_{min}) is defined as the maximum (minimum) h_{ij} over all ordered pairs of nodes.

Theorem 4 (Matthews' Theorem [9]) *For any graph G,*

$$h_{min} \cdot H_N \;\leq\; C(G) \;\leq\; h_{max} \cdot H_N$$

where $H_k = ln(k) + \Theta(1)$ is the k-th harmonic number.

Similarly, in [1] the authors showed the following bound of the partial cover time, PCT, in terms of the hitting time:

Lemma 8 *For any graph G, and $0 \leq c \leq \frac{N-1}{N}$*

$$PCT_c(G) \;<\; 2 \cdot h_{max} \cdot \left\lceil \log_2(\frac{1}{1-c}) \right\rceil \;=\; O(h_{max})$$

In particular, as the maximum hitting time for $M(k, N)$ is known to be $O(N)$ for $k > 2$ [19], Matthew's Theorem and Lemma 8 yield that the cover time and PCT of $M(k, N)$ is $O(N \log N)$ and $O(N)$, respectively, which are tight. In future work, we plan to investigate the maximum hitting time of $G_R(N)$ to obtain tighter analytical bounds on the cover time and PCT of UDGs based on Matthew's Theorem and Lemma 8. In this work, we have investigated the cover time and PCT of UDGs via experiments.

Notice that whether one bounds the cover time or partial cover time based on the hitting time or based on the spectral gap, an $O(\log N)$ efficiency is gained in moving from the cover time to the partial cover time. In fact, as can be observed from the case of the k dimensional grid $M(k, N)$, the $O(\log N)$ gain in efficiency is often tight. Experimental results comparing cover times and PCTs of UDGs from [1] also support this tightness for the case of UDGs, further justifying consideration of the partial cover for information collection in broadcast networks.

6 Resemblance of UDGs

We are concerned with small radii (as a function of N) for which optimal cover time and PCT may yet be achieved. From above discussion we have seen that connectivity and the spectral gap are closely related to the cover time and PCT, and we expect that, for fixed N, as R increases, the cover time and PCT of $G_R(N)$ decrease. We know that $M(3, N)$ has optimal cover time and PCT even though it has constant degree and is not even rapidly mixing. So, we wonder for which radius R would the cover time and PCT of $G_R(N)$ be on the same order as the cover time and PCT, respectively, of $M(3, N)$. As conductance is a measure of connectivity and spectral gap, and $\Phi(G_R(N)) = \Theta(R)$ from Theorem 2, we ask: If $R = \Phi(M(3, N))$ does $G_R(N)$ also have optimal cover time and PCT?

To formalize the implicit notion of resemblance of graphs underlying that question, let us define the following: A UDG $G_R(N)$ *resembles* another graph $G = (V, E)$, with $N = |V|$, iff $R = \Theta(\Phi(G))$. And, as abbreviation, we may write, $G_R(N) \equiv_\Phi G$ to mean that $G_R(N)$ resembles G.

As the spectral gap and mixing time may both be upper bounded and lower bounded by functions of conductance, it is easy to see that two graphs which resemble each other will also have spectral gap and mixing time within close range of each other. For example, if $G_R(N) \equiv_\Phi G$, then $G_R(N)$ is rapidly mixing iff G is rapidly mixing. However, precisely because our notion of resemblance is based essentially on spectral gap, which in general is not sufficient to characterize the cover time and PCT due the non-tightness of Theorem 3 and Lemma 6 for the case of $M(k, N)$ for $k > 2$, we do not necessarily expect that *any* two graphs with the same conductance also have the same cover time or PCT.

But, in the case of UDGs, other special properties improve with increased radius. In particular, it is easy to see that both the average degree δ and the hop diameter D of a UDG are functions of the radius. Specifically, it is easy to see the following:

Corollary 9 *If $G_R(N) \equiv_\Phi M(k, N)$ for some $k \geq 2$ then*

1. $R = \Theta(N^{-\frac{1}{k}})$
2. $D(G_R(N)) = \Theta(N^{\frac{1}{k}}) = D(M(k, N))$
3. $\forall v \in V, \delta(v) = \Theta(N^{\frac{k-2}{k}})$

So, when $G_R(N) \equiv_\Phi M(k, N)$, for any constant $k \geq 2$: Not only do the two graphs have similar conductance and diameter, but also the average degree of the nodes of $G_R(N)$ is orders higher than the degree of any node of $M(k, N)$ (which is just $d = 2k$). This motivates our hypothesis above, paraphrased in terms of resemblance:

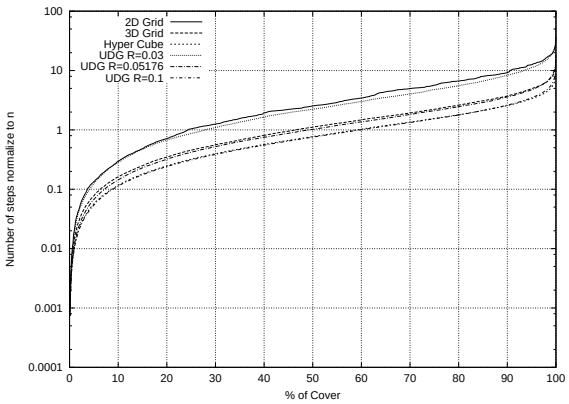

Figure 2: The progress of partial cover time as function of number of steps normalized to N for different graphs

Table 1: Number of steps required for 80% cover of 3D grid and Corresponding UDG

N	R	UDG(R)	3D Grid
1000	0.0859	2.7142	2.7016
1728	0.0704	2.5828	2.6947
2744	0.05965	2.5697	2.7090
4096	0.5176	2.4693	2.6874

Hypothesis 1 *If* $G_R(N) \equiv_\Phi M(k,N)$ *for some* $k > 2$ *then, for any constant* $0 \leq c < 1$*,*

1. $C(G_R(N)) = O(C(M(k,N))) = \Theta(N \log N)$

2. $PCT_c(G_R(N)) = O(PCT_c(M(k,N))) = \Theta(N)$

In particular, this would imply that the radius required for optimal cover time and PCT is $O(N^{-\frac{1}{3}})$, which is significantly lower than $\Theta(1/poly(\log N))$, the radius required for rapid mixing. In the next section, we demonstrate experimental results in support of this hypothesis.

7 Experimental Results

In this section we validate our analytical results and hypotheses using simulations of simple random walks on different graphs. Our random UDGs were constructed by placing N random nodes in a unit area and connecting every

two nodes at distance less than or equal to R. For each experiment we took the average of 100 runs. Unless otherwise specified, the graphs are of size $N = 4096$ nodes. In our experiments based on resemblance, we used the continuous approximation to the conductance as given in the Appendix.

7.1 Cover Time and PCT

Fig. 2 represents the number of steps (time) normalized to N (on log scale) as a function of the progress of the partial cover. The Figure presents three different well known graphs: $2D$-grid $M(2, N)$, $3D$-grid $M(3, N)$, and the hypercube $H(N)$, and three different random $G_R(N)'s$ with $R = 0.03, 0.05176, 0.1$ which closely achieve the same partial cover as the respective known graphs. This validates our intuition that we can achieve any cover time up to an optimal one by increasing R and legitimizes the notion of resemblance for UDGs.

Importantly, note that $R = 0.05176$ corresponds to the radius of the $G_R(4096)$ that resembles $M(3, 4096)$, obtained using the approximation to conductance in equation (15). One may observe from Fig. 2 that the PCT and cover time of the $G_R(4096)$ with $R = 0.05176$ behave very similarly to the PCT and cover time of $M(3, 4096)$. One may further observe from Table 1, that for varying network size N, the 80% PCT of $M(3, N)$ and the $G_R(N)$ that resembles it has very similar PCT. Therefore, Fig. 2 and Table 1 support our Hypothesis 1 and demonstrate that the PCT for the $G_R(N)$ that resembles $M(3, N)$ is almost the same as the PCT of $M(3, N)$, which is optimal, namely $O(N)$ instead of $O(N^{\frac{2}{3}}N)$ which follow Corollary 9 and Corollary 7

In Fig. 2, one may directly observe the sharp increase in the number of steps (time) for every graph as the partial cover approaches the full cover. This confirms the non-negligible gap previously observed by [1] between the order of the cover time and the order of the PCT for UDGs, further justifying consideration of the partial cover.

7.2 Quality of the Random Walk based on Hole Size

As previously stated, the load balancing properties of an almost regular graph are directly related to the graph's mixing rate, the measure of variation distance over time. Fig. 3 compares the variation distance as a function of number of steps (time) for different $G_R(N)$s and 3D-grid $M(3, N)$. Again, as expected, when R increases the variation distance drops faster. Also note that with R less than the radius of the $G_R(N)$ resembling the 3D-grid (0.5176) we get even better mixing time than the 3D-grid. Since this was also the case in Fig. 2 and Table 1 it seems that the UDG that resembles $M(3, N)$ actually has better cover time and mixing time than $M(3, N)$.

When the mixing time is better one would expect that the quality of the partial cover will improve, meaning that the random walk will not leave large contiguous areas in the network uncovered. To make this measurement more precise let $min(v)$ be the minimum distance from v to a visited node in a (partial) random walk. We define the *hole size* of a random walk as the maximum

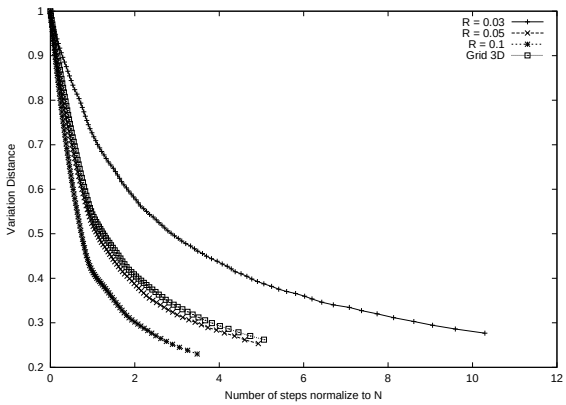

Figure 3: The decrease in variation distance as function of steps up to 95% partial cover

of $min(v)$ over all the nodes in the graph. Note that the hole size is decreasing as the random walk proceeds and more nodes are visited; after cover time it is 0.

Fig. 4 presents the decrease in the hole size as a function of number of steps of random walk for $G_R(N)$ with increasing R. The figure shows that the rate of improvement in the quality is strongly dependent upon R, similarly to the mixing rate. Note that each walk was sampled at 10%, 20%, and up to 100% cover where the hole size is 0. An interesting point to discuss is that this experiment also validates the fact that graphs with different spectral gap (and mixing time) can have the same cover time (such as the 3D grid and the hypercube). For example the graphs with $R > 0.06$ seems to have very similar cover time but nevertheless very different partial cover quality.

Fig. 5 shows the improvement in the hole size after 80% cover as a function of R. Note that it behaves like the function R^{-2}, which further supports the intuition that the hole size is directly related to the mixing rate.

8 Conclusions and Future Work

We have analytically obtained bounds on the mixing time of $G_R(N)$ which is $O(R^{-2} \ln N)$. In particular, the analytical bounds show that the radius required for rapid mixing is $R = O(1/poly(logN))$. Although we also obtained analytical bounds on the cover time and partial cover time of $G_R(N)$ based on the spectral gap, experimental results on the UDGs that resemble $M(3, N)$ indicate that those bounds are not tight. In particular, $R = O(1/N^{1/3})$ is sufficient radius to obtain optimal cover time and partial cover time. However, experimental results on hole size and intuition on load-balancing indicate that for the random walk to have good quality, a short mixing time, or rapid mixing, is needed. Therefore, to be both efficient and have good quality cover and partial cover

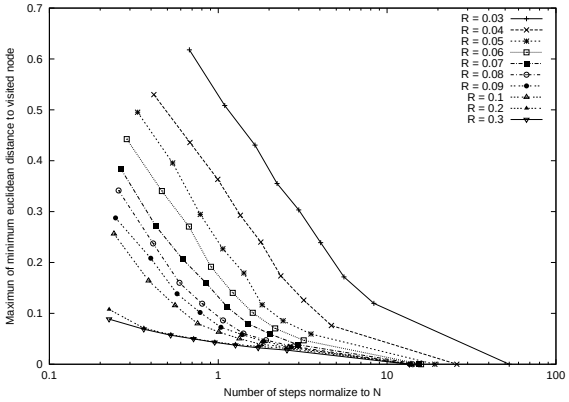

Figure 4: Hole Size as a Function of the Number of Steps Normalized to N for four UDGs with different radii R

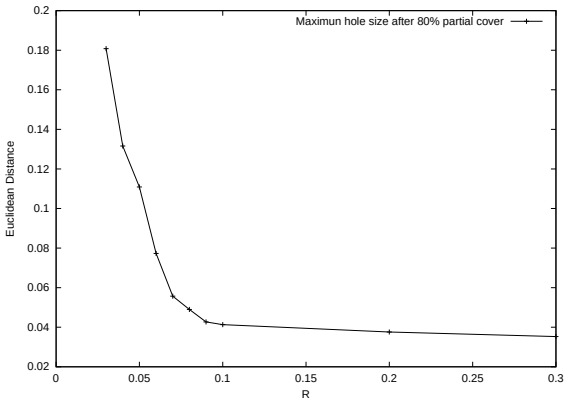

Figure 5: Hole Size at 80% Partial Cover as a Function of radius R

$R = O(1/poly(logN))$ is sufficient.

These bounds may yet be improved or proven analytically. For example, experimental results indicate that the UDG that resembles $M(3, N)$ actually performs better than $M(3, N)$. Moreover, it would be nice to know analytically why the UDG that resembles $M(3, N)$ has optimal cover time and partial cover time. For this insight and in order to improve the bounds, we plan to investigate the electrical resistance of $G_R(N)$, which is intimately connected with the cover time and partial cover time. We also plan to analytically investigate the exact relationship between mixing time and the quality of the random walk in terms of hole size.

Finally, although our random walk does not pose any interference in and of itself, in sensor networks there is always the issue of energy optimization. Whereas finding small radii for which optimal random properties are exhibited implicitly incorporates the idea that large radii impose interference and energy constraints, we would like to explicitly incorporate energy and interference into our current random walk model and find the optimum radius under the new model.

Acknowledgements

We would like to thank Eli Gafni, Yingniang Wu, and Carlos Brito for helpful discussions and corrections.

[Continuous Approximation of Conductance] Follow Fig. 6 Let l be the dividing. A point p in S that is at distance $x < R$ from l neighbors the nodes in the gray area A in the Figure. The size of A is given by $\frac{1}{2}R^2(\theta - \sin\theta)$. (Observe that $\theta = \arccos(\frac{x}{R})$ and A is a function of x.) So p has an expected number of NA edges crossing to $\bar{S}$. Taking the integral over all the points in distance $0 \leq x \leq R$ and assuming that there are $N\Delta x$ nodes in the area $1 \cdot \Delta x$ we get that the expected number of edges crossing from S to $\bar{S}$ is (ignoring the effect of the borders) [6]

[6] Note that as $N \to \infty$ and $R \to 0$ the above bound tightens and approaches equality.

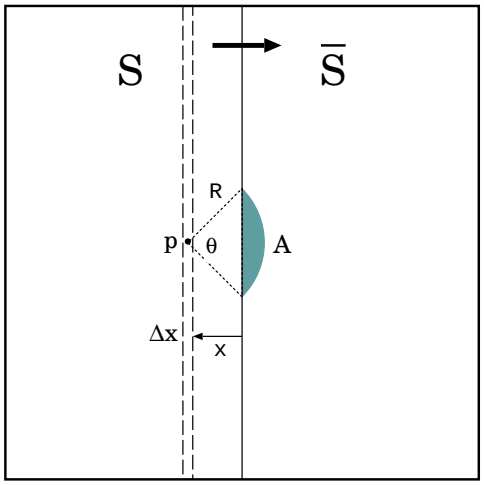

Figure 6: Approximating the Conductance in UDG

$$
\begin{aligned}
E[Cut(S,\bar{S})] &\leq \int_0^R \frac{1}{2}R^2 NAN dx \\
&= \int_0^R \frac{1}{2}R^2 N\big[\arccos(\frac{x}{R}) \\
&\quad -\sin(\arccos(\frac{x}{R})\big]N dx \\
&= \frac{1}{2}R^2N^2\Big[-2R\sqrt{1-\frac{x^2}{R^2}} \\
&\quad +\frac{2}{3}R(1-\frac{x^2}{R^2})^{\frac{3}{2}} + 2x\arccos(\frac{x}{R})\Big]_0^R \\
&= \frac{1}{2}R^2N^2(0-(-2R+\frac{2}{3}R) \\
&= \frac{2}{3}R^3N^2
\end{aligned}
$$

To approximate the conductance we use this upper bound and improve it by assuming the expected degree is $\pi R^2 N$ and by taking out part of the border

effect as we take the integral over the area $(1-R)\cdot\Delta x$ $(R << 1)$

$$\begin{aligned}
\Phi(G_R(N)) &= \frac{Q(S,\bar{S})}{\pi(S)} \\
&= 2\sum_{\substack{x\in S\\ y\in\bar{S}}} \pi(x)P(x,y) \\
&\approx 2\sum_{\substack{x\in S\\ y\in\bar{S}}} \frac{1}{N}\frac{1}{\pi R^2 N} \qquad (15)\\
&\approx 2\frac{2}{3}N^2R^3(1-R)\frac{1}{\pi R^2N^2} \\
&= \frac{4}{3\pi}R(1-R)
\end{aligned}$$

References

[1] C. Avin and C. Brito, "Efficient and robust query processing in dynamic environments using random walk techniques," in *Proceedings of the third international symposium on Information processing in sensor networks*. ACM Press, 2004, pp. 277–286.

[2] S. Shenker, S. Ratnasamy, B. Karp, R. Govindan, and D. Estrin, "Data-centric storage in sensornets," *ACM SIGCOMM Computer Communication Review*, vol. 33, no. 1, pp. 137–142, 2003.

[3] S. Madden, M. J. Franklin, J. M. Hellerstein, and W. Hong, "Tag: a tiny aggregation service for ad-hoc sensor networks," *ACM SIGOPS Operating Systems Review*, vol. 36, no. SI, pp. 131–146, 2002.

[4] A. Heinzelman, W.R.and Chandrakasan and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks," in *Proc. of the 33rd Annual Hawaii Int. Conference on System Sciences*, 2000, pp. 3005–3014.

[5] G. J. Pottie and W. J. Kaiser, "Wireless integrated network sensors," *Communications of the ACM*, vol. 43, no. 5, pp. 51–58, 2000.

[6] D. Kempe, A. Dobra, and J. Gehrke, "Gossip-based computation of aggregate information," in *Proc. of the 44th Annual IEEE Symposium on Foundations of Computer Science*, 2003, pp. 482–491.

[7] C. Gkantsidis, M. Mihail, and A. Saberi, "Random walks in peer-to-peer networks," in *in Proc. 23 Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM). to appear*, 2004.

[8] D. J. Aldous, "On the time taken by random on finite groups to visit every state," *Z. Wahrsch. Verw. Gebiete*, vol. 62, no. 3, pp. 361–374, 1983.

[9] P. Matthews, "Covering problems for Brownian motion on spheres," *Ann. Probab.*, vol. 16, no. 1, pp. 189–199, 1988.

[10] R. Motwani and P. Raghavan, *Randomized algorithms*. Cambridge University Press, 1995.

[11] B. A. Clark and C. J. Colbourn, "Unit disk graphs," *Discrete Mathematics*, vol. 86, pp. 165–177, 1991.

[12] M. Jerrum and A. Sinclair, "The markov chain monte carlo method: an approach to approximate counting and integration," in *Approximations for NP-hard Problems, Dorit Hochbaum ed.* PWS Publishing, Boston, MA, 1997, pp. 482–520.

[13] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, "Gossip and mixing times of random walks on random graphs," unpublished, 2004. http://www.stanford.edu/ boyd/reports/gossip_gnr.pdf.

[14] L. Lovász, "Random walks on graphs: A survey," in *Combinatorics, Paul Erdős is eighty, Vol. 2 (Keszthely, 1993)*, ser. Bolyai Soc. Math. Stud. Budapest: János Bolyai Math. Soc., 1996, vol. 2, pp. 353–397.

[15] A. Sinclair, "Improved bounds for mixing rates of markov chains and multicommodity flow," *Combinatorics, Probability and Computing*, vol. 1, pp. 351–370, 1992.

[16] P. Gupta and P. R. Kumar, "Critical power for asymptotic connectivity in wireless networks," *In Stochastic Analysis, Control, Optimization and Applications: A Volume in Honor of W. H. Fleming, W. M. McEneaney, G. Yin, and Q. Zhang, Eds*, pp. 547–566, 1998.

[17] A. Broder and A. Karlin, "Bounds on the cover time," *J. Theoret. Probab.*, vol. 2, pp. 101–120, 1989.

[18] D. Aldous and J. Fill, *Reversible Markov Chains and Random Walks on Graphs*, unpublished. http://stat-www.berkeley.edu/users/aldous/RWG/book.html.

[19] A. K. Chandra, P. Raghavan, W. L. Ruzzo, and R. Smolensky, "The electrical resistance of a graph captures its commute and cover times," in *Proc. of the twenty-first annual ACM symposium on Theory of computing*. ACM Press, 1989, pp. 574–586.