Mixing Times for Random Walks on Geometric Random Graphs

Stephen Boyd Arpita Ghosh Balaji Prabhakar Devavrat Shah *†
Information Systems Laboratory, Stanford University
Stanford, CA 94305-9510
{boyd, arpitag, balaji, devavrat}@stanford.edu

Abstract

A geometric random graph, $G^d(n,r)$, is formed as follows: place n nodes uniformly at random onto the surface of the d-dimensional unit torus and connect nodes which are within a distance r of each other. The $G^d(n,r)$ has been of great interest due to its success as a model for ad-hoc wireless networks. It is well known that the connectivity of $G^d(n,r)$ exhibits a threshold property: there exists a constant α_d such that for any $\epsilon > 0$, for $r^d < \alpha_d(1-\epsilon)\log n/n$ the $G^d(n,r)$ is not connected with high probability[1] and for $r^d > \alpha_d(1+\epsilon)\log n/n$ the $G^d(n,r)$ is connected *w.h.p.*. In this paper, we study mixing properties of random walks on $G^d(n,r)$ for $r^d(n) = \omega(\log n/n)$. Specifically, we study the scaling of mixing times of the fastest-mixing reversible random walk, and the natural random walk. We find that the mixing time of both of these random walks have the same scaling laws and scale proportional to r^{-2} (for all d). These results hold for $G^d(n,r)$ when distance is defined using any L_p norm. Though the results of this paper are not so surprising, they are non-trivial and require new methods.

To obtain the scaling law for the fastest-mixing reversible random walk, we first explicitly characterize the fastest-mixing reversible random walk on a regular (grid-type) graph in d dimensions. We subsequently use this to bound the mixing time of the fastest-mixing random walk on $G^d(n,r)$. In the course of our analysis, we obtain a tight relation between the mixing time of the fastest-mixing symmetric random walk and the fastest-mixing reversible random walk with a specified equilibrium distribution on an arbitrary graph.

To study the natural random walk, we first generalize a method of [DS91] to bound eigenvalues based on Poincare's inequality and then apply it to the $G^d(n,r)$ graph.

We note that the methods utilized in this paper are novel and general enough to be useful in the context of other graphs.

1 Introduction

A d-dimensional geometric random graph is obtained by placing n nodes uniformly at random on the surface of a d-dimensional unit torus, and connecting nodes within Euclidean distance r of each other. Such a graph is denoted by $G^d(n,r)$[Pen03]. Geometric random graphs have been used successfully in applications where the existence of an edge between two nodes depends on the distance between the nodes. Classically, they have been very useful in percolation, statistical physics, hypothesis testing, cluster analysis, etc. [Pen03]. More recently, the $G^2(n,r)$ graph has been used to model the network connectivity graph for wireless ad-hoc networks and sensor networks [GK00].

In this paper, we study the mixing times of random walks on $G^d(n,r)$. In addition to being of theoretical interest, the mixing time on a graph is directly related to the convergence time of iterative averaging algorithms on that graph, as shown recently in [BGPS04]. In particular, the mixing time of a random walk on $G^d(n,r)$ is connected to the averaging time on a wireless sensor network (modeled as $G^d(n,r)$). This strongly motivates the study of mixing times of random walks on $G^d(n,r)$. As noted in [BGPS04], the natural random walk corresponds to a very simple distributed averaging algorithm. The goal is to compare the performance of this algorithm with the optimal averaging algorithm. This is equivalent to comparing the mixing time of the natural random walk and the fastest-mixing reversible random walk with a uniform stationary distribution. Thus,

*Author names appear in alphabetical order.

†This work is supported in part by a Stanford Graduate Fellowship, and by C2S2, the MARCO Focus Center for Circuit and System Solution, under MARCO contract 2003-CT-888.

[1] In this paper, with high probability (*w.h.p.*) means with probability at least $1 - 1/n^2$.

in this paper, we study the mixing time of the natural random walk and fastest-mixing reversible random walk with uniform stationary distribution.

For completeness, we first state the following definition:

DEFINITION 1. (REVERSIBLE RANDOM WALK) *A random walk on a connected graph with n nodes is defined by the $n \times n$ transition matrix $P = [P_{ij}]$, where P_{ij} is the probability of going from node i to node j. Let $\Pi = [\Pi_i]$ be the stationary distribution, that is, $\Pi P = \Pi$. The random walk is called reversible iff $\Pi_i P_{ij} = P_{ji}\Pi_j$ for all i, j.*

DEFINITION 2. (MIXING TIME) *The mixing time of such a random walk is defined as follows: for any node i define $\Delta_i(t) = \frac{1}{2}\sum_{j=1}^{n} |P_{ij}^t - \Pi_j|$. Then, the mixing time is*

$$T_{\text{mix}}(\epsilon) = \sup_i \inf\{t : \Delta_i(t') \leq \epsilon \ \text{ for all } t' \geq t\}. \quad (1.1)$$

It is well-known that the mixing time of a reversible random walk is related to the second largest eigenvalue in absolute value of P; the following result quantifies this relationship (see survey [Gur00])[2].

LEMMA 1.1. *The mixing time of a random walk with transition matrix P is bounded as follows:*

$$\frac{\lambda_{\max}(P)\log(2\epsilon)^{-1}}{2(1-\lambda_{\max}(P))} \leq T_{\text{mix}}(\epsilon) \leq \frac{\log n + \log \epsilon^{-1}}{1-\lambda_{\max}(P)}, \quad (1.2)$$

where $\lambda_{\max}(P)$ is defined as follows: let $1 = \lambda_1(P) \geq \cdots \lambda_n(P) \geq -1$ be ordered eigenvalues of reversible matrix P, then $\lambda_{\max}(P) = \max\{\lambda_2(P), -\lambda_n(P)\}$.

Later in the paper, we will sometimes refer to $\lambda_{\max}$ as the mixing rate of the random walk and $1 - \lambda_{\max}(P)$ as the spectral gap of P. The *fastest-mixing* or *optimal* (reversible) random walk for a given stationary distribution is the one which has the smallest $\lambda_{\max}$ of all reversible random walks with that stationary distribution [BDX04].

Throughout this paper, when we state results for the mixing time T_{mix}, we will mean $T_{\text{mix}}(\epsilon)$ for $\epsilon = 1/n^{\alpha}$, $\alpha > 0$.

In order to discuss mixing time of a random walk with unique stationary distribution, it is necessary that the graph $G^d(n,r)$ be connected. The following is a well-known result about connectivity of $G^d(n,r)$ [GK00]:

THEOREM 1.1. *Let $r_c(d)$ be such that $nr_c(d)^d = 4\log n$. Then $G^d(n,r)$ is connected with high probability if $r \geq r_c(d)$, and it is not connected with positive probability if $r = o(r_c(d))$.*

For connected $G^d(n,r)$, for $r = \Theta(r_c(d))$, the graph does not have a regular structure, making it hard to study the mixing properties of random walks. Hence, we study mixing times of random walks for $r = \omega(r_c(d))$; for such r, the structure of the graph becomes regular, as shown in Lemma 2.1. Throughout this paper, we only are interested in random walks which are reversible[3], and have uniform stationary distribution. Also, to remove edge effects, we consider $G^d(n,r)$ on the d-dimensional unit torus. The results of this paper also hold if $G^d(n,r)$ is defined on a square or a sphere. The following is the main result of this paper:

THEOREM 1.2. *For $G^d(n,r)$ with $r = \omega(r_c(d))$, with high probability,*

(a) the mixing time of the fastest mixing reversible random walk with uniform stationary distribution is $\Theta\left(r^{-2}\log n\right)$, and

(b) the mixing time of the modified natural random walk, where a node jumps to any of its neighbors (other than itself) with equal probability, and has a self loop of probability $1/2$, is also $\Theta\left(r^{-2}\log n\right)$.

The rest of the paper is organized as follows: In Section 2 we obtain the regularity property of $G^d(n,r)$ for $r = \omega(r_c(d))$ (Lemma 2.1). We prove Theorem 1.2(a) in Section 3 and Theorem 1.2(b) in Section 4. Within Section 3, we first prove the results for $d = 1$ and then extend them to $d \geq 2$. To prove Theorem 1.2(a), given Theorem 1.2(b), we only need to show a lower bound of $\Omega\left(r^{-2}\log n\right)$.

1.1 Related Work

There is a large body of literature on properties of random geometric graphs. Connectivity properties for the $G(n,r)$ were derived in [GK00], where the $G^2(n,r)$ graph was used to model wireless ad-hoc networks as well as sensor networks. The recent book by Penrose [Pen03] contains comprehensive discussions about connectivity, vertex degree distributions, percolation, and many other graph properties for the $G(n,r)$ graph. Recently, sharp thresholds for monotone properties of $G(n,r)$ graphs have been obtained in

[2] For clarity we have stated the result for a uniform stationary distribution; since we are working with the order notation, the results do not change for the stationary distributions we encounter in this paper.

[3] Let $P = [P_{ij}]$ be the transition matrix corresponding to a random walk and $\Pi = [\Pi_i]$ be the corresponding stationary distribution. Then P is reversible iff $\Pi_i P_{ij} = \Pi_j P_{ji}, \ \forall i, j$.

[GRK04]. For some results on covering algorithms for these graphs, see [BBFM03].

Another popular kind of random graph is the Bernoulli random graph $G(n, p)$, which is a graph on n nodes formed by placing a edge between two nodes independently with probability p. Such graphs have been studied extensively; [Bol01] contains a comprehensive treatment of this subject.

2 Regularity of $G^d(n, r)$

In this section, we state a (possibly known) relatively straightforward regularity property of $G^d(n, r)$, which makes the analysis of the mixing time of random walks tractable.

LEMMA 2.1. *For $G^d(n, r)$ with $r = \omega(r_c(d))$, the degree of every node is $\alpha_d n r^d(1 + o(1))$ w.h.p., where $\alpha_d = \frac{\pi^{d/2}}{\Gamma(1+d/2)}$.*

Proof. Let nodes be numbered $i = 1, \cdots, n$. Consider a particular node, say 1. Let random variable X_j be 1 if node j is within distance r of node 1 and 0 otherwise. The X_js are IID Bernoulli with probability $p_d = \alpha_d r^d$ of success (the volume of a $d-$dimensional sphere with radius r is $\alpha_d r^d$). The degree of node 1 is

$$d_1 = \sum_{j=2}^{n} X_j. \tag{2.3}$$

By application of the Chernoff bound we obtain :

$$\begin{aligned} P(|\textstyle\sum_{j=2}^{n} X_j - (n-1)p_d| \geq \delta(n-1)p_d) \leq \\ 2\exp(-\tfrac{\delta^2(n-1)p_d}{2}). \end{aligned} \tag{2.4}$$

If we choose $\delta = \sqrt{\frac{8 \log n}{p_d(n-1)}}$, then the right-hand side in (2.4) becomes $2\exp(-4\log n) = 2/n^4$. So, for $p_d = \omega(\log n/n))$, node 1 has degree

$$\begin{aligned} d_1 &= (n-1)p_d \pm \sqrt{8(n-1)p_d \log n} \\ &\simeq np_d(1 \pm o(1)), \qquad w.p. \geq 1 - \tfrac{2}{n^4}. \end{aligned} \tag{2.5}$$

Using the union bound, we see that

$$\begin{aligned} P(\text{any node has degree} \neq np_d(1 \pm o(1))) &\leq n\tfrac{2}{n^4} \\ &= \tfrac{2}{n^3}. \end{aligned} \tag{2.6}$$

So for large n, *w.h.p.*, all nodes in the $G^d(n, r)$ have degree $np_d(1 \pm o(1))$.

3 Fastest mixing random walk on $G^d(n, r)$

In this section, we characterize the scaling of the fastest mixing random walk on $G^d(n, r)$ with uniform stationary distribution. We first consider the case of $d = 1$, *i.e.* $G^1(n, r)$. This is much easier than the higher dimensional $G^d(n, r)$ with $d \geq 2$. We completely characterize $G^1(n, r)$ with the help of one-dimensional regular graphs. For $G^d(n, r)$ with $d \geq 2$, we obtain a lower bound on the fastest-mixing reversible random walk. Note that since we are interested in reversible random walks with uniform stationary distribution, the transition matrix corresponding to the random walk must be symmetric. (The upper bound of the same order is implied by the natural random walk as in Theorem 2(b).) The remainder of the section is a proof of Theorem 2(a).

3.1 Fastest mixing random walk on $G^1(n, r)$

Let G_k denote the regular graph on n nodes with every node of degree $2k$: place the n nodes on the circumference of a circle, and connect every node to k neighbors on the left, and k on the right. From the regularity lemma, we have that *w.h.p.*, every node in $G^1(n, r)$ has degree $2nr(1 \pm o(1))$. Also, observe that the same technique can be used to show that *w.h.p.* the number of neighbors to the right (ditto left) is $nr(1 \pm o(1))$. Hence, *w.h.p.*the $G^1(n, r)$ is a subgraph of G_k for $k = 4nr$, since for any mapping of the nodes of $G^1(n, r)$ to G_k, an edge between nodes i and j in $G^1(n, r)$ is also present in G_k. Similarly, $G^1(n, r)$ also contains G_l, for $l = (1/2)nr$. Given this, we can now study the problem of finding the optimal random walk on G_k with uniform stationary distribution. We have the following lemma:

LEMMA 3.1. *For k, n such that $k \leq n/4$, the mixing rate of the fastest-mixing symmetric random walk on G_k cannot be smaller than $\cos(2\pi k/n)$.*

Proof. It can be shown using symmetry arguments [PXBD03] that the fastest mixing random walk on G_k with uniform stationary distribution will have a symmetric and circulant transition matrix. (For this simple graph, this can be easily seen using convexity of the second eigenvalue). So we can restrict our attention to the (circulant symmetric) transition matrices

$$P = \begin{bmatrix} p_0\ p_1\ \dots\ p_k\ 0\ \ 0\ \dots\ 0\ p_k\ \dots\ p_2\ p_1 \\ p_1\ p_0\ p_1\ \dots\ p_k\ 0\ \dots\ 0\ 0\ p_k\ \dots\ \ p_2 \\ \vdots \qquad\qquad \vdots \\ \vdots \qquad\qquad \vdots \\ p_1\ \dots\ p_k\ 0\ 0\ \dots\ 0\ \ p_k\ \dots\ p_2\ p_1\ p_0 \end{bmatrix} \tag{3.7}$$

The eigenvalues of this matrix are

$$\mu_m = \sum_{j=0}^{k} p_j e^{-2\pi ijm/n} + \sum_{j=1}^{k} p_j e^{-2\pi i(n-j)m/n)}$$
$$= p_0 + 2\sum_{j=1}^{k} p_j \cos(2\pi jm/n), \quad m = 0, \dots, n-1.$$

For $m = 0$, $\mu_m = 1$, which is the largest eigenvalue. Let $\mathbf{p} = (p_0, p_1, \dots, p_k, p_k, \dots, p_1)$. We are interested in the smallest possible second largest eigenvalue in absolute value, *i.e.*,

$$\begin{array}{lll} \min_{\mathbf{p}} & \max_{m=\{1,\dots,n-1\}} & |\mu_m| \\ & \text{subject to} & \mathbf{1}^T\mathbf{p} = 1, \\ & & \mathbf{p} \succeq 0. \end{array} \tag{3.8}$$

We can obtain a lower bound for the optimal value of (3.8). Now,

$$\begin{array}{rcl} \mu_2 & \le & \max_{m=\{1,\dots,n-1\}} |(\mu_m)| \\ \Rightarrow \min_{\mathbf{p}} \mu_2 & \le & \min_{\mathbf{p}} \max_{m=\{1,\dots,n-1\}} |(\mu_m)|. \end{array} \tag{3.9}$$

The right hand side is the solution of the following linear program with a single total sum constraint:

$$\begin{array}{ll} \min_{\mathbf{p}} & p_0 + 2\sum_{j=1}^{k} p_j \cos(2\pi j/n) \\ \text{s.t.} & \mathbf{1}^T\mathbf{p} = 1 \\ & \mathbf{p} \succeq 0. \end{array} \tag{3.10}$$

For k such that each of the coefficients $\cos(2\pi j/n)$ is positive, *i.e.*, for $k \le n/4$, the smallest coefficient is $\cos(2\pi k/n)$, and so for all such k and n, the minimum value is $\cos(2\pi k/n)$, obtained at $p_k = 1/2$, $p_j = 0$ for all other j. [4] So the fastest mixing random walk on this graph cannot have a mixing rate smaller than $\cos(2\pi k/n)$.

The above result was proved for all $k \le n/4$; however, we will be interested only in those cases where $k = o(n)$, *i.e.*, the graph is not too well connected. For such k, the following lemma allows us to find a 'nearly optimal' transition matrix:

LEMMA 3.2. *For $k = o(n)$, there is a random walk on G_k for which the mixing rate is $\lambda_{\max} = \cos(2\pi k/n) + \Theta(k^4/n^4)$.*

The proof of the lemma is included in the Appendix.

[4] Note that this is only a lower bound: for this $\mathbf{p}$, if k divides n, the second largest eigenvalue is also 1, attained at $m = n/k$.

3.2 Fastest mixing random walk on $G^2(n,r)$ We present the lower bound on the fastest-mixing reversible random walk on $G^2(n,r)$ in this section. The same method can be easily extended to $d \ge 3$. First we characterize the fastest-mixing reversible random walk on a two-dimensional regular graph, G_{kk}, defined as follows: form a lattice on the unit torus, where lattice points are located at $(i/\sqrt{n}, j/\sqrt{n})$, $-\sqrt{n}/2 \le i,j \le +\sqrt{n}/2$, and place the n nodes at these points. An edge between two vertices exists if the L_∞ distance between them is at most $k/\sqrt{n}$. For such G_{kk} the fastest-mixing time scales as follows:

LEMMA 3.3. *The mixing rate of the fastest-mixing reversible random walk on G_{kk} is no smaller than $\cos^2(2\pi k/\sqrt{n})$, that is, the mixing time of the fastest-mixing random walk is such that $T_{\text{mix}} = \Omega(n \log n/k^2)$.*

Proof. As in the one-dimensional case, by symmetry, the optimal transition probability between nodes i and j will depend only on the distance between these nodes. Using this, we can write the transition matrix corresponding to such a symmetric random walk on G_{kk} as the Kronecker (or tensor) product $P_k \otimes P_k$, where $P_k \in \mathbf{R}^{n\times n}$ is as in (3.7). This is not difficult to visualize: for $i,j = 0,\dots,n-1$, $a,b = 1,\dots,n$,

$$(P \otimes P)_{ni+a,nj+b} = P_{i+1,j+1}P_{ab}. \tag{3.11}$$

Now the eigenvalues of $A \otimes B$ are all products of eigenvalues of A and B, so that for $0 \le i,j \le n-1$,

$$\begin{array}{rcl} \lambda_{ij}(P\otimes P) & = & \lambda_i(P)\lambda_j(P) \\ & = & (p_0 + 2\sum_{m=1}^{k} p_m \cos(2\pi \frac{im}{\sqrt{n}}) \\ & \cdot & (p_0 + 2\sum_{m=1}^{k} p_m \cos(2\pi \frac{jm}{\sqrt{n}}). \end{array}$$

The eigenvalue 1 is obtained by setting $i = j = 0$; all other eigenvalues will have absolute value less equal 1. We want to find a lower bound for the second largest eigenvalue in absolute value, call it $\lambda^\star_{\max}$.

As before, choose $i = j = 1$. Then

$$\begin{array}{rcl} \lambda_{11} & \le & \max_{i,j\neq 0} |\lambda_{ij}| \\ \Rightarrow \min_{\mathbf{p}} \lambda_{11} & \le & \min_{\mathbf{p}} \max_{i,j\neq 0} |\lambda_{ij}|, \end{array}$$

so that $\min_{\mathbf{p}} \lambda_{11}$ is a lower bound for $\lambda^\star_{\max}$. Making the assumption again that $k \le \sqrt{n}/4$, the minimizing

$\mathbf{p}$ is the one with $p_k = 1/2$ and $p_i = 0, i \neq k$ (which corresponds to transition probabilities of 1/4 for each of the 4 farthest diagonal nodes, and 0 everywhere else). The value of λ_{11} corresponding to this distribution is $\cos^2(2\pi \frac{k}{\sqrt{n}})$. This is of order $1 - \Theta(\frac{k^2}{n})$, since $\cos^2(2\pi \frac{k}{\sqrt{n}}) = \frac{1}{2} + \frac{1}{2}\cos(2\frac{2\pi k}{\sqrt{n}}) = 1 - \Theta(\frac{k^2}{n})$.

Thus, $(1 - \lambda_{\max}) = O(k^2/n)$. Hence by Lemma 1.1, the corresponding mixing time $T_{\text{mix}} = \Omega(n \log n / k^2)$. [5]

The G_{kk} graph was constructed using the L_∞ distance between vertices. Therefore, the graph formed by placing edges between vertices based on distance measured in *any* L_p norm (for the same k) is a subgraph of G_{kk}, and has a mixing time lower bounded by the mixing time of G_{kk}. Thus our bounds will be valid for the $G(n,r)$ graph constructed according to *any* L_p norm.

Now we'll use the bound on the fastest mixing walk on G_{11} to obtain a bound for $G^2(n,r)$. First we create a new graph $\tilde{G}^2(n,r)$ as follows: place a square grid with squares of side r on the unit torus. Using arguments similar to that used in the proof of Lemma 2.1, each square of area r^2 contains $nr^2(1+o(1))$ nodes for $r = \omega(r_c(d))$. For each of these r^{-2} squares do the following: connect every node in a square to all the nodes in the neighboring 8 squares, as well as the nodes in the same square. Thus, each node is connected to $9nr^2(1+o(1))$ nodes in $\tilde{G}^2(n,r)$. By definition, all edges in $G^2(n,r)$ are present in $\tilde{G}^2(n,r)$ and therefore, the fastest-mixing random walk on $\tilde{G}^2(n,r)$ is at least as fast as that of $G^2(n,r)$. Thus, lower-bounding the fastest-mixing random walk on $\tilde{G}^2(n,r)$ is sufficient to obtain lower bound on the fastest mixing random walk on $G^2(n,r)$.

Now, construct a graph G of r^{-2} nodes as follows: for each square in the square grid used in $\tilde{G}^2(n,r)$, create a node in G. Thus, G has r^{-2} nodes. Two nodes are connected in G if the corresponding squares in the grid are adjacent. Thus, each node is connected to 8 other nodes. Thus, G is a regular graph G_{11} with r^{-2} nodes. In order to use the lower bound on the fastest mixing random walk on G_{11} of r^{-2} nodes (*i.e.*G) as a lower bound on $\tilde{G}^2(n,r)$, we need to show that the fastest-mixing symmetric random walk on $\tilde{G}^2(n,r)$ induces a time-homogeneous reversible random walk on G. This will be implied by the following Lemma.

Lemma 3.4. *There exists a fastest-mixing symmetric random walk on $\tilde{G}^2(n,r)$, whose transition matrix P has the following property: for any two nodes i and j belonging to the same square, $P_{ik} = P_{jk}$ for $k \neq i, j$, and $P_{ii} = P_{jj}$.*

Proof. We prove this by contradiction. Suppose the claimed statement is not true, *i.e.*, there is no transition matrix achieving the smallest $\lambda_{\max}$ with the above property. Since the optimal value of $\lambda_{\max}$ must be attained ([BDX04]), consider such an optimizing P_1, and let i and j be two nodes in the same square for which the above property is not true.

Let A be the permutation matrix with $A_{ij} = A_{ji} = 1, A_{ii} = A_{jj} = 0$, and all other diagonal entries 1 and all other non-diagonal entries 0. Note that A is a symmetric permutation matrix, and therefore $A^{-1} = A^T = A$. Consider the matrix $P_2 = AP_1A$; since $A = A^{-1}$, P_1 and P_2 are similar, and so have the same eigenvalues. Note that since i and j belong to the same square in $\tilde{G}$, they have exactly the same neighbors, and therefore P_2 also respects the graph structure (*i.e.*, $P_{2_{ab}} \neq 0$ only if a and b have an edge between them).

Now, $\lambda_{\max}(P)$ is a convex function of P for symmetric stochastic P ([BDX04]), so

$$\lambda_{\max}\left(\frac{P_1 + P_2}{2}\right) \leq \frac{1}{2}\lambda_{\max}(P_1) + \frac{1}{2}\lambda_{\max}(P_2) = \lambda_{\max}(P_1). \tag{3.12}$$

But $P = (P_2 + P_1)/2$ has the property claimed in the lemma for nodes i and j: $P_{ik} = P_{jk}$ for all $k \neq i, j$, $P_{ii} = P_{jj} = (P_{1_{ii}} + P_{1_{jj}})/2$, and $\lambda_{\max}(P) \leq \lambda_{\max}(P_1)$. We can apply the above procedure recursively (even for multiple rows) to construct a matrix P^* with smallest $\lambda_{\max}$ and the property claimed in the Lemma. This contradicts our assumption and completes the proof.

From Lemma 3.4, we see that under the fastest mixing random walk, the probability of transiting from a node in a square, say S_1, to some neighboring square, say S_2, is the same for all nodes in S_1 and S_2. Thus, essentially we can view the random walk evolving over squares. That is, the fastest random walk on $\tilde{G}^2(n,r)$ induces a random walk on the graph G. By definition of mixing time, the mixing time for this induced random walk on G (with induced equilibrium distribution) certainly lower bounds the mixing time for the random walk on $\tilde{G}^2(n,r)$. Further, the induced random walk is reversible as the random walk was symmetric on $\tilde{G}^2(n,r)$. Therefore, we obtain that the lower bound on mixing time for the fastest-mixing random walk on G implies a lower bound on the mixing time for the fastest-mixing random walk on $\tilde{G}^2(n,r)$. From Lemma 3.3 we have a lower bound of $\Omega(r^{-2}\log n)$ on the mixing time of the fastest-mixing symmetric random walk (i.e. with uniform stationary distribution). From Lemma 3.5 given

[5] It is easy to see that a result similar to Lemma 3.2 can be obtained for $d \geq 2$ using the same method.

below, this in turn implies lower bound of $\Omega(r^{-2}\log n)$ on mixing time of the fastest mixing reversible random walk on $G^2(n,r)$. This completes the proof of 2(a) for $G^2(n,r)$. It is easy to see that the arguments presented above can be readily extended to the case of $d \geq 3$.

LEMMA 3.5. *Consider a connected graph $G = (\{1,\ldots,n\},E)$. Let $T^*_{\text{mix}}(\pi)$ be the mixing time (with $\epsilon = 1/n^\alpha$ for some $\alpha > 0$ as in definition 1) of the fastest mixing reversible random walk on G with stationary distribution π. Let $\beta(\pi) = \max_{i,j}\frac{\pi(i)}{\pi(j)} \leq C$, where C is a constant. Then,*

$$T^*_{\text{mix}}(\pi) = \Omega\left(T^*_{\text{mix}}\left(\frac{1}{n}\mathbf{1}\right)\right), \tag{3.13}$$

i.e., the fastest mixing time for π is no faster than that of the uniform distribution.

Proof. Consider a reversible random walk with stationary distribution π on G and let its transition matrix be R. We will prove the following claim, which in turn implies the statement of the Lemma.

Claim I. There exists a symmetric random walk on graph G with transition matrix S such that

$$T_{\text{mix}}(S) = O(T_{\text{mix}}(R)).$$

Proof of Claim I. For a reversible matrix R, by definition,

$$\pi(i)R(i,j) = \pi(j)R(j,i), \quad \forall\, i,j.$$

Define matrix $P = [P(i,j)]$, where for $i \neq j$,

$$P(i,j) = \begin{cases} R(i,j) & \text{if } \pi(i) \geq \pi(j) \\ \frac{\pi(i)}{\pi(j)}R(j,i) & \text{if } \pi(i) < \pi(j) \end{cases}$$

and $P(i,i) = 1 - \sum_{j\neq i} P(i,j)$.

By definition and reversibility of R, P is a symmetric doubly stochastic matrix. Further, for $i \neq j$, $P(i,j) > 0$ if and only if $R(i,j) > 0$. Hence, P can be viewed as a transition matrix of a symmetric random walk on G, whose stationary distribution is uniform.

Define $Q^R = [Q^R(i,j)]$, where

$$Q^R(i,j) = \pi(i)R(i,j) = \pi(j)R(j,i).$$

Similarly, define $Q^P = \frac{1}{n}P$. Let $\phi : \{1,\ldots,n\} \to \mathbf{R}$ be a non-constant function. Define two quadratic forms, $\mathcal{E}^R$ and $\mathcal{E}^P$, of ϕ, as

$$\begin{aligned} \mathcal{E}^R(\phi,\phi) &= \frac{1}{2}\sum_{i,j}(\phi(i)-\phi(j))^2 Q^R(i,j); \\ \mathcal{E}^P(\phi,\phi) &= \frac{1}{2}\sum_{i,j}(\phi(i)-\phi(j))^2 Q^P(i,j). \end{aligned}$$

Let the variance of ϕ with respect to two different random walks be

$$\begin{aligned} V^R(\phi) &= \frac{1}{2}\sum_{i,j}(\phi(i)-\phi(j))^2\pi(i)\pi(j); \\ V^P(\phi) &= \frac{1}{2}\sum_{i,j}(\phi(i)-\phi(j))^2\frac{1}{n^2}. \end{aligned}$$

Let $\lambda_2(P)$ and $\lambda_2(R)$ denote the second largest eigenvalue of matrices P and R respectively. The minimax characterization of eigenvalues ([HJ85], page 176), gives a bound on the second largest eigenvalue of a reversible matrix $X(=P,R)$ as

$$(1-\lambda_2(X)) = \inf\{\frac{\mathcal{E}^X(\phi,\phi)}{V^X(\phi)} : \phi \text{ non-constant}\}. \tag{3.14}$$

For any π, $\sum_i \pi(i) = 1$, hence $\max_i \pi(i) \geq 1/n$ and $\min_j \pi(j) \leq 1/n$. Further, by the property of π, $\max_{i,j}\frac{\pi(i)}{\pi(j)} = \frac{\max_i \pi(i)}{\min_j \pi(j)} < C$. Hence, for any k,

$$\begin{aligned} \pi(k) &\geq \min_i \pi(i) \geq \frac{\max_i \pi(i)}{C} \geq \frac{1}{nC}, \text{ similarly,} \\ \pi(k) &\leq \max_i \pi(i) \leq C\min_j \pi(j) \leq \frac{C}{n}. \end{aligned}$$

Thus, for any k,

$$\frac{\pi(k)}{1/n} \in \left(\frac{1}{C}, C\right).$$

This implies that

$$\frac{\mathcal{E}^R(\phi)}{\mathcal{E}^P(\phi)} \in \left(\frac{1}{C^2}, C^2\right); \quad \frac{V^R(\phi)}{V^P(\phi)} \in \left(\frac{1}{C^2}, C^2\right).$$

Hence, from (3.14) we obtain $(1-\lambda_2(P)) = \Theta(1-\lambda_2(R))$. Now, we are considering $T_{\text{mix}}(\epsilon)$ for $\epsilon = 1/n^\alpha, \alpha > 0$. Hence, from Lemma 1.1,

$$T_{\text{mix}}(R) = \Theta\left(\frac{\log n}{1-\lambda_{\max}(R)}\right).$$

By definition, $(1-\lambda_{\max}(R)) \leq (1-\lambda_2(R))$. Hence,

$$T_{\text{mix}}(R) = \Omega\left(\frac{\log n}{1-\lambda_2(R)}\right).$$

It is easy to see that random walk on G with symmetric transition matrix $S = (I+P)/2$ has mixing time given by

$$T_{\text{mix}}(S) = \Theta\left(\frac{\log n}{1-\lambda_2(P)}\right).$$

Thus, $T_{\text{mix}}(S) = O(T_{\text{mix}}(R))$. This completes the proof of Claim I and the proof of Lemma 3.5.

Remark: In fact, a stronger result can be proved, which is

$$T^*_{\text{mix}}(\pi) = \Theta\left(T^*_{\text{mix}}\left(\frac{1}{n}\mathbf{1}\right)\right).$$

One part of this has already been proved in the Lemma. The reverse direction is obtained similarly, as follows. Consider any symmetric random walk with transition matrix P, and suppose a stationary distribution π is specified, satisfying $\beta(\pi) = \max_{i,j} \frac{\pi(i)}{\pi(j)} \leq C$, where C is some constant. Then there is a reversible random walk $\bar{R}$ with stationary distribution π, such that $T_{\text{mix}}(\bar{R}) = O(T_{\text{mix}}(P))$. $\bar{R}$ is obtained as follows. Construct a matrix R from P as:

$$R(i,j) = \begin{cases} P(i,j) & \text{if } \pi(i) \leq \pi(j) \\ \frac{\pi(j)}{\pi(i)} P(i,j) & \text{if } \pi(i) > \pi(j), \end{cases}$$

for $i \neq j$, and $R_{ii} = 1 - \sum_{j \neq i} R_{ij}$. R is a stochastic reversible matrix, with stationary distribution π, since $\pi(i)R(i,j) = \pi(j)R(j,i)$. Following the same steps as above, we can conclude that

$$1 - \lambda_2(R) = \Theta(1 - \lambda_2(P)).$$

The matrix $\bar{R} = (I + R)/2$ has the same stationary distribution π and the second largest eigenvalue is $(1 + \lambda_2(R))/2$. Therefore, using Lemma 1.1,

$$T_{\text{mix}}(\bar{R}) = \Theta\left(\frac{\log n}{1 - \lambda_2(R)}\right).$$

As before,

$$\begin{aligned} T_{\text{mix}}(P) &= \Theta\left(\frac{\log n}{1 - \lambda_{\max}(P)}\right) \\ &= \Omega\left(\frac{\log n}{1 - \lambda_2(P)}\right). \end{aligned}$$

Therefore, $T_{\text{mix}}(\bar{R}) = O(T_{\text{mix}}(P))$, and we have the stronger result as claimed in the Remark.

4 Natural random walk on $G^d(n,r)$

In this section, we study the mixing properties of the natural random walk on $G^d(n,r)$. Recall that under the natural random walk, the next node is equally likely to be any of the neighboring nodes. It is well known that under the stationary distribution, the probability of the walk being at node i is proportional to the degree of node i. By Lemma 2.1, all nodes have almost equal degree. Hence the stationary distribution is almost uniform (it is uniform asymptotically). The rest of this section is the proof of Theorem 1.2(b).

Consider a symmetric random walk with transition matrix P as follows: let d_* be the maximum degree of any node in $G^d(n,r)$, then

$$P_{ij} = \begin{cases} 1/d_* & \text{if } i \neq j \text{ are connected} \\ 0 & \text{if } i \neq j \text{ are not connected} \\ 1 - \sum_{j \neq i} P_{ij} & \text{if } i = j. \end{cases}$$

By definition, P is a doubly stochastic symmetric matrix on $G^d(n,r)$. Further, $d^* = \alpha_d n r^d (1 + o(1))$ *w.h.p.* It will be clear to the reader at the end of this section, that using the proof technique of this section, it follows that $T_{\text{mix}}(P) = O(r^{-2} \log n)$. This will imply the upper bound for the proof of Theorem 1.2(a).

4.1 Proof of Theorem 1.2(b):

We use a modification of a method developed by Diaconis-Stroock [DS91] to obtain bounds on the second largest eigenvalue using the geometry of the $G^d(n,r)$.

Note that for $d = 1$, the proof is rather straightforward. The difficulty arises in the case of $d \geq 2$. For ease of exposition in the rest of the section, we consider $d = 2$. Exactly the same argument can be used for $d > 2$. We begin with some initial setup and notation.

Square Grid: Divide the unit torus into a square grid where each square is of area $r^2/16$, i.e. of side length $r/4$. Consider a node in a square. By definition of $G^2(n,r)$, this node is connected to all nodes in the same square and all 8 neighboring squares.

Paths and Distribution: A path between two nodes i and j, denoted by γ_{ij}, is a sequence of nodes $(i, v_1, \cdots, v_{l-1}, j)$, $l \geq 1$, such that $(i, v_1), \cdots, (v_{l-1}, j)$ are edges in $G^2(n,r)$. Let $\gamma = (\gamma_{ij})_{1 \leq i \neq j \leq n}$ denote a collection of paths for all $\binom{n}{2}$ node pairs. Let Γ be the collection of all possible γ. Consider the probability distribution induced on Γ by selecting paths between all node-pairs in the following manner:

- Paths are chosen independently for different node pairs.
- Consider a particular node pair (i,j). We select γ_{ij} as follows: let i belong to square C_0 and j belong to square C_l.
  - If $C_0 = C_l$ or i and j are in the neighboring cells then the path between i and j is (i,j).
  - Else, let $C_1, \ldots, C_{l-1}$, $l \geq 2$ be other squares lying on the straight line joining i and j. Select a node $v_k \in C_k, k = 1, \ldots, l-1$ uniformly at random. Then the path between i and j is $(i, v_1, \ldots, v_{l-1}, j)$.

Under the above setup, we claim the following lemma:

Lemma 4.1. *Under the probability distribution on Γ as described above, the average number of paths passing through an edge is $O(1/r^3)$ w.h.p., where $r = \omega(r_c(d))$.*

Proof. We will compute the average number of paths passing through each square in the order notation. Similar to the arguments of Lemma 2.1, it can be shown that each of the $16/r^2$ squares contains $\frac{nr^2(1+o(1))}{16}$ nodes and each node has degree $\pi nr^2(1 + o(1))$ *w.h.p.*. We restrict our consideration to such instances of $G^2(n, r)$.

Now the total number of paths are $\Theta(n^2)$ since there are $\binom{n}{2}$ node pairs. Each path contains $O(1/r)$ edges, as $O(1/r)$ squares can be lying on a straight line joining two nodes. The total number of squares is $\Theta(1/r^2)$. Hence, by symmetry and regularity, the number of paths passing through each square is $\Theta(n^2r)$. Consider a particular square C. For C, at least $1 - \Theta(r^2)(\approx 1)$ fraction of paths passing through it have endpoints lying in squares other than C. That is, most of the paths passing through C have C as an intermediate square, and not an originating square. Such paths are equally likely to select any of the nodes in C. Hence the average number of paths containing a node, say 1, in C, is $\Theta(n^2r/nr^2) = \Theta(n/r)$. The number of edges between 1 and neighboring squares is $\Theta(nr^2)$. By symmetry, the average load on an edge incident on 1 will be $\Theta(1/r^3)$. This is true for all nodes. Hence, the average load on an edge is $O(1/r^3)$.

Next we will use this setup and Lemma 4.1 to obtain a bound on the second largest eigenvalue using a modified version of Poincare's inequality stated below.

Lemma 4.2. *Consider a natural random walk on a graph $G = (\{1, \ldots, n\}, E)$ with Γ as the set of all possible paths on all node pairs. Let γ_* be the maximum path length (among all paths and over all node pairs), d_* be the maximum node degree, and $|E|$ be total number of edges. Let, according to some probability distribution on Γ, the maximum average load on any of the edges be b, i.e. on average no edge belongs to more than b paths. Then, the second largest eigenvalue, λ_2, is bounded above as*

$$\lambda_2 \leq 1 - \left(\frac{2|E|}{d_*^2\gamma_* b}\right) \tag{4.15}$$

The proof of this Lemma can be found in the Appendix.

From Lemmas 2.1, 4.1, 4.2 and the fact that all paths are of length at most $\Theta(1/r)$, we obtain that the second largest eigenvalue corresponding to the natural random walk on $G^2(n, r)$ is bounded above as:

$$\begin{aligned}\lambda_2 &\leq 1 - \Theta\left(\frac{n^2r^2}{n^2r^4r^{-4}}\right) \\ &= 1 - \Theta(r^2)\end{aligned} \tag{4.16}$$

We would like to note that, for mixing time, we need to show that the smallest eigenvalue (which can be negative), is also $\Theta(r^2)$ away from -1. One well-known way to avoid this difficulty is the following: modify transition probabilities as $Q = \frac{1}{2}(I + P)$. Q and P have the same stationary distribution. By definition, Q has all non-negative eigenvalues, and $\lambda_2(Q) = \frac{1}{2}(1 + \lambda_2(P))$. Thus, the mixing time of the random walk corresponding to Q is governed by $\lambda_2(P)$, and is therefore $\Theta(r^{-2}\log n)$. This random walk Q is the modified natural random walk in Theorem 1.2(b).

Thus, from Lemma 1.1 and (4.16), the proof of Theorem 1.2(b) for $G^2(n, r)$ follows. In general, the above argument can be carried out similarly for $d > 2$ completing the proof of Theorem 1.2(b).

5 Conclusion

We studied the scaling of mixing times for the fastest mixing reversible random walk and modified natural random walks for $G^d(n, r)$. We found that both of them have mixing time of the same order, $\Theta(r^{-2}\log n)$ for all d. In fact these scaling results apply not just for $G^d(n, r)$ constructed with the Euclidean norm, but for any L_p norm.

The methods used in this paper to compute mixing times are novel and we strongly believe that they will be useful in other contexts.

References

[BBFM03] L. Booth, J. Bruck, M. Franceschetti, and R. Meester. Covering algorithms, continuum percolation, and the geometry of wireless networks. *The Annals of Applied Probability*, 13(2):722–41, 2003.

[BDX04] S. Boyd, P. Diaconis, and L. Xiao. Fastest mixing Markov chain on a graph. To appear in *SIAM Review*, problems and techniques section, 2004. Available at `www.stanford.edu/~boyd/fmmc.html`.

[BGPS04] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah. Gossip algorithms: Design, analysis and applications. 2004. Submitted, available at `www.stanford.edu/~devavrat/infocom05.ps`.

[Bol01] B. Bollobas. *Random Graphs*. Cambridge University Press, 2001.

[DS91] P. Diaconis and D. Stroock. Geometric bounds for eigenvalues of Markov chains. *The Annals of Applied Probability*, 1(1):36–61, 1991.

[GK00] P. Gupta and P. R. Kumar. The capacity of wireless networks. *IEEE Transactions on Information Theory*, 46(2):388–404, March 2000.

[GRK04] A. Goel, S. Rai, and B. Krishnamachari. Sharp threshold for monotone properties in geometric random graphs. In *Proc. Conference on Symposium on Theory of Computation*. ACM, 2004.

[Gur00] V. Guruswami. Rapidly mixing markov chains: A comparison of techniques. May 2000. Available at `cs.washington.edu/homes/venkat/pubs/papers.html`.

[HJ85] R. Horn and C. Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, UK, 1985.

[Pen03] M. Penrose. *Random geometric graphs*. Oxford studies in probability. Oxford University Press, Oxford, 2003.

[PXBD03] P. A. Parrilo, L. Xiao, S. Boyd, and P. Diaconis. Symmetry analysis of reversible markov chains. Technical report No. 2003-38, Department of Statistics, Stanford University, 2003.

6 Appendix

Here we present the proofs of Lemma 3.2 and 4.2.

LEMMA 6.1. *For $k = o(n)$, there is a random walk on G_k for which the mixing rate is $\lambda_{\max} = \cos(2\pi k/n) + \Theta(k^4/n^4)$.*

Proof. For simplicity let us assume that $2k$ divides n; it is not difficult to obtain the same results when this is not the case.

Consider the Markov chain with transition probabilities $p_0 = 0$, $p_i = \delta, i = 1, \dots, k-1$, $p_k = 1/2 - (k-1)\delta$. We will show that for a certain δ, small enough, μ_1 is indeed $\lambda_{\max}$, and is away from $\cos(2\pi k/n)$ by $\Theta(k^4/n^4)$.

For the transition matrix P^* corresponding to these probabilities, the eigenvalues are, for $m = 0, \dots, n-1$,

$$\begin{aligned}\mu_m &= 2\textstyle\sum_{i=1}^{k-1} \delta \cos(\frac{2\pi im}{n}) + 2(\frac{1}{2} - (k-1)\delta)\cos(\frac{2\pi km}{n}) \\ &= \cos(\tfrac{2\pi km}{n}) + 2\delta \textstyle\sum_{i=1}^{k-1}(\cos(\frac{2\pi im}{n}) - \cos(\frac{2\pi km}{n}))\end{aligned} \tag{6.17}$$

We want to find the smallest positive δ such that μ_1 is $\lambda_{\max}$ (this is not true, for example, for $\delta = 0$). However, we need δ to be small enough so that the residual term, $2\delta \sum_{i=1}^{k-1}(\cos(2\pi i/n) - \cos(2\pi k/n))$, is small compared to $\cos(2\pi k/n)$.

Since $k = o(n)$ and we hope that δ is small ($o(1)$), we see that the values of m for which $|\mu_m|$ is comparable to μ_1 are those values of m for which $|\cos(2\pi km/n)| = 1$. This happens for $m = \frac{n}{2k}, \frac{n}{k}, \frac{3n}{2k}, \dots, \frac{n}{2}$. (We only need consider values of m until $n/2$, since $\lambda_i = \lambda_{n-i}$.) At all odd multiples of $n/2k$, $\cos(2\pi km/n) = -1$, and for the even multiples, $\cos(2\pi km/n) = 1$. For δ to satisfy $|\mu_m| \leq \mu_1$, we must have for m an even multiple of $n/2k$,

$$\begin{aligned}&1 + 2\delta \textstyle\sum_{i=1}^{k-1}(\cos(\frac{2\pi im}{n}) - 1) \leq \\ &\cos(\tfrac{2\pi k}{n}) + 2\delta \textstyle\sum_{i=1}^{k-1}(\cos\frac{2\pi i}{n} - \cos\frac{2\pi k}{n});\end{aligned} \tag{6.18}$$

and for m an odd multiple of $n/2k$

$$\begin{aligned}&|-1 + 2\delta \textstyle\sum_{i=1}^{k-1}(\cos(\frac{2\pi im}{n}) + 1)| \leq \cos(\frac{2\pi k}{n}) + \\ &\qquad 2\delta \textstyle\sum_{i=1}^{k-1}(\cos\frac{2\pi i}{n} - \cos\frac{2\pi k}{n}), \\ &\Rightarrow 1 - 2\delta \textstyle\sum_{i=1}^{k-1}(\cos(\frac{2\pi im}{n}) + 1) \leq \cos(\frac{2\pi k}{n}) + \\ &\qquad 2\delta \textstyle\sum_{i=1}^{k-1}(\cos\frac{2\pi i}{n} - \cos\frac{2\pi k}{n}).\end{aligned} \tag{6.19}$$

From (6.18), we see that δ must satisfy

$$\delta \geq \frac{\frac{1}{2}(1 - \cos(\frac{2\pi k}{n})}{(k-1)(1 - \cos(\frac{2\pi k}{n})) + \sum_{i=1}^{k-1}\cos(\frac{2\pi i}{n}) + \cos(\frac{2\pi im}{n})} \tag{6.20}$$

for m an odd multiple of $n/2k$, and from (6.19),

$$\delta \geq \frac{\frac{1}{2}(1 - \cos(\frac{2\pi k}{n})}{(k-1)(1 - \cos(\frac{2\pi k}{n})) + \sum_{i=1}^{k-1}\cos(\frac{2\pi i}{n}) - \cos(\frac{2\pi im}{n}))} \tag{6.21}$$

for m a multiple of n/k. So δ can be only as small as the maximum over the specified m of all of these right-hand sides.

Note that the only term dependent on m in each of these expressions is $\sum_{i=1}^{k-1}\cos(2\pi im/n)$. For $m = pn/2k$, p odd,

$$\sum_{i=1}^{k-1}\cos(2\pi im/n) = \sum_{i=1}^{k-1}\cos(\pi ip/k) = 0, \tag{6.22}$$

since $\cos(\pi ip/k) = -\cos(\pi(k-i)p/k)$ for odd p, and if k is even, $\cos(\pi kp/2k) = 0$ also. For $m = qn/k$,

$$\sum_{i=1}^{k-1}\cos(2\pi im/n) = \sum_{i=1}^{k}\cos(2\pi iq/k) - 1 = -1 \tag{6.23}$$

since $\sum_{i=1}^{k}\cos(2\pi iq/k) = 0$ (sum of real parts of the kth roots of unity).

So $\delta = \Theta(k/n^2)$, and returning to (6.17), we see that the residual term in μ_1 is of order $(k/n^2)(k^3/n^2)$, *i.e.*, k^4/n^4, while $\cos(2\pi k/n) \approx 1 - 2\pi^2k^2/n^2$. So the difference between $\lambda_{\max}$ and $\cos(2\pi k/n)$ is $\Theta(k^4/n^4)$.

LEMMA 6.2. *Consider a natural random walk on a graph $G = (\{1, \dots, n\}, E)$ with Γ as the set of all* *possible paths on all node pairs. Let γ_* be the maximum path length (among all paths and over all node pairs), d_* be the maximum node degree, and $|E|$ be total number of edges. Let, according to some probability distribution on Γ, the maximum average load on any of the edges be b, i.e. on average no edge belongs to more than b paths. Then, the second largest eigenvalue, λ_2, is bounded above as*

$$\lambda_2 \leq 1 - \left(\frac{2|E|}{d_*^2 \gamma_* b}\right) \tag{6.24}$$

Proof. The proof follows from a modification of Poincare's inequality (Proposition 1 [DS91]). Before proceeding to the proof, we introduce some notation.

Let $\phi : \{1, \ldots, n\} \to \mathbf{R}$ be a real valued function on the n nodes. Let $\pi = (\pi(i))_{\{1 \leq i \leq n\}}$ denote the equilibrium distribution of the random walk. Let d_i be the degree of node i, then it is well known that $\pi(i) = \frac{d_i}{2|E|} \leq \frac{d_*}{2|E|}$. For node pair (i, j), let

$$Q(i,j) = \pi(i)P_{ij} = \pi(j)P_{ji} = 1/2|E|.$$

Define the quadratic form of function ϕ as

$$\mathcal{E}(\phi, \phi) = \frac{1}{2}\sum_{i,j}(\phi(i) - \phi(j))^2 Q(i,j).$$

Let the variance of ϕ with respect to π be

$$V(\phi) = \frac{1}{2}\sum_{i,j}(\phi(i) - \phi(j))^2 \pi(i)\pi(j).$$

For a directed edge e from $i \to j$, define $\phi(e) = \phi(i) - \phi(j)$ and $Q(e) = Q(i,j)$. First, consider one collection of path $\gamma = (\gamma_{ij})$. Define

$$|\gamma_{ij}|_Q = \sum_{e \in \gamma_{ij}} Q(e)^{-1}.$$

Then, under the natural random walk,

$$|\gamma_{ij}|_Q = |\gamma_{ij}|(2|E|), \tag{6.25}$$

where $|\gamma_{ij}|$ is the length of the path γ_{ij}.

$$\begin{aligned}
V(\phi) &= \frac{1}{2}\sum_{i,j}(\phi(i) - \phi(j))^2 \pi(i)\pi(j) \\
&\overset{(a)}{=} \frac{1}{2}\sum_{i,j}\left(\sum_{e \in \gamma_{ij}}\left(\frac{Q(e)}{Q(e)}\right)^{1/2}\phi(e)\right)^2 \pi(i)\pi(j) \\
&\overset{(b)}{\leq} \frac{1}{2}\sum_{i,j}|\gamma_{ij}|_Q \pi(i)\pi(j)\sum_{e \in \gamma_{ij}} Q(e)\phi(e)^2 \\
&\leq \left(\frac{d_*}{2|E|}\right)^2 \frac{1}{2}\sum_e Q(e)\phi(e)^2 \sum_{\gamma_{ij} \ni e}|\gamma_{ij}|_Q \\
&\overset{(c)}{=} \left(\frac{d_*^2}{2|E|}\right)\frac{1}{2}\sum_e Q(e)\phi(e)^2 \sum_{\gamma_{ij} \ni e}|\gamma_{ij}| \\
&\overset{(d)}{\leq} \left(\frac{d_*^2}{2|E|}\right)\gamma_* \frac{1}{2}\sum_e Q(e)\phi(e)^2 b(\gamma, e), \qquad (6.26)
\end{aligned}$$

where $b(\gamma, e)$ denotes the number of paths passing through edge e under $\gamma = (\gamma_{ij})$. (a) follows by using $\pi(i) \approx 1/n$ for all i and adding as well as subtracting values of ϕ on the nodes of path γ_{ij} for all node pair (i, j) for a give path-set $\gamma = (\gamma_{ij})$. (b) follows by Cauchy-Schwartz inequality. (c) follows from (6.25), and (d) follows from the fact that all path length are smaller than γ_*.

Note that in (6.26), $b(\gamma, e)$ is the only path dependent term. Hence under a probability distribution on Γ (i.e. set of all paths), in (6.26), $b(\gamma, e)$ can be replaced by $b(e)$ where

$$b(e) = \sum_{\gamma \in \Gamma} \Pr(\gamma) b(\gamma, e).$$

Let $b = \max_e b(e)$. Then,

$$\begin{aligned}
V(\phi) &\leq \left(\frac{d_*^2}{2|E|}\right)\gamma_* \frac{1}{2}\sum_e Q(e)\phi(e)^2 b, \qquad (6.27) \\
&= \left(\frac{d_*^2 \gamma_* b}{2|E|}\right)\mathcal{E}(\phi, \phi). \qquad (6.28)
\end{aligned}$$

The minimax characterization of eigenvalue ([HJ85], page 176), gives a bound on the second largest eigenvalue as

$$\lambda_2 = \sup\{1 - \frac{\mathcal{E}(\phi, \phi)}{V(\phi)} : \phi \text{ a non-constant}\}. \tag{6.29}$$

From (6.28) and (6.29), the statement of the Lemma follows.