Internet Mathematics Vol. 8, No. 4: 381–401

Digraph Laplacian and the Degree of Asymmetry

Yanhua Li and Zhi-Li Zhang

Abstract. In this paper we extend and generalize the standard spectral graph theory (or random-walk theory) on undirected graphs to digraphs. In particular, we introduce and define a normalized *digraph Laplacian* (*Diplacian* for short) Γ for digraphs, and prove that (1) its Moore–Penrose pseudoinverse is the discrete Green's function of the Diplacian matrix as an operator on digraphs, and (2) it is the normalized fundamental matrix of the Markov chain governing random walks on digraphs. Using these results, we derive a new formula for computing hitting and commute times in terms of the Moore–Penrose pseudoinverse of the Diplacian, or equivalently, the singular values and vectors of the Diplacian.

Furthermore, we show that the Cheeger constant defined in [Chung 05] is intrinsically a quantity associated with undirected graphs. This motivates us to introduce a metric, the largest singular value of the skewed Laplacian $\nabla = (\Gamma - \Gamma^T)/2$, to quantify and measure the *degree of asymmetry* in a digraph. Using this measure, we establish several new results, such as a tighter bound than that in [Chung 05] on the Markov chain mixing rate, and a bound on the second-smallest singular value of Γ.

1. Introduction

Graphs arising from many applications such as web and online social networks are directed, where the direction of the links contains crucial information.

© Taylor & Francis Group, LLC
ISSN: 1542-7951 print

Random walks are frequently used to model certain dynamic processes on directed or undirected graphs, for example, to reveal important network structural information, such as the importance of nodes using the PageRank algorithm [Brin and Page 98] and community structures using the spectral clustering algorithm [Li et al. 12], or to study ways to explore complex networks efficiently.

Random walks on undirected graphs have been extensively studied and are well understood (see [Lovász 93, Li et al. 11]). They are closely related to spectral graph theory [Chung 06], which has produced powerful tools for studying many important properties of undirected graphs that are of both theoretical and practical significance. Well-known results include bounds on the Cheeger constant and the mixing rate in terms of the second-smallest eigenvalue of the graph Laplacian. On the other hand, there are relatively few similar studies on directed graphs; see, for example, [Chung 05, Zhou et al. 05], where the authors circumvent the "directedness" of digraphs by converting them into undirected graphs through symmetrization. Moreover, some initial attempts [Li and Zhang 10a, Li and Zhang 12] have been made to extend random-walk theory to directed graphs with applications in wireless networking.

In this paper we develop a spectral digraph theory. We introduce the notion of the *Diplacian*, a direct generalization of the graph Laplacian for undirected graphs, denoted by Γ. Instead of using the node degrees as in the case of undirected graphs, the Diplacian is defined using stationary probabilities of the Markov chain governing random walks on digraphs. Furthermore, instead of relying on the positive semidefiniteness of the graph Laplacian matrix for undirected graphs, we establish a key connection between the Diplacian Γ and its Moore–Penrose pseudoinverse [Horn and Johnson 85], denoted by Γ^+, and use the properties of this connection to prove several parallel results for random walks on digraphs. In particular, we show that (1) the Moore–Penrose pseudoinverse Γ^+ of the Diplacian is exactly the discrete Green's function of the Diplacian matrix Γ, acting as an operator on digraphs [Chung and Yau 00], and (2) Γ^+ is the normalized *fundamental matrix* [Aldous and Fill 99] of the Markov chain governing random walks on digraphs.

Based on the connection between Γ^+ and the fundamental matrix, we show how hitting and commute times can be directly computed in terms of the singular values and vectors of the Diplacian. This yields a more direct and efficient way to compute hitting and commute times than existing methods based on the fundamental matrix. More generally, our results suggest a "spectral digraph theory" in which graph properties can be studied using the singular values of the Diplacian in place of the eigenvalues of the Laplacian. In particular, our theory of random walks on digraphs subsumes the existing theory of random walks on undirected graphs as a special case.

Furthermore, we show that the well-known Cheeger constant—generalized to digraphs in [Chung 05]—is fundamentally a quantity intrinsic to undirected graphs, since there are infinitely many digraphs with the same symmetrized undirected graph. Hence, bounds based on the eigenvalues of the symmetrized graph Laplacian do not uniquely capture the properties of digraphs. This leads us to introduce the *degree of asymmetry* to capture the overall "directedness" of digraphs, formally defined as follows: we express a Diplacian Γ in terms of a *symmetric part* $\bar{\mathcal{L}} = (\Gamma + \Gamma^T)/2$ and a *skew-symmetric* part $\nabla = (\Gamma - \Gamma^T)/2$, where $\bar{\mathcal{L}}$ is the symmetrized graph Laplacian for digraphs introduced in [Chung 05], and ∇ is referred to as the skewed Laplacian for digraphs. The largest singular value of ∇, $\delta_{\max}$, is referred to as the *degree of asymmetry*, which provides a quantitative measure of the asymmetry in digraphs. Many key properties of digraphs can then be bounded by the eigenvalues of $\bar{\mathcal{L}}$ and the degree of asymmetry. For instance, by accounting for the asymmetry of digraphs, we are able to obtain a tighter bound than that in [Chung 05] on the irreversible Markov chain mixing rate.

2. Preliminaries: Random Walks on Undirected Graphs

We use a triple $G = (V, E, A)$ to denote an undirected and weighted graph on the node set $V = \{1, 2, \ldots, n\}$. The $n \times n$ nonnegative weight matrix $A = [a_{ij}]$ is symmetric, and it is defined in such a way that $a_{ij} = a_{ji} > 0$ if $(i, j) \in E$, and $a_{ij} = a_{ji} = 0$ otherwise. For $1 \le i \le n$, the degree of node i is $d_i = \sum_{j=1}^{n} a_{ij}$. The volume of G, denoted by $\mathrm{vol}(G)$, is defined as the sum of all node degrees, $d = \sum_{i=1}^{n} d_i$, that is, $\mathrm{vol}(G) = d$.

A random walk on G is a Markov chain defined on G with transition probability matrix $P = [p_{ij}]$, where $p_{ij} = a_{ij}/d_i$. Let $D = \mathrm{diag}[d_i]$ be a diagonal matrix of node degrees. Then $P = D^{-1}A$. Without loss of generality, we assume that the undirected graph G is connected, namely, any node can reach any other node in G. Then it can be shown (see [Aldous and Fill 99]) that the Markov chain is irreducible, and there exists a unique stationary distribution $\{\phi_1, \phi_2, \ldots, \phi_n\}$. Let $\phi = [\phi_i]_{1 \le i \le n}$ be the column vector of the stationary probabilities. Then $\phi^T P = \phi^T$, where the superscript T represents the vector or matrix transpose. Furthermore, this Markov chain on G is *reversible*, namely,

$$\phi_i p_{ij} = \phi_j p_{ji} \quad \text{for every } i, j, \tag{2.1}$$

and

$$\phi_i = \frac{d_i}{\sum_k d_k} = \frac{d_i}{d}, \quad i = 1, 2, \ldots, n. \tag{2.2}$$

Following [Chung 06], we will use the normalized graph Laplacian instead of the unnormalized version $L = D - A$. Given an undirected G, the normalized graph Laplacian of G, which is also called the normalized Laplacian matrix of G, is defined as follows:

$$\mathcal{L} = D^{-1/2}(D - A)D^{-1/2} = D^{1/2}(I - P)D^{-1/2}. \tag{2.3}$$

A key property of the graph Laplacian for an undirected graph is that $\mathcal{L}$ is symmetric and positive semidefinite [Horn and Johnson 85]. Hence all eigenvalues of $\mathcal{L}$ are nonnegative real numbers. In particular, for a connected undirected graph G, $\mathcal{L}$ has rank $n-1$ and has exactly one zero eigenvalue as its smallest eigenvalue. Let $\lambda_1 = 0 < \lambda_2 \leq \cdots \leq \lambda_n$ be the n eigenvalues of $\mathcal{L}$ arranged in increasing order, and let μ_i, $1 \leq i \leq n$, be the corresponding eigenvectors of unit norm. In particular, one can show that the column eigenvector μ_1 of $\mathcal{L}$ associated with the eigenvalue $\lambda_1 = 0$ is given by

$$\mu_1 = \phi^{1/2} = \left[\sqrt{\phi_i}\right] = \left[\frac{\sqrt{d_i}}{\sqrt{d}}\right]. \tag{2.4}$$

Define Λ=diag$[\lambda_1, \ldots, \lambda_n]$, the diagonal matrix formed by the eigenvalues, and $U = [\mu_1, \ldots, \mu_n]$, an orthonormal matrix formed by the eigenvectors of $\mathcal{L}$, where $UU^T = U^T U = I$. It is easy to see that the graph Laplacian $\mathcal{L}$ admits an eigendecomposition [Horn and Johnson 85], namely $\mathcal{L} = U\Lambda U^T$. Using the eigenvalues and eigenvectors of $\mathcal{L}$, we can compute the hitting times and commute times using the following formulas [Lovász 93]:

$$H_{ij} = \sum_{k>1} \frac{d}{\lambda_k}\left(\frac{\mu_{kj}^2}{d_j} - \frac{\mu_{ki}\mu_{kj}}{\sqrt{d_i d_j}}\right) \tag{2.5}$$

and

$$C_{ij} = \sum_{k>1} \frac{d}{\lambda_k}\left(\frac{\mu_{ki}}{\sqrt{d_i}} - \frac{\mu_{kj}}{\sqrt{d_j}}\right)^2, \tag{2.6}$$

where μ_{kj} is the jth entry of the column vector μ_k.

3. Random Walk Theory on Digraphs

In this section, we develop the random walk theory for digraphs. In particular, we generalize the graph Laplacian defined for undirected graphs, and introduce the Diplacian matrix. We prove that the Moore–Penrose pseudoinverse of this Diplacian is exactly equal to a normalized version of the fundamental matrix of the Markov chain governing random walks on digraphs and show that it is also

the Green's function of the Diplacian. Using these connections, we illustrate how hitting and commute times of random walks on digraphs can be directly computed using the singular values and vectors of the Diplacian. We also show that when the underlying graph is undirected, our results reduce to the well-known results for undirected graphs. Hence, our theory includes undirected graphs as a special case.

3.1. Random Walks on Directed Graphs and the Fundamental Matrix

As alluded to earlier, random walks can be defined not only on undirected graphs, but also on digraphs. Let $G = (V, E, A)$ be a weighted digraph defined on the vertex set $V = \{1, 2, \ldots, n\}$, where A is a nonnegative but generally asymmetric weight matrix such that $a_{ij} > 0$ if and only if the directed edge (or arc) (i, j) belongs to E. As before, we will refer to A simply as the adjacency matrix of G. For $i = 1, 2, \ldots, n$, we define the *out-degree* of vertex i, $d_i^+ = \sum_{j=1}^n a_{ij}$, and the *in-degree* of vertex i, $d_i^- = \sum_{j=1}^n a_{ji}$. In general, $d^+ \neq d^-$. However, we have $d = \sum_{i=1}^n d_i^+ = \sum_{i=1}^n d_i^- = \sum_{i=1}^n \sum_{j=1}^n a_{ij}$. As before, we refer to d as the volume of the directed graph G, that is, $\mathrm{vol}(G) = d$. For conciseness, in the following, unless otherwise stated, we refer to the out-degree of a vertex simply as its degree, and use d_i for d_i^+.

Let $D = \mathrm{diag}[d_i]$ be a diagonal matrix of the vertex out-degrees, and define $P = D^{-1}A$. Then $P = [p_{ij}]$ is the transition probability matrix of the Markov chain associated with random walks on G, where at each vertex i, a random walk has probability $p_{ij} = a_{ij}/d_i$ of transiting from vertex i to vertex j if $(i, j) \in E$. We assume that G is *strongly connected*, namely, there is a directed path from each vertex i to every other vertex j. Then the Markov chain P is irreducible, and has a unique stationary probability distribution $\{\phi_i\}$, where $\phi_i > 0$, $1 \leq i \leq n$, namely, $\phi^T P = \phi^T$, where $\phi = [\phi_1, \ldots, \phi_n]^T$ is the column vector of stationary probabilities. In contrast to the case of undirected graphs, the Markov chain associated with random walks on directed graphs is generally irreversible, and (2.1) and (2.2) do not hold on undirected graphs.

For random walks on directed graphs, quantities such as hitting times and commute times can be defined exactly as in the case of undirected graphs. However, since the normalized Laplacian matrix L is so far defined only for undirected graphs, we cannot use the relations (2.5) and (2.6) to compute hitting times and commute times for random walks on directed graphs. On the other hand, using results from standard Markov chain theory, we can express the hitting times and commute times in terms of the fundamental matrix. In [Aldous and Fill 99], the authors define the fundamental matrix $Z = [z_{ij}]$ for an irreducible Markov chain

with transition probability matrix P:

$$z_{ij} = \sum_{t=0}^{\infty} \left(p_{ij}^{(t)} - \phi_j\right), \quad 1 \le i, j \le n, \tag{3.1}$$

where $p_{ij}^{(t)}$ is the (i,j)th entry in the t-step transition probability matrix $P^t = P \cdots P$ (t-fold product).

Let $\Phi = \mathrm{diag}[\phi_i]$ be the diagonal matrix containing the stationary probabilities ϕ_i on the diagonal and let $J = [J_{ij}]$ be the matrix of all 1's, that is, $J_{ij} = 1$ for $1 \le i, j \le n$. We can express Z alternatively as the sum of an infinite matrix series:

$$Z = \sum_{t=0}^{\infty} (P^t - J\Phi) = \sum_{t=0}^{\infty} (P^t - \mathbf{1}\phi^T), \tag{3.2}$$

where $\mathbf{1} = [1, \ldots, 1]^T$ is the column vector of all 1's. Hence we have $J = \mathbf{1} \cdot \mathbf{1}^T$ and $\mathbf{1}^T \Phi = \phi^T$.

While the physical meaning of the fundamental matrix Z may not be obvious from its definition (3.1) or (3.2), it plays a crucial role in computing various quantities related to random walks, or more generally, various stopping-time properties of Markov chains [Aldous and Fill 99]. For instance, the hitting times and commute times of random walks on a directed graph can be expressed in terms of Z as follows (see [Aldous and Fill 99]):

$$H_{ij} = \frac{z_{jj} - z_{ij}}{\phi_j} \tag{3.3}$$

and

$$C_{ij} = \frac{z_{jj} - z_{ij}}{\phi_j} + \frac{z_{ii} - z_{ji}}{\phi_i}. \tag{3.4}$$

In (3.1) and (3.2), the fundamental matrix Z is defined as an infinite sum. We show that Z in fact satisfies a simple relation (3.5), thus can be computed directly using the standard matrix inverse.

Theorem 3.1. *Let P be the transition probability matrix for an irreducible Markov chain. Then its corresponding fundamental matrix Z as defined in* (3.1) *satisfies the following relation:*

$$Z + J\Phi = (I - P + J\Phi)^{-1}. \tag{3.5}$$

Proof. Note that $J\Phi = \mathbf{1}\phi^T$. From $\phi^T P = \phi^T$ and $P\mathbf{1} = \mathbf{1}$, we have $J\Phi P = J\Phi$ and $PJ\Phi = J\Phi$. Using these two relations, it is easy to prove the following equation

by induction:

$$P^m - J\Phi = (P - J\Phi)^m, \text{ for every integer } m > 0. \tag{3.6}$$

Substituting this into (3.2) yields Theorem 3.1. □

Since undirected graphs are a special case of directed graphs, (3.3) and (3.4) provide an alternative way to compute hitting times and commute times for random walks on connected undirected graphs. In this paper we will show that (2.5) and (2.6) are in fact equivalent to (3.3) and (3.4).

3.2. Diplacian and Green's Function for Digraphs

We now generalize the existing spectral graph theory defined for undirected graphs to digraphs by introducing the Diplacian—an appropriately generalized Laplacian matrix—for strongly connected diagraphs. Let $G = (V, E, A)$ be a strongly connected weighted digraph defined on the vertex set $V = \{1, 2, \ldots, n\}$, where in general, the weight (or adjacency) matrix A is asymmetric. A major technical difficulty in dealing with digraphs is that if one naively extends the normalized Laplacian matrix $\mathcal{L} = D^{-1/2}(D - A)D^{-1/2}$ or its unnormalized version $L = D - A$ defined for undirected graphs to digraphs, then $\mathcal{L}$ is in general asymmetric; thus the nice properties such as positive semidefiniteness of $\mathcal{L}$ no longer hold. Past attempts at generalizing spectral graph theory to digraphs have simply symmetrized $\mathcal{L}$, for example by introducing a symmetric matrix $\bar{\mathcal{L}} = I - (\Phi^{1/2}P\Phi^{-1/2} + \Phi^{-1/2}P^T\Phi^{1/2})/2$ [Chung 05, Zhou et al. 05]. Unfortunately, as will be shown in Section 4, such a symmetrized $\bar{\mathcal{L}}$ does not directly capture the unique characteristic of the random walk on the digraph as defined earlier, since a set of diagraphs can have the same $\bar{\mathcal{L}}$.

For a strongly connected digraph G, let $\Phi^{1/2} = \text{diag}[\sqrt{\phi_i}]$. We define the normalized digraph Laplacian matrix[1] (*Diplacian* for short) $\Gamma = [\Gamma_{ij}]$ for the graph G as follows.

Definition 3.2. The *Diplacian* Γ is defined as

$$\Gamma = \Phi^{1/2}(I - P)\Phi^{-1/2}. \tag{3.7}$$

[1] An unnormalized digraph Laplacian is defined as $L = \Phi(I - P)$ in [Boley et al. 11].

Namely, for $1 \le i, j \le n$,

$$\Gamma_{ij} = \begin{cases} 1 - p_{ii} & \text{if } i = j, \\ -\phi_i^{1/2} p_{ij} \phi_j^{-1/2} & \text{if } (i,j) \in E, \\ 0 & \text{otherwise.} \end{cases} \tag{3.8}$$

Treating this Diplacian matrix Γ as an asymmetric operator on a digraph G, we now define the discrete Green's function $\mathcal{G}$ without boundary conditions for digraphs in exactly the same manner as for undirected graphs [Chung and Yau 00], namely, $\mathcal{G}$ is a matrix with its entries indexed by vertices i and j that satisfies the following conditions:

$$[\mathcal{G}\Gamma]_{i,j} = I_{i,j} - \sqrt{\phi_i \phi_j}, \quad 1 \le i, j \le n, \tag{3.9}$$

and expressed in the matrix form

$$\mathcal{G}\Gamma = I - \phi^{1/2} \phi^{1/2^T}. \tag{3.10}$$

In the following we will show that $\mathcal{G}$ is precisely Γ^+, the pseudoinverse of the Laplacian operator Γ on the digraph G. Furthermore, we will relate Γ^+ directly to the fundamental matrix Z of the Markov chain associated with random walks on the digraph G. Before we establish the main result of this paper, we first introduce a bit more notation and then prove the following useful lemma.

Lemma 3.3. *Define $\mathcal{Z} = \Phi^{1/2} Z \Phi^{-1/2}$ as the normalized fundamental matrix and $\mathcal{J} = \Phi^{1/2} J \Phi^{1/2} = \phi^{1/2} \phi^{1/2^T}$. The following relations regarding $\mathcal{Z}$ and $\mathcal{J}$ hold:*

(i) $\mathcal{J} = \mathcal{J}^2$,

(ii) $\mathcal{J}\Gamma = \Gamma\mathcal{J} = \mathcal{J}\mathcal{Z} = \mathcal{Z}\mathcal{J} = \mathbf{0}_{n \times n}$,

(iii) $\Gamma\phi^{1/2} = \mathcal{Z}\phi^{1/2} = \mathbf{0}$, $\phi^{1/2^T}\Gamma = \phi^{1/2^T}\mathcal{Z} = \mathbf{0}^T$.

Proof. These relations can be established using the facts that $J = \mathbf{1}\mathbf{1}^T$, $\mathbf{1}^T\Phi = \phi^T$, $\phi^T J = \mathbf{1}^T$, $\Phi J = \phi\mathbf{1}^T$, $J\Phi J = J$, $\phi^T(I - P) = \mathbf{0}^T$, $(I - P)\mathbf{1} = \mathbf{0}$, $\phi^T Z = \mathbf{0}^T$, and $Z\mathbf{1} = \mathbf{0}$. The last four equalities imply that the matrices $I - P$ and Z have the same left and right eigenvectors, ϕ and $\mathbf{1}$, corresponding to the eigenvalue 0. □

We are now in a position to prove the main theorem of the paper, which states that the Green's function for the Diplacian is exactly its Moore–Penrose pseudoinverse and that it is equal to the normalized fundamental matrix, namely $\mathcal{G} = \Gamma^+ = \mathcal{Z}$.

Theorem 3.4. (Diplacian matrix and Green's function for digraphs.) *Let $G = (V, E, A)$ be a strongly connected digraph with the normalized fundamental matrix $\mathcal{Z}$ as defined in Lemma 3.3, and the Diplacian matrix Γ as defined in* (3.7)*. Then $\mathcal{Z} = \Gamma^+$ is the pseudoinverse of the Diplacian matrix Γ. Furthermore, $\mathcal{Z}$ is the discrete Green's function for Γ, namely,*

$$\mathcal{Z}\Gamma = I - \Phi^{1/2} J \Phi^{1/2} = I - \phi^{1/2}\phi^{1/2^T}, \tag{3.11}$$

where J is the matrix of all 1's and $\phi^{1/2} = [\phi_1^{1/2}, \ldots, \phi_n^{1/2}]^T$ is a column vector.

Proof. From equation (3.5) in Theorem 3.1, we have

$$\mathcal{Z} + \mathcal{J} = (\Gamma + \mathcal{J})^{-1}. \tag{3.12}$$

Multiplying (3.12) on the right by $\Gamma + \mathcal{J}$, and using Lemma 3.3, it is easy to see that

$$\mathcal{Z}\Gamma = I - \mathcal{J}, \tag{3.13}$$

which establishes that $\mathcal{Z}$ is the Green's function of the Diplacian Γ. Similarly, by multiplying (3.12) on the left by $\Gamma + \mathcal{J}$, we can likewise prove that $\Gamma\mathcal{Z} = I - \mathcal{J}$.

Hence $\mathcal{Z}\Gamma = \Gamma\mathcal{Z} = I - \mathcal{J}$ is a real symmetric matrix, which implies $(\Gamma\mathcal{Z})^T = \Gamma\mathcal{Z}$ and $(\mathcal{Z}\Gamma)^T = \mathcal{Z}\Gamma$. Furthermore, since $\mathcal{J}\mathcal{Z} = 0$, equation (3.13) yields $\mathcal{Z}\Gamma\mathcal{Z} = \mathcal{Z}$. Similarly, since $\Gamma\mathcal{J} = \Phi^{1/2}(I - P)J\Phi^{1/2} = 0$, equation (3.13) yields $\Gamma\mathcal{Z}\Gamma = \Gamma$. This establishes that $\mathcal{Z}$ satisfies the four conditions of the matrix pseudoinverse. Hence $\mathcal{Z}$ is also the Moore–Penrose pseudoinverse of Γ. Therefore, $\mathcal{G} = \mathcal{Z} = \Gamma^+$. □

3.3. Computing Hitting and Commute Times for Digraphs Using the Diplacian

Using the relationship between the Diplacian Γ, its pseudoinverse Γ^+, and the normalized fundamental matrix $\mathcal{Z}$, we can now express the hitting times and commute times of random walks on digraphs in terms of Γ^+, or alternatively in terms of the singular values and singular vectors of the Diplacian matrix Γ.

From $Z = \Phi^{-1/2}\mathcal{Z}\Phi^{1/2} = \Phi^{-1/2}\Gamma^+\Phi^{1/2}$, and using (3.3) and (3.4), we can compute the hitting times and commute times for random walks on digraphs directly in terms of the entries of Γ^+.

Theorem 3.5. (Computing hitting and commute times using Γ^+.) *The hitting times and commute times of random walks on strongly connected digraphs can be computed using the pseudoinverse of the Diplacian matrix Γ^+ as follows:*

$$H_{ij} = \frac{\Gamma^+_{jj}}{\phi_j} - \frac{\Gamma^+_{ij}}{\sqrt{\phi_i\phi_j}} \tag{3.14}$$

and

$$C_{ij} = H_{ij} + H_{ji} = \frac{\Gamma^+_{jj}}{\phi_j} + \frac{\Gamma^+_{ii}}{\phi_i} - \frac{\Gamma^+_{ij}}{\sqrt{\phi_i\phi_j}} - \frac{\Gamma^+_{ji}}{\sqrt{\phi_i\phi_j}}, \tag{3.15}$$

where Γ^+_{ij} *is the* (i,j)*th entry of* Γ^+*, and* ϕ_i *is the stationary probability of vertex* i*.*

For undirected graphs, we show that (3.14) and (3.15) reduce to (2.5) and (2.6). This can be seen from the fact that for undirected graphs, $\Gamma = \mathcal{L}$ is symmetric and positive semidefinite. Hence the singular value decomposition of $\mathcal{L}$ is the same as the eigendecomposition of $\mathcal{L}$.

For digraphs we can express Γ^+_{ij} directly in terms of the singular values and left/right singular vectors of the generally asymmetric Diplacian matrix Γ. Let σ_i, u_i, and v_i be the ith singular value and the corresponding left and right singular vectors of Γ arranged in increasing order, where $\|u_i\|_2 = 1$ and $\|v_i\|_2 = 1$, $i = 1, 2, \ldots, n$. In particular, $0 = \sigma_1 < \sigma_2 \leq \cdots \leq \sigma_n$. Hence $\Gamma = U\Sigma V^T$, where $\Sigma = \mathrm{diag}[\sigma_i]$, $U = [u_1, \ldots, u_n]$, $V = [v_1, \ldots, v_n]$, $UU^T = I$, $VV^T = I$, and $\Gamma^+ = V\Sigma^+U^T$, where $\Sigma^+ = \mathrm{diag}[\sigma_i^+]$. Therefore, $\Gamma^+_{ij} = \sum_{k>1} \frac{1}{\sigma_k} v_{ki}u_{kj}$. Plugging these equations into (3.14) and (3.15), we can express the hitting times and commute times for random walks in a digraph in terms of the singular values and left and right singular vectors of the Diplacian matrix Γ as follows:

$$H_{ij} = \sum_{k>1} \frac{1}{\sigma_k} \left(\frac{v_{kj}u_{kj}}{\phi_j} - \frac{v_{ki}u_{kj}}{\sqrt{\phi_i\phi_j}} \right) \tag{3.16}$$

and

$$C_{ij} = \sum_{k>1} \frac{1}{\sigma_k} \left(\frac{v_{kj}u_{kj}}{\phi_j} + \frac{v_{ki}u_{ki}}{\phi_i} - \frac{v_{ki}u_{kj}}{\sqrt{\phi_i\phi_j}} - \frac{v_{kj}u_{ki}}{\sqrt{\phi_j\phi_i}} \right). \tag{3.17}$$

4. Degree of Asymmetry, Generalized Cheeger Constant, and Bounds on the Mixing Rate

In this section we explore the relation between the Diplacian Γ and the symmetrized Laplacian $\bar{\mathcal{L}}$. We first show that the symmetrized Laplacian matrix $\bar{\mathcal{L}}$ and the Cheeger constant $h(G)$ as defined in [Chung 05] are in a sense primarily determined by an undirected graph associated with random walks with transition probability matrix $\bar{P} = (P + \Phi^{-1}P^T\Phi)/2$ and thus cannot capture the unique characteristics of each individual digraph. As a result, we investigate two questions: (1) how can the "degree of asymmetry" of a digraph be quantified and

measured? and (2) how does the degree of asymmetry affect crucial properties of a digraph such as the mixing rate? In the following, we propose one metric—the largest singular value of $\nabla = (\Gamma - \Gamma^T)/2$—as a measure of the degree of asymmetry in a digraph. We show that by explicitly accounting for the degree of asymmetry, we can obtain generally tighter bounds on quantities (for example, mixing rate) associated with random walks (or Markov chains) on digraphs.

4.1. The Degree of Asymmetry and Relations to the Symmetrized Laplacian

In [Chung 05], the symmetrized Laplacian matrix for digraphs, $\bar{\mathcal{L}} = \left(\Gamma + \Gamma^T\right)/2$, is introduced. It generalizes the Cheeger constant to digraphs and bounds it in terms of the second smallest eigenvalue of $\bar{\mathcal{L}}$. In the following we show that the symmetrized Laplacian $\bar{\mathcal{L}}$ and the Cheeger constant introduced by Chung are in fact two quantities intrinsic to undirected graphs.

Theorem 4.1. *Given a digraph G with transition probability matrix P, there exist infinite digraphs that have the same stationary distribution matrix Φ and the same symmetrized transition probability matrix $\bar{P} = (P + \Phi^{-1}P^T\Phi)/2$. As a result, all these graphs have the same symmetrized Laplacian matrix and Cheeger constant.*

Proof. We prove the theorem by construction. Given a digraph $G = (V, E, A)$ with transition probability matrix P, all the digraphs G' with transition probability P' given by

$$P'(\alpha) = \alpha P + (1-\alpha)\Phi^{-1}P^T\Phi \tag{4.1}$$

form an infinite digraph set, denoted by $\mathcal{G}_G$, where $\alpha \in [0, 1]$.

It is easy to check that every $P'(\alpha)$ defined in (4.1) is nonnegative and satisfies $\phi^T P'(\alpha) = \phi^T$ and $P'(\alpha)\mathbf{1} = \mathbf{1}$. Thus $P'(\alpha)$ represents a transition probability matrix of a random walk with stationary distribution ϕ.

For every $G' \in \mathcal{G}_G$, the Diplacian matrix is given by $\Gamma' = \Phi^{1/2}(I - P')\Phi^{-1/2}$, and the symmetrized Laplacian is determined uniquely by $\bar{P}$, since we have

$$\bar{\mathcal{L}} = \frac{\Gamma' + \Gamma'^T}{2} = \Phi^{1/2}\left(I - \frac{P + \Phi^{-1}P^T\Phi}{2}\right)\Phi^{-1/2} = \Phi^{1/2}(I - \bar{P})\Phi^{-1/2}. \tag{4.2}$$

In particular, when $\alpha = \frac{1}{2}$, $P'(\frac{1}{2}) = \bar{P}$ represents the undirected graph $\bar{G}$.

For every $S \subset N = \{1, \ldots, n\}$, define an n-element vector f_S as $f_S(i) = 1/F_\phi(S)$, $i \in S$, and $f_S(i) = -1/F_\phi(\bar{S})$, $i \in \bar{S}$, where $F_\phi(S) := \sum_{i\in S}\phi_i$ and $F_\phi(\bar{S}) := \sum_{i\in\bar{S}}\phi_i$ are the circulation functions [Chung 05]. Define $x_S = \Phi^{-1/2}f_S$.

Then we have

$$x_S^T \Gamma x_S = f_S^T \Phi (I - P) f_S = \sum_{i \in S, j \in \bar{S}} \phi_i p_{ij} = F_\phi(\partial S),$$

$$x_S^T \Gamma x_S = x_S^T \Gamma^T x_S = \frac{1}{2}\big(x_S^T \Gamma x_S + x_S^T \Gamma^T x_S\big) = x_S^T \bar{\mathcal{L}} x_S. \tag{4.3}$$

Hence, we know that for all the graphs $G' \in \mathcal{G}_G$, the circulation functions $F_\phi(\bar{S})$, $F_\phi(S)$, and $F_\phi(\partial S)$ depend only on the partition $S \vee \bar{S} = V$. Therefore, every graph $G' \in \mathcal{G}_G$ has the same Cheeger constant, that is,

$$\min_S \frac{F_\phi(\partial S)}{\min\{F_\phi(S), F_\phi(\bar{S})\}} = h(G') = h(\bar{G}). \tag{4.4}$$

□

To capture the "degree of asymmetry" in a digraph, we express Γ as a sum of a symmetric part and a skew-symmetric part:

$$\Gamma = \bar{\mathcal{L}} + \nabla, \tag{4.5}$$

where we define the skew-symmetric part $\nabla = (\Gamma - \Gamma^T)/2$ as the skewed Laplacian matrix of the digraph. Note that $\Gamma^T = \bar{\mathcal{L}} + \nabla^T = \bar{\mathcal{L}} - \nabla$. Hence ∇ captures the difference between Γ and its transpose, which induces a reserved Markov chain or random walk. When Γ is symmetric, then $\nabla = 0$. Let $0 = \sigma_1 \le \sigma_2 \le \cdots \le \sigma_n$ denote the singular values in increasing order of Γ. Likewise, let $0 = \bar{\lambda}_1 \le \bar{\lambda}_2 \le \cdots \le \bar{\lambda}_n$ denote the eigenvalues of $\bar{\mathcal{L}}$, and $0 = \delta_1 \le \delta_2 \le \cdots \le \delta_n = \delta_{\max}$ the singular values of ∇. The following relations hold among them (see [Bhatia 97]):

$$\bar{\lambda}_i \le \sigma_i \le \bar{\lambda}_i + \delta_n, \quad i = 1, 2, \ldots, n. \tag{4.6}$$

From equation (4.6), we see that $\sigma_i - \bar{\lambda}_i \le \delta_n$, $i = 1, \ldots, n$. When the graph is undirected, we have $\Gamma = \bar{\mathcal{L}}$. Thus $\delta_n = 0$ and $\sigma_i = \lambda_i$, $i = 1, \ldots, n$. We therefore propose the largest singular value of ∇, $\delta_n = \delta_{\max}$, as a measure of the degree of asymmetry in the underlying digraph. Note that $\delta_n = \|\nabla\|$, where $\|\cdot\|$ is the operator bound norm of a matrix:

$$\begin{aligned}\|\nabla\| &= \sup_{\|x\|=1} \|\nabla x\|_2 = \sup_{\|y\|=\|x\|=1} |\langle y, \Gamma x\rangle - \langle y, \Gamma^T x\rangle| \\ &= \sup_{\|y\|=\|x\|=1} |\langle y, \Gamma x\rangle - \langle x, \Gamma y\rangle|\end{aligned}$$

(see [Bhatia 97], pp. 6, 91). On the other hand, $\langle x, \Gamma x\rangle - \langle x, \Gamma^T x\rangle = 0$ for every x.

In the following, we relate and bound δ_n—the degree of asymmetry—to two other important quantities associated with the Markov chain on a digraph: the *digraph gap* $g(G)$ defined below and the second-largest singular value of the transmission probability matrix P.

Given a digraph G, the circulation function $F_\phi(\cdot)$, where $F_\phi(i,j) = \phi_i p_{ij}$, obeys the flow conservation law at every node of the digraph: $\sum_k F_\phi(k,i) = \sum_j F_\phi(i,j)$ for all i. Now define the *digraph gap*

$$g(G) = \max_S \sum_{i\in S} \Big| \sum_{j\in \bar{S}} (F_\phi(i,j) - F_\phi(j,i)) \Big|,$$

which quantifies the maximum difference between two bipartite subgraphs S and $\bar{S}$ among all partitions. We have the following theorem relating the degree of asymmetry with $g(G)$ and $\sigma_{n-1}(\mathcal{P})$, the second-largest singular value of $\mathcal{P} = \Phi^{1/2} P \Phi^{-1/2}$.

Theorem 4.2. (Bounds on the degree of asymmetry.) *We have*

$$2g(G) \le \delta_n \le \lambda_{n-1}^{1/2}(\mathcal{P}^T\mathcal{P}) = \sigma_{n-1}(\mathcal{P}), \tag{4.7}$$

where $\mathcal{P} = \Phi^{1/2} P \Phi^{-1/2}$.

Proof. For a subset of vertices $S \subset N = \{1,\ldots,n\}$, define two n-element vectors y_S and z_S as

$$y_S(i) = \begin{cases} \phi_i^{1/2} & \text{if } i \in \bar{S} \text{ or } g(i,\bar{S}) > 0, \\ -\phi_i^{1/2} & \text{otherwise,} \end{cases}$$

and

$$z_S(i) = \begin{cases} \phi_i^{1/2} & \text{if } i \in \bar{S} \text{ or } g(i,\bar{S}) < 0, \\ -\phi_i^{1/2} & \text{otherwise,} \end{cases}$$

where $g(i,\bar{S}) = \sum_{j\in\bar{S}} (p_{ij}\phi_i - p_{ji}\phi_j)$. Then we have $\|y_S\| = 1$, $\|z_S\| = 1$, and the lower bound of δ_n is obtained as

$$2g(G) = \max_S y_S^T \nabla z_S \le \delta_n. \tag{4.8}$$

Moreover, we have the following equation:

$$\nabla^T\nabla = \frac{1}{2}(\mathcal{P}^T\mathcal{P} + \mathcal{P}\mathcal{P}^T) - \frac{1}{4}(\mathcal{P}^T + \mathcal{P})^2. \tag{4.9}$$

It is easy to check that $\phi^{1/2}$ is the left and right eigenvector of ∇ corresponding to the smallest eigenvalue 0 and the eigenvector of $\mathcal{P}$ corresponding to the largest eigenvalue 1. We also know that $(\mathcal{P}^T + \mathcal{P})^2$ is positive semidefinite. Therefore,

for every $n \times 1$ column vector $x \perp \phi^{1/2}$, we have

$$\begin{aligned}\delta_n^2 &= \max_{x \perp \phi^{1/2}} x^T \nabla^T \nabla x \le \frac{1}{2} \max_{x \perp \phi^{1/2}} x^T (\mathcal{P}^T \mathcal{P} + \mathcal{P}\mathcal{P}^T) x \le \max_{x \perp \phi^{1/2}} x^T \mathcal{P}^T \mathcal{P} x \\ &= \lambda_{n-1}(\mathcal{P}^T \mathcal{P}), \\ \delta_n &\le \lambda_{n-1}^{1/2}(\mathcal{P}^T \mathcal{P}) = \sigma_{n-1}(\mathcal{P}).\end{aligned} \tag{4.10}$$

Now (4.8) and (4.10) yield Theorem 4.2. □

Theorem 4.3 below relates and bounds the second-smallest singular value σ_2 of Γ in terms of the degree of asymmetry δ_n, the Cheeger constant, and the second-smallest eigenvalue $\bar{\lambda}_2$ of $\bar{\mathcal{L}}$.

Theorem 4.3. (Relations among σ_2, $\bar{\lambda}_2$, δ_n, and the Cheeger constant.) *Given a strongly connected graph $G = (V, E, A)$ and its Laplacian matrix $\Gamma = \Phi^{1/2}(I - P)\Phi^{-1/2}$, we have the bounds for the second-smallest singular value of Γ as*

$$\frac{h^2(G)}{2} \le \sigma_2 \le \left(1 + \frac{\delta_n}{\bar{\lambda}_2}\right) \cdot 2h(G). \tag{4.11}$$

When the graph is undirected, we have $h^2(G)/2 \le \sigma_2 = \bar{\lambda}_2 \le 2h(G)$, which is exactly the same as the bounds obtained in [Chung 05].

Proof. (i) Proof of the upper bound: Let σ_2 be the second-smallest (or the smallest nonzero) singular value of the Diplacian matrix Γ, and denote the singular decomposition of the Moore–Penrose pseudoinverse of the Diplacian by $\Gamma^+ = V\Sigma^+ U^T$, where each diagonal entry of Σ^+ is $\sigma_i^+ = 1/\sigma_i$ if $\sigma_i \ne 0$, and $\sigma_i^+ = 0$ otherwise. Then the largest singular value of Γ^+ is $\max_i \sigma_i^+ = 1/\sigma_2$.

Let S denote a subset of vertices $S \subset V$, and let $\bar{S}$ denote the complement of S, that is, $\bar{S} = V - S$. Then we define

$$x_0 = \Phi^{1/2} f_S, \quad y_0 = \Phi^{1/2}(I - P) f_S = \Gamma x_0,$$

where $f : V \to \mathbb{R}$ is defined as follows:

$$f_S(i) = \begin{cases} \dfrac{1}{F_\phi(S)} & \text{if } i \in S, \\ -\dfrac{1}{F_\phi(\bar{S})} & \text{if } i \in \bar{S}, \end{cases} \tag{4.12}$$

with $F_\phi(S) = \sum_{i \in S} \phi_i$ and $F_\phi(\bar{S}) = \sum_{i \in \bar{S}} \phi_i$. Then for each i $(1 \le i \le n)$, we have $\langle v_i, x_0 \rangle = \sigma_i^+ \langle u_i, y_0 \rangle$, which indicates that $\langle v_i, x_0 \rangle$ and $\langle u_i, y_0 \rangle$ have the same sign.

Then we have

$$\sigma_2 = \frac{1}{\max_i \sigma_i^+} \leq \left(\frac{\sum_{i=1}^n \langle v_i, x_0 \rangle \sigma_i^+ \langle u_i, y_0 \rangle}{\sum_{i=1}^n \langle v_i, x_0 \rangle \langle u_i, y_0 \rangle} \right)^{-1} = \frac{x_0^T V U^T y_0}{x_0^T \Gamma^+ y_0}. \tag{4.13}$$

The denominator of (4.13) can be rewritten as

$$\begin{aligned} x_0^T \Gamma^+ y_0 &= f_S^T \Phi^{1/2} \Gamma^+ \Phi^{1/2} (I - P) f_S = f_S^T \Phi Z (I - P) f_S = f_S^T \Phi f_S \\ &= F_\phi(S) F_\phi(\bar{S}) \geq 0, \end{aligned} \tag{4.14}$$

where Z is the fundamental matrix. The third equality holds because we have $Z(I - P) = I - J\Phi$ by Theorem 3.1, and f_S is orthogonal to ϕ, that is, $J\Phi f_S = \phi\phi^T f_S = 0$, where J is a matrix of all 1's. Thus we have $(I - J\Phi) f_S = f_S - J\Phi f_S = f_S$. From (4.13) and (4.14), we have

$$\begin{aligned} \sigma_2 &\leq \frac{x_0^T V U^T y_0}{F_\phi(S) F_\phi(\bar{S})} = \frac{x_0^T V U^T y_0}{x_0^T y_0} \cdot \frac{x_0^T y_0}{F_\phi(S) F_\phi(\bar{S})} = \frac{x_0^T V U^T y_0}{x_0^T y_0} \cdot \frac{f_S^T \Phi (I - P) f_S}{F_\phi(S) F_\phi(\bar{S})} \\ &\leq \frac{x_0^T V U^T y_0}{x_0^T y_0} \cdot 2h(G) = c_0 \cdot 2h(G), \end{aligned} \tag{4.15}$$

where $h(G)$ is the Cheeger constant of the digraph $G = (V, E, A)$.

Now, we are in a position to bound

$$c_0 = \frac{x_0^T V U^T y_0}{x_0^T y_0} \leq 1 + \frac{\delta_n}{\bar{\lambda}_2}.$$

Let $R = VU^T$. Since $x_0^T \Gamma x_0 = x_0^T \Gamma^T x_0$, we can rewrite c_0 as

$$c_0 = \frac{x_0^T R y_0}{x_0^T y_0} = \frac{x_0^T R \Gamma x_0}{x_0^T \Gamma x_0} = \frac{x_0^T V \Sigma V^T x_0}{x_0^T \bar{\mathcal{L}} x_0},$$

where both $R\Gamma$ and $\bar{\mathcal{L}}$ are symmetric positive semidefinite matrices. Denote the eigendecomposition of $\bar{\mathcal{L}}$ by $\bar{\mathcal{L}} = W \Lambda W^T = C^T C$, where $C = \Lambda^{1/2} W^T$, and define $z_0 = C x_0$. Then we have $x_0^T C^T (C^+)^T = x_0^T (I - \mathcal{J}) = x_0^T$, with $x_0 \perp \phi^{1/2}$ and $\mathcal{J} = \phi^{1/2} \phi^{1/2^T}$. As a result, we have

$$c_0 = \frac{z_0^T C^{+T} R \Gamma C^+ z_0}{z_0^T z_0} = \frac{z_0^T T z_0}{z_0^T z_0}. \tag{4.16}$$

Since $T = C^{+T} R \Gamma C^+$ is symmetric, it has real eigenvalues. By substituting $z = Cx$ in $Tz = \lambda z$, we prove that T and $\bar{\mathcal{L}}^+ R\Gamma$ have the same eigenvalues:

$$\begin{aligned} &Tz - \lambda z = \bar{\mathcal{L}}^+ R \Gamma x - \lambda x = \mathbf{0}, \text{ if } x \perp \phi^{\frac{1}{2}}, \\ &Tz - \lambda z = \bar{\mathcal{L}}^+ R \Gamma (I - \mathcal{J}) x - \lambda (I - \mathcal{J}) x = \bar{\mathcal{L}}^+ R \Gamma \mathbf{0} - \lambda \mathbf{0} = \mathbf{0}, \text{ if } x = \alpha \phi^{\frac{1}{2}}, \end{aligned} \tag{4.17}$$

where $\alpha \in \mathbb{R}$. From (4.17) and (4.16), we have

$$\begin{aligned} c_0 &\le \max_z \frac{z^T T z}{z^T z} = \lambda_n(T) = \lambda_n(\bar{\mathcal{L}}^+ R\Gamma) = \lambda_n(R\Gamma\bar{\mathcal{L}}^+) \le \sigma_n(R\Gamma\bar{\mathcal{L}}^+) \\ &= \sigma_n(\Gamma\bar{\mathcal{L}}^+) = \sigma_n((\bar{\mathcal{L}} + \nabla)\bar{\mathcal{L}}^+) \le 1 + \frac{\delta_n}{\bar{\lambda}_2}, \end{aligned} \tag{4.18}$$

where the last two inequalities hold because of [Seber 07, Theorem 6.72(b)(iii), p. 118, and Theorem 6.80, p. 120].

From (4.15) and (4.18), we have

$$\sigma_2 \le \left(1 + \frac{\delta_n}{\bar{\lambda}_2}\right) \cdot 2h(G).$$

In particular, when the graph is undirected, we have $\bar{\mathcal{L}} = \Gamma$ and $1 + \delta_n/\bar{\lambda}_2 = 1$. Thus the above upper bound of σ_2 reduces to $\bar{\lambda}_2 \le 2h(G)$, which subsumes the results obtained in [Chung 05].

(ii) Proof of lower bound: It is proven in [Chung 05, Chung 06] that the second-smallest eigenvalue $\bar{\lambda}_2$ of the symmetrized Laplacian matrix $\bar{\mathcal{L}}$ can be bounded by the Cheeger constant as $\bar{\lambda}_2 \ge h^2(G)/2$. Since for any directed graph $G = (V, E, A)$, we have $\sigma_i \ge \bar{\lambda}_i$ $(1 \le i \le n)$, based on [Bernstein 05, Theorem 5.11.25, p. 355], we have $\sigma_2 \ge \bar{\lambda}_2 \ge h^2(G)/2$. □

Finally, we introduce a generalized Cheeger constant, $\tilde{h}(G)$, defined as

$$\tilde{h}(G) = \min_S \frac{\|\Gamma x_S\|}{\|x_S\|} = \min_{x_S \perp \phi^{1/2}} \frac{(x_S^T \Gamma^T \Gamma x_S)^{1/2}}{(x_S^T x_S)^{1/2}}, \tag{4.19}$$

where for every $S \subset N = \{1, 2, \ldots, n\}$, $x_S = \Phi^{-1/2} f_S$ is as defined above. We see that the generalized Cheeger constant thus defined minimizes the 2-norm of the circulations across bipartite subgraphs S and $\bar{S}$, whereas $h(G)$ minimizes the 1-norm (the sum of absolute values) of the circulations across S and $\bar{S}$. Clearly, $\sigma_2 \le \tilde{h}(G)$.

4.2. Bounding the Mixing Rate of Random Walks on Digraphs

In this section, using mixing rating bounds as an example, we show that by considering the degree of asymmetry, we obtain a better bound for the mixing rate of random walks on digraphs.

The mixing rate is a measure of how fast a random walk converges to its stationary distribution. Many papers have studied the problem of bounding the mixing rate of random walks (or reversible Markov chains) on undirected graphs, such as [Boyd et al. 05, Jerrum and Son 02].

Relatively few papers [Chung 05, Fill 91, Mihail 89] have addressed the problem of bounding the mixing rate of Markov chains (or random walks) on digraphs. In bounding the convergence rate from an initial distribution to the stationary distribution of a Markov chain with the transition probability matrix P, the χ-square distance [Chung 05, Fill 91] is commonly used, which is defined as follows:

$$\chi(t) = \max_{i \in V(G)} \left(\sum_{j \in V(G)} \frac{(P^t(i,j) - \phi_j)^2}{\phi_j} \right)^{1/2}. \tag{4.20}$$

Since not all random walks on strongly connected digraphs are convergent, Chung defines a lazy random walk on G with transition probability matrix $P_L = (I + P)/2$ and derives the following bound on the mixing rate of this lazy random walk using the χ-square distance [Chung 05], where a bound using a closely related total variance distance is also derived. Define $M = \Phi^{1/2} P_L \Phi^{-1/2}$. Then $\chi(t)^2 \leq \varepsilon^t \max_i \phi_i^{-1}$, where

$$\varepsilon = \max_{f \perp \phi^{1/2}} \frac{f^T M^T M f}{f^T f} = \max_{f \perp \phi^{1/2}} \frac{\|Mf\|^2}{\|f\|^2} \leq 1 - \frac{\bar{\lambda}_2}{2}. \tag{4.21}$$

From Theorem 4.1, we know that this bound leads to the same upper bound for all digraphs with the same $\bar{\mathcal{L}}$. By accounting for the degree of asymmetry, we obtain a generally tighter upper bound on $\|Mf\|^2/\|f\|^2$ as follows, which in turn yields a tighter bound on $\chi(t)$.

Theorem 4.4. *For an irreducible Markov chain P and a lazy random walk $P_L = (I + P)/2$, we have*

$$\varepsilon = \max_{f \perp \phi^{1/2}} \frac{\|Mf\|^2}{\|f\|^2} \leq 1 - \frac{\bar{\lambda}_2}{2} - \frac{1 - \sigma_{n-1}^2(\mathcal{P})}{4} \tag{4.22}$$

$$\leq \frac{3}{4} - \frac{\bar{\lambda}_2}{2} + \frac{(\sigma_{n-1}(\bar{\mathcal{P}}) + \delta_n)^2}{4}, \tag{4.23}$$

where

$$M = \Phi^{1/2} P_L \Phi^{-1/2} \quad \textit{and} \quad \bar{\mathcal{P}} = \frac{\mathcal{P} + \mathcal{P}^T}{2} = I - \bar{\mathcal{L}}.$$

Proof. We first note that $M^T M$ can be rewritten as

$$\begin{aligned} M^T M &= \left(I - \frac{\Gamma^T}{2}\right)\left(I - \frac{\Gamma}{2}\right) = \frac{3}{4} I - \frac{\Gamma + \Gamma^T}{4} + \frac{\mathcal{P}^T \mathcal{P}}{4} \\ &= \left(I - \frac{\bar{\mathcal{L}}}{2}\right) - \frac{I - \mathcal{P}^T \mathcal{P}}{4}. \end{aligned}$$

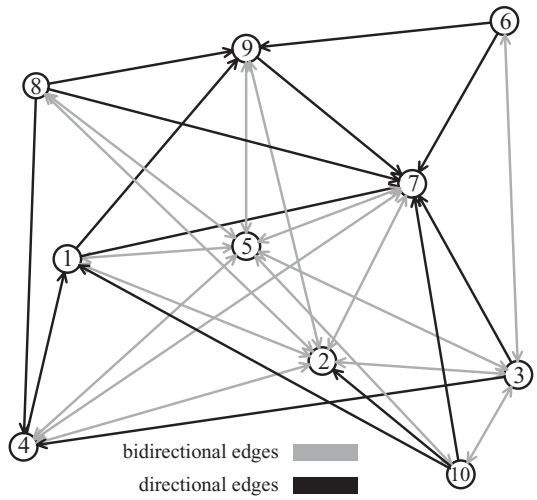

Figure 1. Knoke's data [Knoke and Kuklinski 82, Hanneman and M. Riddle 05] on information exchanges among organizations operating in the social welfare field. The degree of asymmetry of this digraph is 0.3096.

Since $f \perp \phi^{1/2}$ is the left and right singular vectors of Γ corresponding to the largest singular value 1 and is also the eigenvector of $\bar{\mathcal{L}}$ corresponding to the smallest eigenvalue 0, by applying the min–max theorem [Bhatia 97, Theorem 6.58(e), p. 108], we prove (4.22). Then $\sigma_{n-1}^2(\mathcal{P})$ in (4.22) can be further bounded as follows:

$$\sigma_{n-1}^2(\mathcal{P}) \leq \sigma_{n-1}^2(\bar{\mathcal{P}} + \nabla) \leq (\sigma_{n-1}(\bar{\mathcal{P}}) + \delta_n)^2, \tag{4.24}$$

which yields (4.23). □

We remark that since $\sigma_{n-1}(\mathcal{P}) \leq \sigma_n(\mathcal{P}) = 1$, $1 - \sigma_{n-1}^2(\mathcal{P}) \geq 0$ holds, thus the bound in (4.22) is tighter than Chung's bound in (4.21). In (4.23), we further bound $\sigma_{n-1}(\mathcal{P})$ in terms of the degree of asymmetry and the singular value of $\bar{\mathcal{P}}$. In particular, when $\mathcal{P}$ is symmetric, the underlying graph is undirected and the bound in (4.23) boils down to (4.22).

4.3. Numerical Analysis

In this section, we study a real directed network in terms of its degree of asymmetry δ_n and the mixing rate bound. Figure 1 presents the information flow network with ten formal organizations concerned with social welfare issues in one Midwestern American city, where the data were collected by Knoke and Kuklinski [Knoke and Kuklinski 82, Hanneman and M. Riddle 05].

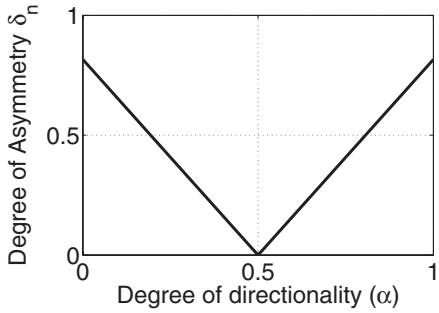

Figure 2. Degree of asymmetry.

It is easy check that the Knoke information exchange digraph is strongly connected; that is, every node can reach every other node in the graph. It has degree of asymmetry $\delta_n = 0.3096$. Using (4.1), we construct a set of new digraphs with α ranging from 0 to 1. As we can see from Figure 2, the new graph $G(\alpha)$ becomes more symmetric with smaller δ_n, when the parameter α gets closer to 1/2. On the other hand, when $|\alpha - \frac{1}{2}|$ is larger, or equivalently α is closer to 0 or 1, the graph becomes more asymmetric, with larger δ_n.

Moreover, as we change α, we compute various mixing rate bounds on ϵ, such as our bounds (4.22) and (4.23), and the bound (4.21) obtained in [Chung 05]. Figure 3 shows the comparisons between mixing rate bounds over the degree of asymmetry in the Knoke information exchange digraph, where (4.21) provides an invariant bound for all digraphs constructed by (4.1), because the construction preserves the stationary distribution and the symmetrized Laplacian matrix $\bar{\mathcal{L}}$. On the other hand, (4.22) and (4.23) bound the random-walk mixing rates on

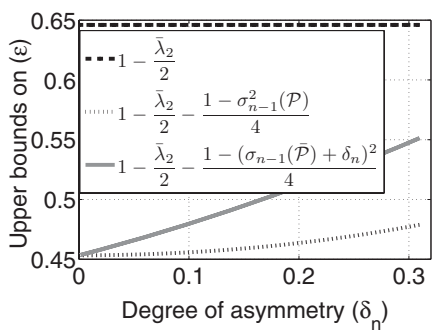

Figure 3. Comparing mixing rate bounds.

those digraphs more precisely, where digraphs that are more symmetric with lower δ_n have lower mixing rate bounds.

Acknowledgments We would like to thank Fan Chung Graham for her useful feedback and comments, which improved the presentation of this work. We also thank the reviewers for very constructive and detailed comments.

This work was supported in part by National Science Foundation grants CNS-0905037 and CNS-1017647, the DTRA grant HDTRA1-09-1-0050, and a University of Minnesota DTI grant. A preliminary version of the results in this paper appeared in [Li and Zhang 10b].

References

[Aldous and Fill 99] D. Aldous and J. A. Fill. "Reversible Markov Chains and Random Walks on Graphs." Available at http://www.stat.berkeley.edu/~aldous/RWG/book.html, 1999.

[Bernstein 05] D. S. Bernstein. *Matrix Mathematics: Theory, Facts, and Formulas with Application to Linear Systems Theory*. Princeton University Press, 2005.

[Bhatia 97] R. Bhatia. *Matrix Analysis*. Springer, 1997.

[Boley et al. 11] D. Boley, G. Ranjan, and Z.-L. Zhang. "Commute Times for a Directed Graph Using an Asymmetric Laplacian." *Linear Algebra and Its Applications* 435:2 (2011), 224–242.

[Boyd et al. 05] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah. "Mixing Times for Random Walks on Geometric Random Graphs." In *SIAM Workshop on Analytic Algorithmics and Combinatorics (ANALCO), Vancouver, 2005*, pp. 240–249, 2005.

[Brin and Page 98] S. Brin and L. Page. "The Anatomy of a Large-Scale Hypertextual Web Search Engine." *Computer Networks and ISDN Systems* 30 (1998), 107–117.

[Chung 05] F. R. K. Chung. "Laplacians and the Cheeger Inequality for Directed Graphs." *Annals of Combinatorics* 9 (2005), 1–19.

[Chung 06] F. R. K. Chung. *Spectral Graph Theory*, CBMS Regional Conference Series in Mathematics 92, 2006.

[Chung and Yau 00] F. R. K. Chung and S. T. Yau. "Discrete Green's Functions." *Journal of Combinatorial Theory, Series A* 91:1 (2000), 191–214.

[Fill 91] J. A. Fill. "Eigenvalue Bounds on Convergence to Stationarity for Nonreversible Markov Chains, with an Application to the Exclusion Process." *Ann. Appl. Probab.* 1 (1991), 62–87.

[Hanneman and M. Riddle 05] R. A. Hanneman and M. Riddle. *Introduction to Social Network Methods*. University of California, Riverside (published in digital form at http://faculty.ucr.edu/~hanneman/), 2005.

[Horn and Johnson 85] R. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, 1985.

[Jerrum and Son 02] M. Jerrum and J.-B. Son. "Spectral Gap and Log-Sobolev Constant for Balanced Matroids." In *43rd Annual IEEE Symposium on Foundations of Computer Science (FOCS 2002)*, pp. 721–729, 2002.

[Knoke and Kuklinski 82] D. Knoke and J. H. Kuklinski. *Network Analysis*. Sage Publications, 1982.

[Li and Zhang 10a] Y. Li and Z.-L. Zhang. "Random Walks on Digraphs: A Theoretical Framework for Estimating Transmission Costs in Wireless Routing." In *INFOCOM 2010: Proceedings of the 29th IEEE International Conference on Computer Communications*, pp. 1–9, 2010.

[Li and Zhang 10b] Y. Li and Z.-L. Zhang. "Random Walks on Digraphs: The Generalized Digraph Laplacian and the Degree of Asymmetry." In *LNCS WAW 2010: Proceedings of the 7th Workshop on Algorithms and Models for the Web Graph, Stanford, CA, 2010*, pp. 74–85, 2010.

[Li et al. 11] Y. Li, Z.-L. Zhang, and D. Boley. "The Routing Continuum from Shortest-path to All-path: A Unifying Theory." In *31st International Conference on Distributed Computing Systems*, pp. 847–856, 2011.

[Li and Zhang 12] Y. Li and Z.-L. Zhang. "Random Walks and Green's Function on Digraphs: A Framework for Estimating Wireless Transmission Costs." *IEEE/ACM Transactions on Networking*, PP(99) (2012), 1–14.

[Li et al. 12] Y. Li, Z.-L. Zhang, and J. Bao. "Mutual or Unrequited Love: Identifying Stable Clusters in Social Networks with Uni- And Bi-directional Links." In *LNCS WAW 2012: Proceedings of the 9th Workshop on Algorithms and Models for the Web Graph, Halifax, Nova Scotia, Canada, 2012*, pp. 113–125, 2012.

[Lovász 93] L. Lovász. "Random Walks on Graphs: A Survey." *Combinatorics* 2 (1993), 1–46.

[Mihail 89] M. Mihail. "Conductance and Convergence of Markov Chains: A Combinatorial Treatment of Expanders." In *FOCS 1989: Proceedings of the IEEE 30th Annual Symposium on Foundations of Computer Science*, pp. 526–531, 1989.

[Seber 07] G. A. F. Seber. *A Matrix Handbook for Statisticians*. Wiley-Interscience, 2007.

[Zhou et al. 05] D. Zhou, J. Huang, and B. Schölkopf. "Learning from Labeled and Unlabeled Data on a Directed Graph." In *ICML '05: Proceedings of the 22nd International Conference on Machine Learning*, pp. 1036–1043, 2005.

Yanhua Li, Computer Science and Engineering Department, University of Minnesota, 4-192 Keller Hall, 200 Union Street SE, Minneapolis, MN 55455 (yanhua@cs.umn.edu)

Zhi-Li Zhang, Computer Science and Engineering Department, University of Minnesota, 4-192 Keller Hall, 200 Union Street SE, Minneapolis, MN 55455 (zhzhang@cs.umn.edu)