The Interplay Between Dynamics and Networks: Centrality, Communities, and Cheeger Inequality

Rumi Ghosh
Robert Bosch
rumi.ghosh@gmail.com

Shang-Hua Teng
Computer Science
USC
shanghua@usc.edu

Kristina Lerman and
Xiaoran Yan
Information Sciences Institute
USC
lerman/xiaoran@isi.edu

ABSTRACT

We study the interplay between a *dynamic process* and the *structure* of the network on which it is defined. Specifically, we examine the impact of this interaction on the quality-measure of network clusters and node centrality. This enables us to effectively identify network communities and important nodes participating in the dynamics. As the first step towards this objective, we introduce an umbrella framework for defining and characterizing an ensemble of dynamic processes on a network. This framework generalizes the traditional *Laplacian framework* to continuous-time biased random walks and also allows us to model some epidemic processes over a network. For each dynamic process in our framework, we can define a function that measures the quality of every subset of nodes as a potential cluster (or community) with respect to this process on a given network. This subset-quality function generalizes the traditional *conductance measure* for graph partitioning. We partially justify our choice of the quality function by showing that the classic Cheeger's inequality, which relates the conductance of the best cluster in a network with a spectral quantity of its Laplacian matrix, can be extended from the Laplacian-conductance setting to this more general setting.

1. INTRODUCTION

Two fundamental problems in network analysis involve identifying central nodes and communities in a network. The goal of centrality identification is to find central or important nodes, for example, those that control the flow of information on the network. The objective of community detection is to discover subsets of well-interacting nodes in a given network. Both of these graph mining approaches have tremendous applications in different areas, such as social network, biological network, World Wide Web analysis. For example, the centrality measure Page Rank [20] is one of the backbones of Google's search algorithm. Similarly, community detection or graph clustering is useful in potential drug-target identification [29].

The first step towards finding potentially meaningful communities is to define a *quality function* that measures the degree to which a subset of network nodes tend to interact among themselves. To this end, measures such as clustering coefficients, subgraph density, and conductance, were introduced to express the quality of a subset as a good cluster. Similarly, many measures of *centrality* including Page Rank and eigenvector centrality are used to discover important nodes. However, to define a useful quality or centrality function, one needs to determine how the network structure models the interaction among nodes. Unfortunately, this can be quite challenging and is often ignored. Indeed, the aforementioned measures apply different interpretations. As a result, current network analysis methods often make implicit assumptions about the dynamic processes (and the resultant interaction) among nodes and use these assumptions to define the quality of a cluster or centrality of a node.

Explicitly taking the dynamic process into account, a node's centrality at any time describes its participation in the flow taking place on the network [5, 15]. Similarly, communities are groups of nodes that interact more frequently with each other (according to the rules of the dynamic process) than with nodes from other communities [16]. In fact, this view of modeling is not new. For instance, in choosing *conductance* as a measure of cluster quality, one may in fact implicitly assume that *unbiased random walk* is taking place on the network [12, 26, 8, 9]. Under this assumption, a measure of centrality is the *heat kernel page rank* [6].

Other dynamic processes such as the spread of information, ideas, or epidemics, induce different interactions from the unbiased random walks. An epidemic [13] is a stochastic process that, unlike a random walk, attempts to transition to (i.e., infect) every neighbor of a node. Epidemic dynamics may be specified by the *replicator* operator [16], whose stationary distribution defines *eigenvector centrality* [4, 10]. It is natural to think that the centrality of a node depends on the specifics of the dynamic process, which together with the network topology influence its activity level. For example, nodes that are visited most frequently by a random walk (specified by the heat kernel page rank) are different from nodes that are infected most often during an epidemic (specified by eigenvector centrality). Moreover, epidemics could lead to a different community formation of a network [16].

Our Contributions.

We study the interplay between a dynamic process and the underlying network on which it unfolds. We focus on the impact of this interaction on the emergence of central

nodes and the formation of communities in the network, and on the design of efficient algorithms for their identification.

General framework for dynamic processes: We present an umbrella framework for describing dynamic processes on a network that generalizes the traditional Laplacian framework for diffusion and random walks. Recall that a random walk on a network is a stochastic dynamic process that transitions from a node to a random neighbor of that node. It defines a Markov chain that can be specified by the normalized Laplacian of the network. Our framework (Section 2) attempts to capture the class of dynamic processes that evolve in time according to a rule that can be defined by generalizing the normalized Laplacian.

Formal analysis of interaction dynamics: Our framework defines a class of dynamic processes with relatively simple characterization parameters, which enables rigorous analysis of the impact of these parameters on the measures of the community quality and node centrality. Its inclusion of diffusion and random walks allows us to build on the insights from previous work on mathematical analysis of random walks, conductance measures for clustering, and spectral notions for centrality, in order to derive our generalization that reflects the interaction between the dynamic process and its underlying network. We are also able to use it to define new processes whose properties may offer useful insights into community finding and graph partitioning. Although this particular framework cannot express every dynamic process of interest, the generalization allows us to model some epidemic processes over a network. Thus, we can use these well- established special cases (such as random walks and epidemics) to clarify the relationship between the parameters of the dynamic processes and their induced community qualities and centralities.

Generalized conductance: We extend conductance to a general quality-measure for clusters that reflects the dynamic process on the network. For each dynamic process, we use its characterization parameters to define this function. It measures the quality of every subset (of nodes) as a potential cluster (community) with respect to this process on the given network. This subset-quality measure generalizes traditional conductance and provides a continuous family of measures for subsets in a graph. Recall that, for each set $S \subset V$ in a weighted graph $G = (V, E)$, the *conductance* of S, $\phi_G(S)$, is equal to the ratio of the degree of the connectivities between S to $V - S$ to total degree of connectivities of S (or the total degree of connectivities of $V - S$, whichever is smaller). As each subset-quality function measures different degree to which a subset of nodes in a graph tend to cluster together and its coherence as a community, a comparative study of these subset-measures could be useful in the design and evaluation of community detection algorithms.

Generalized Cheeger's inequality: In both theory and practice, a basic problem in community detection is to justify why one quality function is better than another for modeling communities in a network. Unfortunately, this is a highly challenging problem for rigorous reasoning and justification. The interplay of dynamics and network topology further adds to its difficulty. In this paper, we will build on the wisdoms demonstrated in the past decades in spectral graph theory. To partially justify our choice of the quality function, we prove that the classical Cheeger's inequality can be extended to our generalized settings. Recall (also see Section 3) that the Cheeger's inequality relates a spec-

tral quantity of the Laplacian matrix of the network to the conductance of the best cluster in the network. Particular, it shows that if λ is the second smallest eigenvalue of the normalized Laplacian matrix of G and $S \subset V$ is a subset in G with the smallest conductance, then

$$\left(\phi_G(S)\right)^2 / 2 \leq \lambda \leq 2 \cdot \phi_G(S). \tag{1}$$

Centrality, quality and community: We will show that the same relation (as in Eq. 1) holds for the generalized conductance and the extreme eigenvalue of the linear operator for each dynamic process in our framework. Each function that measures community quality can then be used to decompose the network into potential communities.

In addition, like previous work in random walks, we can also relate the convergence to stationary distribution with the quality measure in each of the dynamics in this framework. For example, the time taken by the random walk to reach its stationary distribution is bounded by conductance [11]. The Cheeger inequality provides this connection via the extreme eigenvalue of the normalized Laplacian. Similarly, using generalized Cheeger inequalities, we also show that the time it takes for each of the dynamic processes to reach its stationary is bounded by the corresponding quality function.

Efficient spectral and local partitioning algorithms: Like previous studies in Laplacian-based spectral graph theory, such as Spielman-Teng [26], Andersen-Chung-Lang [2], and Andersen-Peres [3], our analysis also leads to efficient spectral and local algorithms for identifying provably good clusters based on these new quality functions.

We hope this study will help lead to better approaches for defining and understanding the general interaction between dynamics and networks. Due to lack of space we are unable to provide the proofs of many theorems in this paper, which will be available in the extended version of this paper.

2. DYNAMIC PROCESSES ON NETWORKS

We represent a network as a weighted, undirected graph $G = (V, E, \boldsymbol{A})$, where for $i, j \in V$, $\boldsymbol{A}[i, j] = a_{i,j}$ assigns an (affinity) non-negative weight to each edge. We follow the tradition that $a_{i,j} = 0$ if and only if $(i, j) \notin E$; i.e., $\boldsymbol{A}$ is the weighted adjacency matrix. By convention we assume it is symmetric and $a_{i,i} = 0$ for all $i \in V$. In the discussion below, the *(weighted) degree* of node $i \in V$ is defined as the total weight of edges incident on it, that is, $d_i = \sum_j a_{i,j}$.

2.1 A Family of Dynamic Processes

In a dynamic process, each network node i has a dynamic variable θ_i associated with it, which can change its value based on interactions with its neighbors. The values of the variables evolve in time according to the rules of the dynamic process. We consider linear dynamic processes of the kind:

$$\frac{d\boldsymbol{\theta}}{dt} = -\boldsymbol{\mathcal{L}}\boldsymbol{\theta}, \tag{2}$$

where $\boldsymbol{\theta}$ is a column vector containing the θ entries and $\boldsymbol{\mathcal{L}}$ is a symmetric positive semi-definite matrix, the *spreading operator*, which defines the details of the dynamic process.

As discussed in the introduction, we focus on dynamic processes on a network that generalize the traditional normalized Laplacian for diffusion and random walks. Recall that the *symmetric normalized Laplacian matrix* of a weighted

graph $G = (V, E, \boldsymbol{A})$ is defined as

$$\boldsymbol{D}^{-1/2}(\boldsymbol{D} - \boldsymbol{A})\boldsymbol{D}^{-1/2},$$

where $\boldsymbol{D}$ is the diagonal matrix defined by $(d_1, ..., d_n)$. We study the properties of a dynamic process whose spreading operator can be written as:

$$\boldsymbol{\mathcal{L}} = \boldsymbol{T}^{-1/2}\boldsymbol{D_W}^{-1/2}(\boldsymbol{D_W} - \boldsymbol{W})\boldsymbol{D_W}^{-1/2}\boldsymbol{T}^{-1/2}. \qquad (3)$$

Here $\boldsymbol{T}$ is the $n \times n$ diagonal matrix of *node delay factors*. Its ith diagonal element τ_i represents the average delay of node i. Another generalization from the traditional Laplacian is the use of the *interaction matrix* $\boldsymbol{W}$ instead of the adjacency matrix $\boldsymbol{A}$. In theory, $\boldsymbol{W}$ can be any $n \times n$ symmetric positive-definite matrix; however, here we restrict our attention to scaling transformations of the adjacency matrix $\boldsymbol{A}$. We also assume that the spreading operator is *properly scaled*: specifically, we assume that $\tau_i \geq 1$, for all $i \in V$. Note that the degree matrix $\boldsymbol{D_W}$ is now also defined in terms of the interaction matrix, that is $d_{\boldsymbol{W}\,i} = \sum_j w_{i,j}$. We name this spreading operator the *generalized Laplacian*.

We study how the interplay between dynamics and topology affects clustering as $\boldsymbol{T}$ and $\boldsymbol{W}$ vary. For a better intuition, it helps to consider the random walk Laplacian matrix:

$$\boldsymbol{\mathcal{L}}^{RW} = (\boldsymbol{D_W} - \boldsymbol{W})\boldsymbol{D_W}^{-1}\boldsymbol{T}^{-1}.$$

This matrix is similar to our symmetric normalized Laplacian matrix, with identical spectral properties. As its name indicates, this operator defines a continuous-time unbiased random walk on the interaction graph $\boldsymbol{W}$, which according to [15] is equivalent to biased random walks on the original adjacency matrix $\boldsymbol{A}$. Under this interpretation, τ_i is the mean delay time of the random walk on node i before a transition, assuming a simple Poisson process. This intuition leads to two orthogonal parametrizations of a dynamic process: namely $\boldsymbol{W}$ controls the distribution of walk trajectories and $\boldsymbol{T}$ controls the delay time of node transitions along each trajectory. While we use symmetric operators for mathematical convenience in definitions and proofs, it is more intuitive to think from the random walk perspective.

While the generalized Laplacian does not cover all dynamic processes of interest, this family of spreading operators includes some well-known ones, such as the Laplacian or normalized Laplacian, as well as a continuous family of new operators that are not as well studied. It also contains certain operators for modeling epidemics. The consideration of this family of operators is also partially motivated by our recent experimental work in understanding network centrality [10, 16]. We conclude this subsection with some special cases that can be described in this framework.

Normalized Laplacian.
If the interaction matrix is the original adjacency matrix $\boldsymbol{W} = \boldsymbol{A}$, and node delay factor is simply the identity matrix $\boldsymbol{T} = \boldsymbol{I}$, then we recover the *normalized symmetric Laplacian*:

$$\boldsymbol{\mathcal{L}} = \boldsymbol{I} - \boldsymbol{D}^{-1/2}\boldsymbol{A}\boldsymbol{D}^{-1/2}.$$

(Scaled) Graph Laplacian.
When $\boldsymbol{W} = \boldsymbol{A}$, $\boldsymbol{T} = d_{max}\boldsymbol{D}^{-1}$, the spreading operator corresponds to the (scaled) graph *Laplacian*

$$\boldsymbol{\mathcal{L}} = \frac{1}{d_{max}}(\boldsymbol{D} - \boldsymbol{A}).$$

This operator is often used to describe *heat diffusion*-like processes.

Replicator.
Let $\boldsymbol{v}$ be the eigenvector of $\boldsymbol{A}$ associated with its largest eigenvalue λ_{max}: $\boldsymbol{A}\boldsymbol{v} = \lambda_{max}\boldsymbol{v}$. We can then construct a diagonal matrix $\boldsymbol{V}$ whose elements are the components of the eigenvector $\boldsymbol{V}$. Let us scale the adjacency matrix according to $\boldsymbol{W} = \boldsymbol{V}\boldsymbol{A}\boldsymbol{V}$ and use it as the interaction matrix. Setting the node delay factor to identity, the spreading operator is:

$$\boldsymbol{\mathcal{L}} = \boldsymbol{I} - \boldsymbol{D_W}^{-1/2}\boldsymbol{W}\boldsymbol{D_W}^{-1/2} = \boldsymbol{I} - \frac{1}{\lambda_{max}}\boldsymbol{A}$$

This operator is known as the replicator matrix $\boldsymbol{R}$, and it models *epidemic diffusion* on a graph [16]. It is simply the normalized Laplacian of the interaction graph $\boldsymbol{V}\boldsymbol{A}\boldsymbol{V}$ [24]. Using the random walk intuition, an unbiased random walk on this interaction graph is equivalent to a maximum entropy random walk on the original graph given by the adjacency matrix $\boldsymbol{A}$ [15].

Unbiased Adjacency Matrix.
Reweighing each edge by the inverse of square root of the degrees of the endpoints gives the unbiased adjacency matrix $\boldsymbol{W} = \boldsymbol{D}^{-1/2}\boldsymbol{A}\boldsymbol{D}^{-1/2}$. Then, the degree of node i in the scaled graph is $d_{\boldsymbol{W}\,i} = \sum_{j \in V} \boldsymbol{W}[i,j]$. Letting $\boldsymbol{T} = d_{\boldsymbol{W}\,max}\boldsymbol{D_W}^{-1}$ gives a *unbiased Laplacian matrix*:

$$\boldsymbol{\mathcal{L}} = \frac{1}{d_{\boldsymbol{W}\,max}}(\boldsymbol{D_W} - \boldsymbol{W}).$$

Similarly, many other operators can be expressed using this framework.

2.2 Network Dynamics and Centrality

Before considering the impact of spreading operators on communities, we first examine the node centralities of each operator at any time, and their stationary distribution after convergence. The community quality functions that we use are partially inspired by the stationary distribution.

Solution of Equation 2 gives the weight distribution of the dynamic process across nodes at any time based on the initial weight distribution. In the rest of the paper, for convenience we will refer to this instantaneous distribution $\boldsymbol{\theta}(t)$ as *time-dependent centrality*. Its stationary distribution is the conventional *stationary centrality*, or centrality for short.

$$\boldsymbol{\theta}(t) = e^{-\boldsymbol{\mathcal{L}}t} \cdot \boldsymbol{\theta}_0 = \sum_{k=0}^{\infty} \frac{(-t)^k}{k!}\boldsymbol{\mathcal{L}}^k\boldsymbol{\theta}_0 \qquad (4)$$

If the dynamic process converges when starting from $\boldsymbol{\theta}_0$, then $\lim_{t \to \infty} \boldsymbol{\theta}(t)$ is proportional to $\boldsymbol{\pi}$:

$$\pi_i = \frac{\sqrt{d_{\boldsymbol{W}\,i}\tau_i}}{\sum_j \sqrt{d_{\boldsymbol{W}\,j}\tau_j}} \qquad (5)$$

For example, the stationary distribution of $\boldsymbol{R}$ is $\boldsymbol{v}$, also known as the *eigenvector centrality*. Eigenvector centrality was introduced by Bonacich [4] to explain the importance of actors in a social network based on the importance of the actors to which they were connected, and it gives the stationary distribution of a simple epidemic at the epidemic threshold [27, 10].

2.3 Network Dynamics and Communities

In network clustering and community detection, one would like to identify subsets of nodes $S \subseteq V$ that are more similar, or behave more similarly, to each other than to nodes in other subsets. A standard approach to clustering involves defining an objective function that measures the *quality* of a cluster. For a subset $S \subseteq V$, let $\bar{S} = V \setminus S$ to denote the complement of S, which consists of nodes that are not in S. Let $\text{cut}(S, \bar{S}) = \sum_{i \in S, j \in \bar{S}} a_{i,j}$ denote the total affinity weights of all edges used by S to connect with the outside world. Let $\text{vol}(S) = \sum_{i \in} d_i = \sum_{i \in S, j \in V} a_{i,j}$ denote the volume of all affinity weights involving vertices in S.

One popular heuristic to measure the quality of a subset S as a potential good cluster (or a community) [12, 26, 8] is to use the ratio of these two quantities:

$$\phi(S) = \frac{\text{cut}(S, \bar{S})}{\min(\text{vol}(S), \text{vol}(\bar{S}))} \qquad (6)$$

For example, a subset that (approximately) minimizes this quantity — the *conductance* of S — is a desirable cluster, as it maximizes the fraction of affinities within the subset. If interactions among nodes are proportional to their affinity weights, then a set with small conductance also means that its members interact significantly more with each other than with members not in the subset. Other well-known quality functions are normalized cut [23] and ratio-cut, given by

$$\frac{\text{cut}(S, \bar{S})}{\text{vol}(S)} + \frac{\text{cut}(S, \bar{S})}{\text{vol}(\bar{S})} \quad \text{and} \quad \frac{\text{cut}(S, \bar{S})}{\min(|S|, |\bar{S}|)},$$

respectively. The smallest achievable such ratio is known as the *isoperimetric number*.

Algorithmically, once a quality function is selected, one can then perform a partitioning-based algorithm or mathematical programming-based method to find a cluster or clusters that optimizes the quality function.

For the Normalized Laplacian paradigm (or when the underlying process is a random walk), there exists a relationship between clustering and dynamics: in a good (i.e., low conductance cut) cluster, a random walk starting within a cluster seldom transitions outside the cluster [17, 23, 22, 26]. Therefore, the presence of a good cluster implies that it will take a random walk a long time to reach its stationary distribution. We generalize this notion with a claim that every dynamic process has an associated function that measures the quality of the cluster with respect to that process. Optimizing the quality function leads to cohesive communities, i.e., groups of nodes that the dynamic process seldom leaves.

Generalized Conductance.
Consider a dynamic process defined by a spreading operator $\mathcal{L} = \boldsymbol{T}^{-1/2} \boldsymbol{D}_{\boldsymbol{W}}^{-1/2} (\boldsymbol{D}_{\boldsymbol{W}} - \boldsymbol{W}) \boldsymbol{D}_{\boldsymbol{W}}^{-1/2} \boldsymbol{T}^{-1/2}$. For a set $S \subseteq V$, let $\text{vol}_{\mathcal{L}}(S) = \sum_{i \in S} d_{\boldsymbol{W} i} \tau_i$ be the generalized volume under our framework. We define the quality of a set S with respect to $\mathcal{L}$ as:

$$h_{\mathcal{L}}(S) = \frac{\text{cut}(S, \bar{S})}{\min\left(\text{vol}_{\mathcal{L}}(S), \text{vol}_{\mathcal{L}}(\bar{S})\right)}$$

$$= \frac{\sum_{i \in S, j \in \bar{S}} w_{i,j}}{\min\left(\sum_{i \in S} d_{\boldsymbol{W} i} \tau_i, \sum_{i \in \bar{S}} d_{\boldsymbol{W} i} \tau_i\right)} \qquad (7)$$

We also define,

$$\phi_{\mathcal{L}} = \min_S h_{\mathcal{L}}(S) \qquad (8)$$

Using the random walk perspective, the numerator measures the random jumps across communities, while the denominator ensures a balanced bisection. The generalized volume can be interpreted as the total time a random walk stays within a community after convergence, as it is proportional to both node degrees and node delay factors.

Below, we summarize the induced special cases discussed in the previous subsection.

Normalized Laplacian.
$\boldsymbol{W} = \boldsymbol{A}$ and $\boldsymbol{T} = \boldsymbol{I}$, and hence $h_{\mathcal{L}}(S)$ is the conductance.

(Scaled) Graph Laplacian.
$\boldsymbol{W} = \boldsymbol{A}$ and $\boldsymbol{T} = d_{max} \boldsymbol{D}^{-1}$, hence

$$h_{\mathcal{L}}(S) = \frac{\text{cut}(S, \bar{S})}{\min(d_{\max}|S|, d_{max}|\bar{S}|)} = \frac{1}{d_{\max}} \cdot \frac{\text{cut}(S, \bar{S})}{\min(|S|, |\bar{S}|)},$$

This is the ratio cut scaled by $1/d_{\max}$.

Replicator.
$\boldsymbol{W} = \boldsymbol{V} \boldsymbol{A} \boldsymbol{V}$ and $\boldsymbol{T} = \boldsymbol{I}$. Recall $\boldsymbol{v}$ is the eigenvector of $\boldsymbol{W}$ associated with the largest eigenvalue $\lambda_{\max}$. The redefined cut size is $\sum_{i \in S, j \in \bar{S}} f_{ij} = \sum_{i \in S, j \in \bar{S}} v_i v_j$. Therefore,

$$h_{\mathcal{L}}(S) = \frac{\sum_{i \in S, j \in \bar{S}} v_i v_j}{\lambda_{\max} \min\left(\sum_{i \in S} v_i^2, \sum_{i \in \bar{S}} v_i^2\right)}$$

Since the degree of a node in an interaction graph is $d_{\boldsymbol{W} i} = \sum_j f_{ij} = \lambda_{\max} v_i^2$, the generalized conductance of the Replicator is simply the conductance of the interaction graph [24].

Unbiased Adjacency Matrix.
$\boldsymbol{W} = \boldsymbol{D}^{-1/2} \boldsymbol{A} \boldsymbol{D}^{-1/2}$. and $\boldsymbol{T} = d_{\boldsymbol{W} max} D_{\boldsymbol{W}}^{-1}$. The associated quality function is

$$h_{\mathcal{L}}(S) = \frac{1}{d_{\boldsymbol{W} max}} \cdot \frac{\sum_{i \in S, j \in \bar{S}} \frac{a_{i,j}}{\sqrt{d_i d_j}}}{\min(|S|, |\bar{S}|)}.$$

We call this quality function *unbiased cohesion*.

3. GENERALIZED CHEEGER'S INEQUALITY AND SPECTRAL PARTITIONING

Cheeger inequality states that

$$\phi_G^2/2 \le \lambda_1 \le 2\phi_G$$

where λ_1 is the second largest eigenvalue of the normalized Laplacian, $\mathcal{L} = \boldsymbol{I} - \boldsymbol{D}^{-1/2} \boldsymbol{W} \boldsymbol{D}^{-1/2}$, and ϕ_G is conductance. The relationship between conductance and spectral properties of the Laplacian enables the use of its eigenvectors for partitioning graphs, particularly the nearest-neighborhood graphs and finite-element meshes [25].

In this section, we generalize Cheeger's inequality to any spreading operator under our framework and its associated generalized conductance (given by Eq. 8). Our generalization of Cheeger's inequality comes with algorithmic consequences. It leads to spectral partitioning algorithms that are efficient in finding low conductance cuts for a given operator.

3.1 Generalized Cheeger Inequality

THEOREM 1. *(Generalized Cheeger Inequality)*
Consider the dynamic process described by a (properly scaled)

spreading operator $\mathcal{L} = T^{-1/2}D_W^{-1/2}(D_W - W)D_W^{-1/2}T^{-1/2}$. Let $\lambda_0 \leq \lambda_1 \leq ... \leq \lambda_{n-1}$ be the eigenvalues of $\mathcal{L}$. Then $\lambda_0 = 0$ and λ_1 satisfies the following inequalities:

$$\phi_{\mathcal{L}}(G)^2/2 \leq \lambda_1 \leq 2\phi_{\mathcal{L}}(G)$$

where $\phi_{\mathcal{L}}(G)$ is given by Eq. 8.

PROOF. We prove the theorem by following the approach for proving the classic Cheeger's inequality (see [8]).

Let $(\tau_1, ..., \tau_n)$ be the diagonal entries of T. Note that $v_0 = T^{1/2}D_W^{1/2} \cdot \mathbf{1}$, where $\mathbf{1}$ denotes the vector of all 1's, is an eigenvector of $\mathcal{L}$ associated with eigenvalue $\lambda_0 = 0$. Let $\mathrm{vol}_{\mathcal{L}}(S) = \sum_{i \in S} d_i \tau_i$ for $S \subseteq V$, where for clarity we abuse the notation d_i and use it as $d_{W\,i}$. Suppose f is the eigenvector associated with λ_1. Then, $f \perp v_0$. Consider vector g such that $g[u] = f[u]/\sqrt{d_u \tau_u}$. The fact that $f \perp v_0$ then implies $\sum_v g[v] d_v \tau_v = 0$. Then,

$$\phi_{\mathcal{L}}(G) = \min_S \frac{\mathrm{cut}(S, \bar{S})}{\min(\mathrm{vol}_{\mathcal{L}}(S), \mathrm{vol}_{\mathcal{L}}(\bar{S}))}$$

$$\lambda_1 = \frac{f^T \mathcal{L} f}{f^T f} = \frac{\sum_{u,v \in V}\left(\frac{f[u]}{\sqrt{d_u \tau_u}} - \frac{f[v]}{\sqrt{d_v \tau_v}}\right)^2 w_{u,v}}{\sum_v f[v]^2}$$

$$= \frac{\sum_{u,v \in V}(g[u] - g[v])^2 w_{u,v}}{\sum_v g[v]^2 d_v \tau_v}$$

Instead of sweeping the vertices of G according to the eigenvector f itself, we sweep the vertices of the graph G according to g by ordering the vertices of G so that

$$g[v_1] \geq g[v_2] \geq \cdots \geq g[v_n]$$

and consider sets $S_i = \{v_1, \cdots, v_i\}$ for all $1 \leq i \leq n$.

Similar to [8], we will eventually only consider the first "half" of the sets S_i during the sweeping: Let r denote the largest integer such that $\mathrm{vol}_{\mathcal{L}}(S_r) \leq \mathrm{vol}_{\mathcal{L}}(V)/2$. Note that

$$\sum_v (g[v] - g[v_r])^2 d_v \tau_v$$

$$= \sum_v g[v]^2 d_v \tau_v + g[v_r]^2 d_v \tau_v \geq \sum_v g[v]^2 d_v \tau_v.$$

where the first equation follows from $\sum_v g[v]d_v\tau_v = 0$. We denote the positive and negative part of $g - g[v_r]$ as g_+ and g_- respectively:

$$g_+[v] = \begin{cases} g[v] - g[v_r], & \text{if } g[v] \geq g[v_r]. \\ 0, & \text{otherwise.} \end{cases} \quad (9)$$

$$g_-[v] = \begin{cases} |g[v] - g[v_r]|, & \text{if } g[v] \leq g[v_r]. \\ 0, & \text{otherwise.} \end{cases} \quad (10)$$

Now

$$\lambda_1 = \frac{\sum_{u,v \in V}(g[u] - g[v])^2 w_{u,v}}{\sum_v g[v]^2 d_v \tau_v}$$

$$\geq \frac{\sum_{u,v \in V}(g_+[u] - g_+[v])^2 w_{u,v} + (g_-[u] - g_-[v])^2 w_{u,v}}{\sum_v (g_+[v]^2 + g_-[v]^2) d_v \tau_v}$$

$$\geq \min\left[\frac{\sum(g_+[u] - g_+[v])^2 w_{u,v}}{\sum_v g_+[v]^2 d_v \tau_v}, \frac{\sum(g_-[u] - g_-[v])^2 w_{u,v}}{\sum_v g_-[v]^2 d_v \tau_v}\right]$$

Without loss of generality, we assume the first ratio is at most the second ratio, and will mostly focus on the vertices

$\{v_1, ..., v_r\}$ in the first "half" of the graph in the analysis below. Thus,

$$\lambda_1 \geq \frac{\sum_{u,v}(g_+[u] - g_+[v])^2 w_{u,v}}{\sum_v g_+[v]^2 d_v \tau_v}$$

$$\geq \frac{\left(\sum_{u,v}(g_+^2[u] - g_+^2[v])w_{u,v}\right)^2}{\left(\sum_v g_+[v]^2 d_v \tau_v\right)\left(\sum_{u,v}(g_+[u] + g_+[v])^2 w_{u,v}\right)}$$

which follows from the Cauchy-Schwartz inequality.

We now separately analyze the numerator and denominator. To bound the denominator, we will use the following property of τ_i: Because $\mathcal{L}$ is properly scaled, $\tau_i \geq 1$ for all $i \in V$. Therefore,

$$\sum_{u,v}(g_+[u] + g_+[v])^2 w_{u,v} \leq \sum_{u,v} 2(g_+^2[u] + g_+^2[v])w_{u,v}$$

$$= 2 \sum_{u \in V} g_+^2[u] d_u$$

$$\leq 2 \sum_{u \in V} g_+^2[u] d_u \tau_u.$$

Hence, the denominator is at most

$$2\left(\sum_{u \in V} g_+^2[u] d_u \tau_u\right)^2.$$

To bound the numerator, we consider subsets of nodes $S_i = \{v_1, \cdots, v_i\}$ for all $1 \leq i \leq r$ and define $S_0 = \emptyset$. First note that

$$\mathrm{vol}_{\mathcal{L}}(S_i) - \mathrm{vol}_{\mathcal{L}}(S_{i-1}) = d_{v_i} \tau_{v_i}. \quad (11)$$

By the definition of $\phi_{\mathcal{L}}(G)$, we know $\phi_{\mathcal{L}}(G) \leq \min_i h_{\mathcal{L}}(S_i)$ for all $1 \leq i \leq r$, where recall the function $h_S(\mathcal{L})$ is defined by Eq. 7. Since $\mathrm{vol}_{\mathcal{L}}(S_i) \leq \mathrm{vol}_{\mathcal{L}}(\bar{S}_i)$ for all $1 \leq i \leq r$, we have

$$\mathrm{cut}(S_i, \bar{S}_i) \geq \phi_{\mathcal{L}} \cdot \mathrm{vol}_{\mathcal{L}}(S_i) \quad (12)$$

By orienting vertices according to $v_1, ..., v_n$, we can express the numerator

$$\mathrm{Num} = \left(\sum_{u,v}(g_+^2[u] - g_+^2[v])w_{u,v}\right)^2$$

$$= \left(\sum_{i<j}\sum_{k=0}^{j-i-1}\left(g_+^2[v_{i+k}] - g_+^2[v_{i+k+1}]\right)w_{v_i,v_j}\right)^2$$

$$= \left(\sum_{i=1}^{n-1}\left(g_+^2[v_i] - g_+^2[v_{i+1}]\right) \cdot \mathrm{cut}(S_i, \bar{S}_i)\right)^2$$

Collecting (v_i, v_{i+1}) terms

$$\geq \left(\sum_{i=1}^{n-1}\left(g_+^2[v_i] - g_+^2[v_{i+1}]\right) \cdot \phi_{\mathcal{L}} \cdot \mathrm{vol}_{\mathcal{L}}(S_i)\right)^2$$

By Eqn: 12

$$= \phi_{\mathcal{L}}^2 \cdot \left(\sum_{i=1}^{n} g_+^2[v_i] \cdot (\mathrm{vol}_{\mathcal{L}}(S_i) - \mathrm{vol}_{\mathcal{L}}(S_{i+1}))\right)^2$$

By Eqn. 11 and $g_+(v_n) = 0$

$$= \phi_{\mathcal{L}}(G)^2 \cdot \left(\sum_{i=1}^{n} g_+^2[v_i] \cdot d_{v_i} \tau_i\right)^2.$$

Combining the bounds for the numerator and the denominator, we obtain $\lambda_1 \leq \phi_{\mathcal{L}}^2/2$ as stated in the theorem. The right hand side of the theorem follows from the same argument for the standard Cheeger Inequality. $\quad\square$

3.2 Spectral Partitioning for Generalized Conductance

Given a weighted graph $G = (V, E, \boldsymbol{A})$ and a operator $\mathcal{L}$, we can use the standard sweeping method in the proof of Theorem 1 to find a partition $(S, \bar{S})$. This procedure is described in Algorithm 1.

Algorithm 1 Spectral Dynamics Clustering $(G, \mathcal{L})$

Input: weighted network: $G = (V, E, \boldsymbol{A})$, and spreading operator $\mathcal{L}$ defined by the interaction matrix $\boldsymbol{W}$ and the node delay factor $\boldsymbol{T}$.
Output partition: $(S, \bar{S})$
Algorithm

- Find the eigenvector $\boldsymbol{f}$ of $\mathcal{L} = \boldsymbol{T}^{-1/2}\boldsymbol{D}_{\boldsymbol{W}}^{-1/2}(\boldsymbol{D}_{\boldsymbol{W}} - \boldsymbol{W})\boldsymbol{D}_{\boldsymbol{W}}^{-1/2}\boldsymbol{T}^{-1/2}$ associated with the second smallest eigenvalue of $\mathcal{L}$.

- Let vector $\boldsymbol{g}$ be $g[u] = f[u] \cdot \sqrt{d_{\boldsymbol{W}_u}\tau_u}$.

- Order the vertices of G into $(v_1, ..., v_n)$ such that $g[v_1] \geq g[v_2] \geq ... \geq g[v_n]$.

- Sweeping: For each $S_i = \{v_1, ..., v_i\}$, compute

$$h_{\mathcal{L}}(S_i) = \frac{\text{cut}(S_i, \bar{S}_i)}{\min\left(\text{vol}_{\mathcal{L}}(S_i), \text{vol}_{\mathcal{L}}(\bar{S}_i)\right)}.$$

- Output the S_i with the smallest $h_{\mathcal{L}}(S_i)$.

Before stating the quality guarantee of the above algorithm, we quickly discuss its implementation and running time. The most expensive step is the computation of the eigenvalue vector f associated with the second smallest eigenvalue of $\mathcal{L}$. While one can use standard numerical methods to find an approximation of this eigenvector – the analysis would depend on the separation of the second and the third eigenvalue of $\mathcal{L}$. Since $\mathcal{L}$ is a diagonally scaled normalized Laplacian matrix, one can use the nearly-linear-time Laplacian solvers (e.g., by Spielman-Teng [26] or Koutis-Miller-Peng [14]) to solve linear systems in $\mathcal{L}$.

Following [26], let us consider the following notion of spectral approximation of $\mathcal{L}$: Suppose $\lambda_1(\mathcal{L})$ the second smallest eigenvalue of $\mathcal{L}$. For $\epsilon \geq 0$, $\bar{f}$ is an *ϵ-approximate second eigenvector* of $\mathcal{L}$ if $\bar{f} \perp \boldsymbol{D}^{1/2}\boldsymbol{T}^{1/2} \cdot \boldsymbol{1}$, and

$$\frac{\bar{f}^T \mathcal{L} \bar{f}}{\bar{f}^T \bar{f}} \leq (1 + \epsilon) \cdot \lambda_1(\mathcal{L}).$$

The following proposition follows directly from the algorithm and Theorem 7.2 of [26] (using the solver from [14]).

PROPOSITION 1. *For any interaction graph $G = (V, E, \boldsymbol{W})$ and node scaling factor $\boldsymbol{T}$, and $\epsilon, p > 0$, with probability at least $1 - p$, one can compute an ϵ-approximate second eigenvector of operator $\mathcal{L}$ in time*

$$O\left(|E| \log n \log \log n \log(1/p) \log(1/\epsilon)/\epsilon\right).$$

To use this spectral approximation algorithm (and in fact any numerical approximation to the second eigenvector of $\mathcal{L}$) in our spectral partitioning algorithm for the dynamics, we will need a strengthened theorem of Theorem 1.

THEOREM 2. *(Extended Cheeger Inequality with Respect to Rayleigh Quotient)*
For any interaction graph $G = (V, E, \boldsymbol{W})$ and node scaling factor $\boldsymbol{T}$, (whose diagonals are $(\tau_1, ..., \tau_n)$), for any vector $\boldsymbol{u}$ such that $\boldsymbol{u} \perp \boldsymbol{D}^{1/2}\boldsymbol{T}^{1/2} \cdot \boldsymbol{1}$, if we order the vertices of G into $(v_1, ..., v_n)$ such that $g[v_1] \geq ... \geq g[v_n]$, where $\boldsymbol{g} = (\boldsymbol{DT})^{-1/2} \cdot \boldsymbol{u}$ then

$$\frac{(\min_i h_{\mathcal{L}}(S_i))^2}{2} \leq \frac{\boldsymbol{u}^T \mathcal{L} \boldsymbol{u}}{\boldsymbol{u}^T \boldsymbol{u}},$$

where $\mathcal{L} = \boldsymbol{T}^{-1/2}\boldsymbol{D}_{\boldsymbol{W}}^{-1/2}(\boldsymbol{D}_{\boldsymbol{W}} - \boldsymbol{W})\boldsymbol{D}_{\boldsymbol{W}}^{-1/2}\boldsymbol{T}^{-1/2}$ and $S_i = \{v_1, ..., v_i\}$.

The next theorem then follows directly from Proposition 1, Theorem 2 and the definition of ϵ-approximate second eigenvector of $\mathcal{L}$ that provide a guarantee of the quality of the algorithm of this subsection.

THEOREM 3. *For any interaction graph $G = (V, E, \boldsymbol{W})$ and node scaling factor $\boldsymbol{T}$, (whose diagonals are $(\tau_1, ..., \tau_n)$), one can compute in time*

$$O(|E| \log n \log \log n \log(1/\epsilon)/\epsilon)$$

a partition $(S, \bar{S})$ such that

$$
\begin{aligned}
h_{\mathcal{L}}(S) &= \frac{\sum_{v \in S, u \in \bar{S}} w_{u,v}}{\min\left(\sum_{v \in S} d_{\boldsymbol{W}_v}\tau_v, \sum_{v \in \bar{S}} d_{\boldsymbol{W}_v}\tau_v\right)} \\
&\leq \sqrt{2(1 + \epsilon)\lambda_1(\mathcal{L})}
\end{aligned}
$$

where $\boldsymbol{T}^{-1/2}\boldsymbol{D}_{\boldsymbol{W}}^{-1/2}(\boldsymbol{D}_{\boldsymbol{W}} - \boldsymbol{W})\boldsymbol{D}_{\boldsymbol{W}}^{-1/2}\boldsymbol{T}^{-1/2}$, $w_{u,v}$ is the $(u, v)^{th}$ entry of the interaction matrix $\boldsymbol{W}$, and $\lambda_1(\mathcal{L})$ is the second smallest eigenvalue of $\mathcal{L}$. Consequently,

$$h_{\mathcal{L}}(S) \leq 2\sqrt{(1 + \epsilon)\phi_{\mathcal{L}}(G)}$$

$$= 2(1 + \epsilon)\sqrt{\min_{S^* \in V} \frac{\sum_{v \in S^*, u \in \bar{S}^*} w_{u,v}}{\min\left(\sum_{v \in S^*} d_{\boldsymbol{W}_v}\tau_v, \sum_{v \in \bar{S}^*} d_{\boldsymbol{W}_v}\tau_v\right)}}.$$

4. LOCAL CLUSTERING ALGORITHMS

In the analysis of massive networks, it is essential to identify subsets of nodes whose induced sub-graphs have "significant" structural coherence without examining the entire network. Using clustering as an example, Spielman and Teng [26] introduced a framework of *local algorithms* for network analysis: given an input node, a local algorithm can only explore the neighbors of the nodes it has already accessed. It may occasionally access some random nodes in the network. The complexity of the local algorithm is then measured by the total number of accesses it makes, as well as the computations it performs.

In spite of lacking global access, Spielman and Teng showed that local algorithms can be effective in identifying good clusters when conductance is used as quality measure. Their work was subsequently improved by Andersen-Chung-Lang [2], Andersen-Pere [3], and Chung [6, 7]. Chung [6, 7] applied the sweeping method to the heat-kernal page rank to identify a random walk-based local cluster. Similarly, we propose a

dynamic-dependent sweeping procedure to discover a local cluster of provably good generalized conductance. Due to length limit, we only state relevant theorems in this paper. For details of proofs, please refer to the full version on arXiv.

4.1 Convergence and Local $h_{\mathcal{L}}$-Clustering

For an starting vector $\boldsymbol{\mu}$, let $\theta_{t,\boldsymbol{\mu}}[i]$ denote the value of the dynamics of vertex $i \in V$ at time t when the initial vector is $\boldsymbol{\mu}$. In other words, $\boldsymbol{\theta}_{t,\boldsymbol{\mu}} = (\theta_{t,\boldsymbol{\mu}}[1], ..., \theta_{t,\boldsymbol{\mu}}[n])$ is the solution of Eq. 4 with $\boldsymbol{\theta}_0 = \boldsymbol{\mu}$.

Let $\boldsymbol{\pi} = (\pi[1], ..., \pi[n])$ denote

$$\boldsymbol{\pi} = \left(\sqrt{d_{\boldsymbol{W}_1}\tau_1}, ..., \sqrt{d_{\boldsymbol{W}_n}\tau_n} \right).$$

If the dynamic process converges when starting from $\boldsymbol{\mu}$, then $\lim_{t\to\infty} \boldsymbol{\theta}_{t,\boldsymbol{\mu}}$ is proportional to $\boldsymbol{\pi}$, i.e., for a constant z depending only on $\boldsymbol{\mu}$,

$$\lim_{t\to\infty} \theta_{t,\boldsymbol{\mu}}[i] = z_{\boldsymbol{\mu}} \cdot \pi[i] = z_{\boldsymbol{\mu}} \cdot \sqrt{d_{\boldsymbol{W}_i}\tau_i}. \tag{13}$$

We will also use $\boldsymbol{\pi}_{\boldsymbol{\mu}}$ to denote $\lim_{t\to\infty} \boldsymbol{\theta}_{t,\boldsymbol{\mu}}$.

In this spirit, we define the *fractional volume* of a subset $S \subseteq V$ be $\mathrm{fvol}_{\mathcal{L}}(S) = \sum_{j \in S} \sqrt{d_{\boldsymbol{W}_j}\tau_j}$. Following the analysis of Theorem 1, let $\tau_{min} = \min_i \tau_i$, $\tau_{max} = \max_i \tau_i$ and $\boldsymbol{\mu}_S[i] = \frac{\sqrt{d_{\boldsymbol{W}_i}\tau_i}}{\sum_j \sqrt{d_{\boldsymbol{W}_j}\tau_j}}$ for $i \in S$ and is 0 otherwise. Like in the traditional Laplacian-conductance framework, we will establish that the existence of a community $S \subset V$ with small $h_{\mathcal{L}}(S)$ underscores why the dynamics may not converge rapidly from the inital distribution.

We examine the evolution of the dynamics defined by a network $G = (V, E, \boldsymbol{A})$ and a spreading operator $\mathcal{L}$. Particularly, we are interested in estimating the rate that the dynamic process converges to its stationary distribution from some initial distribution, and its relation to the generalized conductance $h_{\mathcal{L}}$ that we defined (Eqn. (7)). As each dynamic process leads to its own notion of (time-dependent) centrality, we will study and compare the corresponding community structures by its generalized conductance.

Our local clustering algorithm (Algorithm 2) explores this interplay. Here we assume the starting vector $\boldsymbol{\mu}_u$ is seeded in a single node u, i.e., it has a single one entry for node u, and other entries are all zeros. Next theorem provides a performance guarantee for the local algorithm.

THEOREM 4 (LOCAL $h_{\mathcal{L}}$-CLUSTERING). *For any* $0 \le \phi \le 1$, *suppose S is a subset in a interaction graph* $G = (V, E, \boldsymbol{W})$ *with* $\mathrm{vol}_{\mathcal{L}}(S) \le \frac{1}{12} \cdot \frac{\tau_{min}}{\tau_{max}} \cdot \mathrm{vol}_{\mathcal{L}}(V)$ *and* $h_{\mathcal{L}}(S) \le \phi^2$. *Then, there exists a subset $S_S \subseteq S$ with* $\mathrm{fvol}_{\mathcal{L}}(S_S) \ge \mathrm{fvol}_{\mathcal{L}}(S)/4$ *such that for any $u \in S_S$, Algorithm 2 with $s \ge 2\mathrm{fvol}_{\mathcal{L}}(S)$ and $\epsilon = \phi^2 \cdot \frac{\tau_{min}}{\tau_{max}} \cdot \frac{1}{10 \cdot s}$ will in time $O\left(s\phi^{-2} \cdot \frac{\tau_{max}}{\tau_{min}} \cdot \log n \right)$ return a set S_S with the properties that (1) the conductance of S_S is at most* $O\left((1+\epsilon)\phi \cdot \sqrt{\log s + \log \frac{\tau_{max}}{\tau_{min}}} \right)$ *and (2)* $\mathrm{fvol}(S_S) \le 2s$.

Theorem 4 follows from our proofs in the following subsection 3 together with Andersen-Chung-Lang's analysis [2] of the push-based approximation scheme. Similar to its global counterpart in Theorem 3, Theorem 4 uses Theorem 2 in place of Theorem 1 to address the approximation of the dynamic process. During the approximate dynamic process with the Andersen-Chung-Lang push scheme, only $O(1/\epsilon)$ nodes have non-zero entries.

Algorithm 2 Local $h_{\mathcal{L}}$-Clustering$(G, \mathcal{L}, u, \phi, s, \epsilon)$

Input: Network: $G = (V, E, \boldsymbol{A})$, spreading operator $\mathcal{L}$ defined by the interaction matrix $\boldsymbol{W}$ and the node delay factor $\boldsymbol{T}$. The starting node: u, quality bound: ϕ, targeted fractional volume: s, and rounding approximation ϵ
Output: subset S_S
Algorithm

1. Set $t = \phi^{-2} \cdot \tau_{min} \cdot \log 2$.

2. Apply the push-based approximation scheme of Andersen-Chung-Lang [2] with rounding parameter ϵ to find an approximation $\tilde{\boldsymbol{\theta}}_{t,u,\epsilon}$ to $\boldsymbol{\theta}(t,u) = e^{-\mathcal{L}t}\boldsymbol{\mu}_u$ and adaptively maintain the set $\tilde{V}_\epsilon = \{v : \tilde{\theta}_{t,u,\epsilon}[v] > 0\}$.

3. Order the vertices of $\tilde{V}_\epsilon$ into $\{v_1, v_2, \cdots\}$ such that $\frac{\theta_{t,u}[v_1]}{\sqrt{d_{\boldsymbol{W}_1}\tau_1}} \ge \frac{\theta_{t,u}[v_2]}{\sqrt{d_{\boldsymbol{W}_2}\tau_2}} \ge \cdots \ge \frac{\theta_{t,u}[v_n]}{\sqrt{d_{\boldsymbol{W}_n}\tau_n}}$

4. Sweeping: For each $S_i = S_{t,u}[i] = \{v_1, ..., v_i\}$ such that $\mathrm{fvol}S_i \le 2s$, compute

$$h_{\mathcal{L}}(S_i) = \frac{\mathrm{cut}(S_i, \bar{S}_i)}{\min\left(\mathrm{vol}_{\mathcal{L}}(S_i), \mathrm{vol}_{\mathcal{L}}(\bar{S}_i)\right)}.$$

5. Output the S_i with the smallest $h_{\mathcal{L}}(S_i)$.

4.2 Rate of Convergence and Local Sweeping

Define $\Theta_{t,\boldsymbol{\mu}}(S) = \sum_{i \in S} \theta_{t,\boldsymbol{\mu}}[i]\sqrt{\frac{d_{\boldsymbol{W}_i}}{\tau_i}}$. We first establish a lemma showing that while the quantity $\Theta_{t,\boldsymbol{\mu}}(S)$ is reducing during every step of the dynamic process, the derivative depends on the generalized volume of S, and is bounded by a factor proportional to the generalized conductance $h_{\mathcal{L}}(S)$ of S given by Eqn. (7).

LEMMA 1. *For $S \subset V$ with $\mathrm{fvol}_{\mathcal{L}}(S) \le \mathrm{fvol}_{\mathcal{L}}(V)/2$ we have $\frac{d\Theta_{t,\boldsymbol{\mu}}(S)}{dt} \le 0$. Moreover,* $\left| \frac{d\Theta_{t,\boldsymbol{\mu}}(S)}{dt} \right| \le \frac{h_{\mathcal{L}}(S)}{\tau_{min}}$.

Lemma 1 then provides a lower bound on the fractional volume of a subset $S_S \subset S$ from whom the dynamic process does not converge rapidly. Let $\boldsymbol{\theta}_{t,u}$ denote $\boldsymbol{\theta}_{t,\boldsymbol{\mu}_u}$, $\boldsymbol{\pi}_u$ denote $\boldsymbol{\pi}_{\boldsymbol{\mu}_u}$, $\Theta_{t,u}(S) = \Theta_{t,\boldsymbol{\mu}_u}(S)$.

COROLLARY 1. *For any subset $S \subset V$ in a interaction graph $G = (V, E, \boldsymbol{W})$, let*

$$S_S = \left\{ u \in S : \Theta_{t,u}(S) \ge \frac{1}{4} \cdot \Theta_{0,u}(S) \cdot e^{-\frac{1}{\tau_{min}} \cdot h_{\mathcal{L}}(S)}t \right\}.$$

If $t \ge \frac{\tau_{min}}{h_{\mathcal{L}}(S)}$ then $\mathrm{fvol}_{\mathcal{L}}(S_S) \ge \mathrm{fvol}_{\mathcal{L}}(S)/4$.

Following the intuition of the proof of the extended Cheeger inequality (Theorem 1) and the traditional Laplacian analysis [17, 26, 8, 7], we consider a sweeping process based on the time-dependent centrality with starting vertex u. At time t, the sweeping is performed according to

$$\frac{\theta_{t,u}[v_1]}{\sqrt{d_{\boldsymbol{W}_1}\tau_1}} \ge \frac{\theta_{t,u}[v_2]}{\sqrt{d_{\boldsymbol{W}_2}\tau_2}} \ge \cdots \ge \frac{\theta_{t,u}[v_n]}{\sqrt{d_{\boldsymbol{W}_n}\tau_n}}$$

by considering $S_{t,u}[i] = \{v_1, \cdots, v_i\}$. Like in Chung [6], for a targeted size s, we define the *s-local $h_{\mathcal{L}}$-value* denoted by $\kappa_{t,u,s}$ as: $\kappa_{t,u,s} = \min_{i:\mathrm{fvol}_{\mathcal{L}}(S_{t,u}[i]) \le 2s} h_{\mathcal{L}}(S_{t,u}[i])$.

LEMMA 2. *For any subset $S \subset V$ in a interaction graph $G = (V, E, \boldsymbol{W})$ of fractional volume $s = \mathrm{fvol}_{\boldsymbol{\mathcal{L}}}(S) \leq \mathrm{fvol}_{\boldsymbol{\mathcal{L}}}(V)/4$, we have*

$$\Theta_{t,u}(S) - \Pi_u(S) \leq \frac{1}{\tau_{min}} \mathrm{fvol}_{\boldsymbol{\mathcal{L}}}(S) \cdot e^{-t \frac{\kappa^2_{t,u,s}}{4}}. \qquad (14)$$

.

The following lemma is a consequence of Lemma 2 and Corollary 1, which ultimately leads to Theorem 4.

LEMMA 3 (LOCAL $h_{\boldsymbol{\mathcal{L}}}$-SWEEPING). *For any $0 \leq \phi \leq 1$, suppose S is a subset in a interaction graph $G = (V, E, \boldsymbol{W})$ with $\mathrm{vol}_{\boldsymbol{\mathcal{L}}}(S) \leq \frac{1}{12} \cdot \frac{\tau_{min}}{\tau_{max}} \cdot \mathrm{vol}_{\boldsymbol{\mathcal{L}}}(V)$ and $h_{\boldsymbol{\mathcal{L}}}(S) \leq \phi^2$. Then, there exists a subset $S_S \subseteq S$ with $\mathrm{fvol}(S_S) \geq \mathrm{fvol}(S)/4$ such that for any $u \in S_S$, the sweeping by using the vector $\boldsymbol{\theta}_{t,u}$ with $t = O(\phi^{-2} \cdot \tau_{min})$ will find a set S_S with s-local $h_{\boldsymbol{\mathcal{L}}}$-value at most $O\left(\phi \cdot \sqrt{\log s + \log(\tau_{max}/\tau_{min})}\right)$.*

5. EXPERIMENTS

Through experiments we demonstrate the difference in the centralities and communities detected in a graph under different dynamic processes. Interestingly, even this simple class of processes can lead to divergent views about who the central nodes are and what are the cohesive clusters for a collection of widely studied real-world networks.

We study how the dynamic processes defined in this paper affect centrality and spectral partitioning. Our framework actually offers more freedom in designing node delay factor $\boldsymbol{T}$ and the interaction graph $\boldsymbol{W}$ and is thus a much more powerful tool. We will study these possibilities in greater detail and use the proposed local clustering algorithm for real-world applications in future work.

Table 1 lists the networks we study empirically, and their properties. We treat all networks as undirected. These networks come from different domains, and embody a variety of dynamic processes and interactions, from real world friendships (Zachary karate club [30]), to social network (Facebook [18]), to electrical power distribution (Power Grid [28]), to intra-text (Word Adjacency [19]) and inter-text (Political Blogs [1]) links.

The House of Representatives network is built from the 98th United States House of Representatives voting data [21]. Unlike the previously studied variants, here we use a special version taking account of all 908 votes. The resulting network is dense and has an unusually flat degree distribution. Originally analyzed by Smith et al. [24], this network better differentiates between the dynamics under our framework.

5.1 Centrality Profiles

We first study node (stationary) centrality rankings resulting from different dynamic processes. By construction, centrality of a node converges to Eq. 5. Given a graph, the degree sequence is fixed, and stationary centrality depends solely on the node delay factor $\boldsymbol{T}$. A centrality profile is a curve of stationary centrality values for each node given a spreading operator. Figure 1 shows the *centrality profiles* of network nodes. To improve visualization, nodes are ordered by their centrality according to the normalized Laplacian matrix, and they are rescaled to fall within the same range.

Except for "House of Representatives" and "Power Grid" networks, the centrality profiles on the other four data sets are very similar. They all have heavy-tailed degree distributions, thus their centralities under the normalized Laplacian
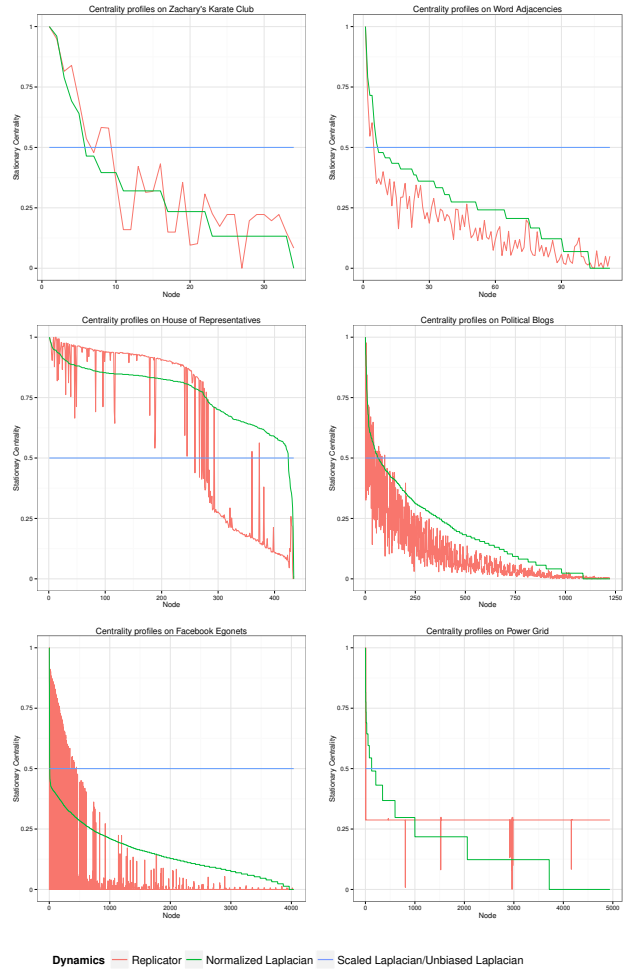

Figure 1: Centrality profiles for different networks
Comparison of stationary centrality on all six networks. The x-axes represent node indices which has been fixed for all dynamics on each network in descending order of the "Normalized Laplacian" centrality. On the y-axes, the convergent centrality has been normalized to the interval [0,1] for all dynamics on all networks.
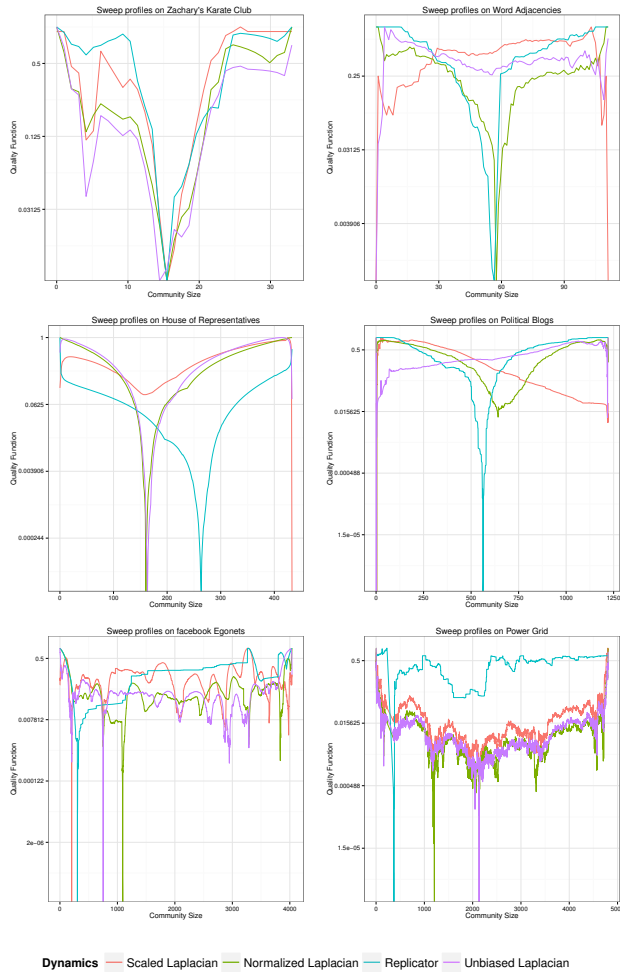

Figure 2: Optimal bisections for $\boldsymbol{\mathcal{L}}$ (left) and $\boldsymbol{R}$ (right) on Political Blogs
Laplacian has the exact same bisection as Normalized Laplacian, while Unbiased Adjacency leads to similar results as Replicator. Notice the correspondence to their sweep profiles.

Table 1: Specifications of benchmark networks

Name	#nodes	#edges	Diameter	Clustering	Properties
Zachary's Karate Club	34	78	5	0.588	A toy social network
Word Adjacency	112	425	5	0.19	A disassortative network
House of Representatives	434	51033	4	0.882	A dense network with even degrees
Political Blogs	1490	16714	9	0.21	Has two "cores" with "whiskers"
Facebook Egonets	4039	88234	17	0.303	Combined ego networks
Power Grid	4941	6594	46	0.107	A large sparse network

behave similarly. Replicator centralities follow largely the same trend. There are important differences, however. For example, in the "House of Representatives" network, nodes in the range $[1, 50]$ considered to be highly central according to the normalized Laplacian are judged not as important by the Replicator. Since centralities and communities are closely related under our framework, this would eventually lead to the differences in community detection, as we show in the next subsection. The centrality scores of nodes in the "Power Grid" network are all the same except for a few nodes. The lack of information comes from the extreme sparsity of the network.

5.2 Communities and Sweep Profiles

We use the *sweep profile* to study differences in network partitioning using different spreading operators for the real world networks listed in Table 1. Given a spreading operator, a sweep profile is a curve of quality functions following Algorithm 1. Let f be the eigenvector associated with the second smallest eigenvalue of $\mathcal{L}$. The vertices are ordered $(v_1, ..., v_n)$ such that $g[v_1] \geq ... \geq g[v_n]$ where $g[u] = f[u]/\sqrt{d_{\boldsymbol{W}_u}\tau_u}$. The x-axis plots $S_i = \{v_1, ..., v_i\}$. The y-axis plots

$$h_{\mathcal{L}}(S_i) = \frac{\mathrm{cut}(S_i, \bar{S}_i)}{\min\left(\mathrm{vol}_{\mathcal{L}}(S_i), \mathrm{vol}_{\mathcal{L}}(\bar{S}_i)\right)}.$$

The smallest $h_{\mathcal{L}}(S_i)$ gives a provably good cluster S_i according to Theorem 3.

Figure 3 shows the sweep profile for each network under different dynamics. To improve visualization, we rescale community profiles to lie within the same range. Sweep profiles provide us an interesting perspective into the differences in the communities identified according to different dynamics. With the exception of Zachary's karate club, there are large differences both in the shape of the profile and, more importantly, in the location of its minimum, which corresponds to the optimal bisection of the network.

In the "Political Blogs" network, both regular and normalized versions of Laplacian split a "whisker" community from the core (Left of Figure 2). However, the Replicator identifies a community of more than 500 nodes (Right of Figure 2). While Replicator is less susceptible to "whiskers", on the "Power Grid" and "Facebook Egonets" networks, it only identifies small communities within the "core". Instead, unbiased Laplacian succeeds in finding more balanced cuts of the networks. In the "House of Representatives" network, the Replicator puts "swing nodes" into a different community from the other operators. Recall that this difference was also reflected in its centrality profiles. In future work, we plan to analyze the reasons behind these differences in more detail and in a dynamic setting.

Figure 3: Sweep profiles for different networks
Sweep profiles produced according to Algorithm 1 on all six networks. The x-axes indicate the sweep point, or community size of one of the bi-sections. On the logarithmic y-axes, the quality function has been normalized to the same range for all dynamics on each network.

6. CONCLUSION

The dynamics of a process occurring on a network can be succinctly described in terms of a spreading operator. The generalized Laplacian operator presented in this paper can describe the spreading operators associated with known processes, such as random walks and epidemics, but also new ones, such as the unbiased Laplacian. We generalize the relationships between the properties of random walks and normalized graph Laplacian, to other dynamic processes. Each operator leads to a distribution that gives centrality of nodes with respect to that process. In addition, we show that the generalized conductance measuring community quality with respect to the dynamic process is related to the eigenvalues of the spreading operator describing that process through a Cheeger-like inequality. These relationships can be used for spectral graph partitioning. Nodes within the same partition interact more with each other via the dynamic process than with nodes in other partitions. As in previous local clustering algorithms of Spielman-Teng, Andersen-Chung-Lang and Andersen-Peres [26, 3, 2], the mathematical structure underlying the generalized Cheeger's inequality enables efficient local clustering algorithms for the more general community qualities, whose running time is essentially linear in the size of the cluster it outputs, and does not depend on the size of the whole graph.

While our framework is flexible enough to represent several important types of dynamic processes, it does not represent all possible processes, for example, those that are locally non-conservative. In order to describe such dynamics, an even more general framework is needed. We conjecture, however, that the more general spreading operators will still obey the Cheeger-like inequality, and that other theorems presented in this paper can be extended to these processes.

Acknowledgments

7. REFERENCES

[1] L. A. Adamic and N. Glance. The political blogosphere and the 2004 US election: divided they blog. In *Proceedings of the 3rd international workshop on Link discovery*, pages 36–43. ACM, 2005.

[2] R. Andersen, F. Chung, and K. Lang. Using PageRank to Locally Partition a Graph. *Internet Math.*, 4(1):1–128, 2007.

[3] R. Andersen and Y. Peres. Finding Sparse Cuts Locally Using Evolving Sets. In *STOC*, pages 235–244. ACM, 2009.

[4] P. Bonacich and P. Lloyd. Eigenvector-like measures of centrality for asymmetric relations. *Social Networks*, 23(3):191–201, 2001.

[5] S. Borgatti. Centrality and network flow. *Social Networks*, 27(1):55–71, Jan. 2005.

[6] F. Chung. The heat kernel as the pagerank of a graph. *Proceedings of the National Academy of Sciences*, 104(50):19735–19740, Dec. 2007.

[7] F. Chung. A local graph partitioning algorithm using heat kernel pagerank. *Internet Mathematics*, 6(3):315–330, Jan. 2009.

[8] F. R. K. Chung. *Spectral Graph Theory (CBMS Regional Conference Series in Mathematics, No. 92)*. American Mathematical Society, Feb. 1997.

[9] J.-C. Delvenne, S. N. Yaliraki, and M. Barahona. Stability of graph communities across time scales. *ArXiv e-prints*, Dec. 2008.

[10] R. Ghosh and K. Lerman. Parameterized centrality metric for network analysis. *Physical Review E*, 83(6):066118, June 2011.

[11] M. Jerrum and A. Sinclair. Conductance and the rapid mixing property for Markov chains: the approximation of permanent resolved. In *STOC*, pages 235–244. ACM, 1988.

[12] R. Kannan, S. Vempala, and A. Vetta. On Clusterings: Good, Bad and Spectral. *J. ACM*, 51(3):497–515, May 2004.

[13] D. Kempe, J. Kleinberg, and E. Tardos. Maximizing the spread of influence through a social network. In *KDD '03*, pages 137–146. ACM, 2003.

[14] I. Koutis, G. L. Miller, and R. Peng. Approaching Optimality for Solving SDD Linear Systems. In *FOCS*, pages 235–244. IEEE, 2010.

[15] R. Lambiotte, R. Sinatra, J.-C. Delvenne, T. S. Evans, M. Barahona, and V. Latora. Flow graphs: Interweaving dynamics and structure. \pre, 84(1):017102, July 2011.

[16] K. Lerman and R. Ghosh. Network Structure, Topology and Dynamics in Generalized Models of Synchronization. *Physical Review E*, 86(026108), 2012.

[17] L. Lovász. *Random Walks on Graphs: A Survey*, pages 353–397. 1993.

[18] J. J. McAuley and J. Leskovec. Learning to Discover Social Circles in Ego Networks. In *NIPS*, volume 272, pages 548–556, 2012.

[19] M. E. J. Newman. Finding community structer in networks using the eigenvectors of matrices. *Physical Review E*, 74(3), 2006.

[20] L. Page, S. Brin, R. Motwani, and T. Winograd. The pagerank citation ranking: Bringing order to the web, 1999.

[21] K. Poole. Voteview Website. http://voteview.com/house98.htm.

[22] M. Rosvall and C. T. Bergstrom. Maps of random walks on complex networks reveal community structure. *Proceedings of the National Academy of Sciences*, 105(4):1118–1123, Jan. 2008.

[23] J. Shi and J. Malik. Normalized Cuts and Image Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(8):888–905, 2000.

[24] L. M. Smith, K. Lerman, C. Garcia-Cardona, A. G. Percus, and R. Ghosh. Spectral Clustering with Epidemic Diffusion. *CoRR*, abs/1303.2663, 2013.

[25] D. A. Spielman and S.-H. Teng. Spectral partitioning works: Planar graphs and finite element meshes. In *IEEE FOCS*, pages 96–105, 1996.

[26] D. A. Spielman and S.-H. Teng. Nearly-linear Time Algorithms for Graph Partitioning, Graph Sparsification, and Solving Linear Systems. In *STOC*, pages 81–90. ACM, 2004.

[27] Y. Wang, D. Chakrabarti, C. Wang, and C. Faloutsos. Epidemic spreading in real networks: an eigenvalue viewpoint. In *Proceedings of the 22nd Symposium on Reliable Distributed Systems*, pages 25–34, Los Alamitos, CA, USA, Oct. 2003. IEEE.

[28] D. J. Watts and S. H. Strogatz. Collective dynamics of 'small-world'networks. *nature*, 393(6684):440–442, 1998.

[29] M. A. Yildirim, K.-I. Goh, M. E. Cusick, A.-L. Barabasi, and M. Vidal. Drug-target network. *Nat Biotech*, 25:1119–1126, Oct. 2007.

[30] W. W. Zachary. An Information Flow Model for Conflict and Fission in Small Groups. *Journal of Anthropological Research*, 33(4):452–473, 1977.