IMPROVED CHEEGER'S INEQUALITY AND ANALYSIS OF LOCAL GRAPH PARTITIONING USING VERTEX EXPANSION AND EXPANSION PROFILE[*]

TSZ CHIU KWOK[†], LAP CHI LAU[†], AND YIN TAT LEE[‡]

Abstract. We prove two generalizations of the Cheeger's inequality. The first generalization relates the second eigenvalue to the edge expansion and the vertex expansion of the graph G, $\lambda_2 = \Omega(\phi^V(G)\phi(G))$, where $\phi^V(G)$ denotes the robust vertex expansion of G and $\phi(G)$ denotes the edge expansion of G. The second generalization relates the second eigenvalue to the edge expansion and the expansion profile of G, for all $k \geq 2$, $\lambda_2 = \Omega(\phi_k(G)\phi(G)/k)$, where $\phi_k(G)$ denotes the k-way expansion of G. These show that the spectral partitioning algorithm has better performance guarantees when $\phi^V(G)$ is large (e.g., planted random instances) or $\phi_k(G)$ is large (instances with few disjoint nonexpanding sets). Both bounds are tight up to a constant factor. Our approach is based on a method to analyze solutions of Laplacian systems, and this allows us to extend the results to local graph partitioning algorithms. In particular, we show that our approach can be used to analyze personal pagerank vectors and to give a local graph partitioning algorithm for the small-set expansion problem with performance guarantees similar to the generalizations of Cheeger's inequality. We also present a spectral approach to prove similar results for the truncated random walk algorithm. These show that local graph partitioning algorithms almost match the performance of the spectral partitioning algorithm, with the additional advantages that they apply to the small-set expansion problem and their running time could be sublinear. Our techniques provide common approaches to analyze the spectral partitioning algorithm and local graph partitioning algorithms.

Key words. spectral graph theory, graph expansion, Cheeger's inequality, random walks

AMS subject classifications. 05C50, 05C81, 05C85

DOI. 10.1137/16M1079816

1. Introduction. Let $G = (V, E)$ be a complete weighted graph and $n := |V|$. For simplicity, we assume that the graph is regular and the total weight on each vertex is one throughout.[1] Let $w(S, T)$ be the total weight of the edges with one vertex in S and another vertex in T. The edge expansion of a set $S \subseteq V$ and the edge expansion of a graph G are defined as

$$\phi(S) := \frac{w(S, \overline{S})}{|S|} \quad \text{and} \quad \phi(G) := \min_{S:|S| \leq |V|/2} \phi(S),$$

where $\overline{S} := V - S$ is the complement set. Let $L = I - A$ be the Laplacian matrix of G, where I and A are the identity and the adjacency matrix of G, with eigenvalues

890

of L being $0 = \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n \leq 2$. Cheeger's inequality [Che70, AM85, Alo86] bounds the edge expansion of G using the second eigenvalue of L,

$$\frac{1}{2}\lambda_2 \leq \phi(G) \leq \sqrt{2\lambda_2}.$$

It is useful in bounding the edge expansion of a graph and also bounding the mixing time of random walks [HLW06]. The proof of Cheeger's inequality gives an efficient algorithm to find a set with expansion at most $\sqrt{2\lambda_2}$, and we will refer to this algorithm as the spectral partitioning algorithm (also known as the sweep cut algorithm on the second eigenvector). A recent generalization [KLLOT13] of Cheeger's inequality bounds the edge expansion of G using the second and the kth eigenvalues of L for any $k \geq 2$,

$$\phi(G) = O(k)\frac{\lambda_2}{\sqrt{\lambda_k}}.$$

This provides a better analysis of the spectral partitioning algorithm in practical instances of image segmentation and data clustering.

1.1. Our results. We prove two new generalizations of Cheeger's inequality. These provide better analyses of the spectral partitioning algorithm when some expansion parameters of the graph are large. We also prove similar bounds for the personal pagerank algorithm and the truncated random walk algorithm. These give local graph partitioning algorithms for the small-set expansion problem with improved Cheeger's guarantees. Our techniques provide common approaches to analyzing the spectral partitioning algorithm and local graph partitioning algorithms.

1.1.1. Vertex expansion. The first generalization bounds the second eigenvalue of L by the edge expansion and the vertex expansion of G. We define the robust vertex expansion by a slight modification of that from Kannan, Lovász, and Montenegro [KLM06]. For $S \subseteq V$, let $N_{1/2}(S) := \max\{|T| \mid T \subseteq \overline{S} \text{ and } w(S,T) \leq \frac{1}{2}w(S,\overline{S})\}$. In [KLM06], the definition is $\tilde{N}_{1/2}(S) = \min\{|T| \mid T \subseteq \overline{S} \text{ and } w(S,T) \geq \frac{1}{2}w(S,\overline{S})\}$. It is easy to see that $N_{1/2}(S)$ and $\tilde{N}_{1/2}(S)$ differ by at most 1. Define

$$\phi^V(S) := \frac{N_{1/2}(S)}{|S|} \quad \text{and} \quad \phi^V(G) := \min_{S:|S|\leq|V|/2} \phi^V(S)$$

as the robust vertex expansion[2] of G. Also define

$$\Psi(S) := \phi(S) \cdot \phi^V(S) \quad \text{and} \quad \Psi(G) := \min_{S:|S|\leq|V|/2} \Psi(S)$$

as the minimum product of the edge expansion and the robust vertex expansion. The following is a generalization of Cheeger's inequality using robust vertex expansion.

THEOREM 1.1.

$$\lambda_2 = \Omega(\min\{\Psi(G), \phi(G)\}).$$

COROLLARY 1.2.

$$\lambda_2 = \Omega(\phi^V(G) \cdot \phi(G)).$$

---

[2]Note that the usual definition of vertex expansion, defined as $\min_{S:|S|\leq|V|/2} N(S)/|S|$, is too sensitive to edges of tiny weights (e.g., adding a complete graph with tiny edge weight will change $\phi^V(G)$ to one). One could replace the constant $1/2$ in the definition of $N_{1/2}(S)$ by another constant, say, 0.99, so that the definition of robust vertex expansion is closer to the definition of (ordinary) vertex expansion while we can still obtain similar results.

Note that $\phi^V(S) \geq \frac{1}{2}\phi(S)$ and so Corollary 1.2 is a generalization of Cheeger's inequality. Observe that $\phi^V(S)$ could be much larger than $\phi(S)$ when the edges crossing S spread out. For example, randomly generated instances such as those in the planted partition model [Bop87, McS01] have $\phi^V(G) = \Omega(1)$, and thus Theorem 1.1 implies that the spectral partitioning algorithm is a constant factor approximation algorithm for those instances.[3] Another interesting example is the hypercube.[4]

1.1.2. Expansion profile. The δ-small-set expansion ($0 < \delta \leq 1/2$) of G and the k-way expansion ($k \geq 2$) of G are defined as

$$\phi_\delta(G) := \min_{S:|S|\leq\delta|V|} \phi(S) \quad \text{and} \quad \phi_k(G) := \min_{S_1,\ldots,S_k:\ S_i\cap S_j=\emptyset\ \forall i\neq j} \ \max_{1\leq i\leq k} \phi(S_i).$$

The curve $\phi_\delta(G)$ for $0 < \delta \leq 1/2$ is defined by Lovász and Kannan [LK99] and is called the expansion profile of G. Note that $\phi(G) = \phi_{1/2}(G) = \phi_2(G)$. The following is a generalization of Cheeger's inequality using k-way expansion.

THEOREM 1.3. *For all $k \geq 2$,*

$$\lambda_2 = \Omega\left(\frac{1}{k} \cdot \phi_k(G) \cdot \phi(G)\right).$$

COROLLARY 1.4. *For all $\delta \leq 1/2$,*

$$\lambda_2 = \Omega\left(\delta \cdot \phi_\delta(G) \cdot \phi(G)\right).$$

Both Theorem 1.1 and Theorem 1.3 are tight up to a constant factor. Both proofs of Theorems 1.1 and 1.3 show that the spectral partitioning algorithm achieves the performance guarantees, i.e., the algorithm would output a set S with $\Psi(S) = O(\lambda_2)$ and $\phi(S) = O(k\lambda_2/\phi_k(G))$, respectively. These imply that the spectral partitioning algorithm is an $O(1/\phi^V(G))$-approximation and an $O(k/\phi_k(G))$-approximation for edge expansion.

1.1.3. Local partitioning algorithms for small-set expansion. Our proof techniques allow us to use the same approach to analyze the local partitioning algorithm using personal pagerank vectors [ACL06]. Given a parameter $\alpha \in (0,1]$ and a vertex s, the personal pagerank vector $r_{s,\alpha} \in \mathbb{R}^n$ is the unique solution to the equation $r_{s,\alpha} = \alpha\chi_s + (1-\alpha)Wr_{s,\alpha}$, where χ_s is the indicator vector which has value 1 at s and 0 elsewhere, and $W = (I+A)/2$ is the transition matrix of the lazy random walks.

THEOREM 1.5. *For any (unknown target) set $S \subseteq V$, there is a polynomial time randomized algorithm to find a set S' with*
   1. *$\phi(S') = O(\phi(S)\log(|S|)/\phi^V(G))$ and $|S'| = O(|S|\log|S|)$,*
   2. *$\phi(S') = O(k\phi(S)\log(|S|)/\phi_k(G))$ and $|S'| = O(|S|\log|S|)$*

---

[3]For example, in a planted k-partition instance where there are k subsets of size n/k with probability p having an edge between two vertices in the same subset and probability q having an edge between two vertices in different subsets for $q \ll p$, the improved Cheeger's inequality only proves a $O(k)$-approximation while Theorem 1.1 proves a $O(1)$-approximation.

[4]For hypercubes, it is known that the edge expansion is $\Omega(1/\log n)$, the vertex expansion is $\Omega(1/\sqrt{\log n})$ [Har66], and the product of the edge expansion and the vertex expansion is $\Omega(1/\log n)$ [Mar74]. We believe that the same bounds hold for robust vertex expansion, $\phi^V(G) = \Omega(1/\sqrt{\log n})$ and $\Psi(G) = \Omega(1/\log n)$, but we don't know of a proof yet. If that's true, Corollary 1.2 will give a bound of $\Omega(1/\log^{3/2}(n))$ on the second eigenvalue, and Theorem 1.1 will give the correct bound of $\Omega(1/\log(n))$, while Cheeger's inequality only gives a bound of $\Omega(1/\log^2 n)$.

by computing $r_{s,\alpha}$ for a random vertex $s \in S$ with $\alpha = O(\phi(S))$ and returning a level set of $r_{s,\alpha}$. For unweighted d-regular graphs, there is a local implementation with running time $O(d|S| \log(|S|)/\phi(S) + |S| \log^2 |S|)$.

Theorem 1.5 implies that the personal pagerank algorithm is an $O(\log(|S|)/\phi^V(G))$-approximation and an $O(k \log(|S|)/\phi_k(G))$-approximation for the small-set expansion problem where the output set size is bounded within a logarithmic factor of the target set size.

We also present a spectral approach to prove that the local graph partitioning algorithm using truncated random walks has similar performance guarantees as the spectral partitioning algorithm. Let $p_{s,t} := W^t \chi_s$ be the probability distribution vector after t steps of lazy random walks starting from the vertex s.

THEOREM 1.6. *For any (unknown target) set $S \subseteq V$, there is a polynomial time randomized algorithm to find a set S' with*
   1. *$\phi(S') = O(k\phi(S)/(\epsilon\phi_k(G)))$ and $|S'| = O(|S|^{1+\epsilon})$,*
   2. *$\phi(S') = O(k\phi(S) \log(|S|)/\phi_k(G))$ and $|S'| = O(|S|)$,*
by computing $p_{s,t}$ for a random vertex $s \in S$ with $t = O(\log(|S|)/\phi(S))$ for 1 and $t = O(1/\phi(S))$ for 2 and returning a level set of $p_{s,t}$. For unweighted d-regular graphs, there is a local implementation with running time $O(d\epsilon^2|S|^{1+\epsilon} \log^2(|S|)/\phi(S)^3)$.

Theorem 1.6 implies that the truncated random walks algorithm is an $O(k/\phi_k(G))$-approximation or an $O(k \log(|S|)/\phi_k(G))$-approximation for the small-set expansion problem, with different trade-offs of the output set size.

Our results provide improved analyses of local graph partitioning algorithms when the vertex expansion or the k-way expansion is large and provide theoretical justification of their good empirical performances in applications such as image segmentation and data clustering (see [ZLM13] and the references therein). The results show that the performances of local graph partitioning algorithms almost match that of Theorems 1.1 and 1.3 (within at most an $O(\log(|S|))$-factor in the approximation guarantee), with the additional advantages that they apply to the small-set expansion problem (giving bicriteria approximations for $\phi_\delta(G)$) and also that their running time could be sublinear in the graph size (when d and $|S|$ are small enough).

1.2. Comparisons with related work.

1.2.1. Generalizations of Cheeger's inequality. There are several recent generalizations of Cheeger's inequality using higher eigenvalues of the Laplacian matrix. The first generalization by Arora, Barak, and Steurer [ABS10] relates higher eigenvalues to small-set expansions,

$$\phi_{O(k^{-1/100})}(G) = O\left(\sqrt{\lambda_k \log_k n}\right),$$

and they use it to design a subexponential time algorithm for approximating unique games. The second generalization by Louis et al. [LRTV12] and Lee, Oveis Gharan, and Trevisan [LOT12] relates higher eigenvalues to k-way expansion (a stronger requirement than small-set expansion),

$$\tag{1} \frac{1}{2}\lambda_k \le \phi_k(G) \le O\left(\sqrt{\lambda_{2k} \log k}\right),$$

and this justifies the use of higher eigenvalues in k-way graph partitioning. Then there is a generalization by Kwok et al. [KLLOT13] relating higher eigenvalues to the ordinary edge expansion,

$$\tag{2} \phi(G) \le O(k)\frac{\lambda_2}{\sqrt{\lambda_k}},$$

which shows that the spectral partitioning algorithm performs better in instances with λ_k large for a small k.

Instead of using higher eigenvalues to give better bounds on expansion parameters, our results use expansion parameters to give better bounds on the second eigenvalue. We remark that the techniques developed in [KLLOT13] could be used to prove Theorem 1.3 (see Appendix A), but our approach is quite different and could be used to prove Theorem 1.1 and to extend Theorem 1.3 to analyze personal pagerank vectors. We also note that our proof of Theorem 1.3 can be used to prove (2) using a graph powering trick as described in [KL14] (see Appendix A).

1.2.2. Local graph partitioning algorithms. Local graph partitioning algorithms are useful in finding a small nonexpanding set in a large graph, as their running times are only weakly dependent on the graph size and could be sublinear time. All known algorithms are based on some random walks related processes. The first local graph partitioning algorithm is a truncated random walk algorithm by Spielman and Teng [ST13], which returns a set S' with $\phi(S') = O\left(\sqrt{\phi(S)\log^3 n}\right)$ with work-to-volume ratio $O(\text{polylog}(n)/\phi^2(S))$. The second algorithm is a personal pagerank algorithm by Andersen, Chung, and Lang [ACL06], which returns a set S' with $\phi(S') = O(\sqrt{\phi(S)\log(|S|)})$ with work-to-volume ratio $O(\text{polylog}(n)/\phi(S))$. The evolving set process is used by Andersen and Peres [AP09] to further improve the work-to-volume ratio to $O(\text{polylog}(n)/\sqrt{\phi(S)})$ while having the same performance guarantee as in [ACL06]. Using a better analysis of the escaping probability of random walks, Oveis Gharan and Trevisan [OT12] (see also [KL12]) showed that the $\sqrt{\log(|S|)}$ factor in the performance ratio can be removed, thereby almost matching the guarantee of Cheeger's inequality. They combined this with the evolving set process to find a set S' with $\phi(S') = O(\sqrt{\phi(S)/\epsilon})$, $|S'| = O(|S|^{1+\epsilon})$ and work-to-volume ratio $O(|S|^{\epsilon}\text{polylog}(n)/\sqrt{\phi})$.

Our contribution is to show that the performance of some simple local graph partitioning algorithms (truncated random walks, personal pagerank) almost match that of the improved Cheeger's inequalities. These provide the first analyses showing that random walk based algorithms perform better when $\phi^V(G)$, $\phi_k(G)$ or $\lambda_k(G)$ is large, with similar performances to the spectral partitioning algorithm while having additional features. We note that Zhu, Lattanzi, and Mirrokni [ZLM13] gave a better analysis of the personal pagerank algorithm when the internal expansion of the target set is large; our results are related but incomparable.

1.2.3. Analysis of mixing time. The notion of expansion profile was introduced by Lovász and Kannan [LK99] in the study of mixing times of random walks. They proved that the mixing time is upper bounded by $\int_{1/n}^{1/2} \frac{dx}{x\Phi(x)^2}$, where $\Phi(x) = \min_{0 \le \delta \le x} \phi_\delta$, which is a better bound on the mixing time when the average conductance is large (e.g., small sets expand well in geometric graphs).

Our work is inspired by their paper and some subsequent work [KLM06, MP05], both in the proof techniques (discussed in the next subsection) and in the definitions. The robust vertex expansion and its expansion profile are studied in [KLM06], where better bounds on the mixing time are proved in a similar form to the average conductance bound above. In particular, it implies the mixing time is bounded by $O(\log(n)/\Psi(G))$, and thus $\lambda_2 \ge \Omega(\Psi(G)/\log(n))$. We note that Morris and Peres also proved a lower bound on the second eigenvalue (Theorem 15 in [MP05]) using a parameter related to vertex expansion, but their definition is incomparable to ours.

Our contribution is to directly bound the second eigenvalue (not the mixing time) using the expansion parameters and our bounds are independent of n. Also, the bound that we prove using ϕ_k is considerably stronger. Using ϕ_δ, the average conductance bound only gives $1/\lambda_2 \leq O(\log(\delta n)/\phi_\delta^2 + \log(1/\delta)/\phi^2)$, not improving on Cheeger's inequality even when $\phi_\delta = \Omega(1)$ for constant δ, while Corollary 1.4 gives an $O(1)$-approximation when $\phi_\delta = \Omega(1)$ for constant δ.

1.3. Technical overview. The proofs are inspired by the work of Lovász and Kannan [LK99]. We observe that their method is useful in analyzing the solution to a Laplacian system ($Lx = b$), and can be extended to study both the second eigenvectors ($Lx = \lambda x$) and the personal pagerank vectors.

The high-level approach is to look at the solution vector $x \in \mathbb{R}^n$ with $x_1 \geq x_2 \geq \cdots \geq x_n$ and relate the (slow) decrease of x_i to the (large) expansion of the level sets in this vector. Similar to [LK99], we define a jumping sequence of indices $1 = m_0, m_1, m_2, \ldots$ such that $x_{m_i} - x_{m_{i+1}}$ is inversely proportional to the expansion of the level set $[1, m_i]$ (see Lemma 2.1). Using the Laplacian equation of the second eigenvector, we use an inductive argument to show that if the expansion of all level sets is $\Omega(\sqrt{\lambda_2})$, then the values of x_i decrease slowly enough such that $x_{n/2} > 0$ (see Lemma 2.2), contradicting that x is orthogonal to the all-one vector. We remark that this gives a new and quite different proof of Cheeger's inequality (e.g., without using the Cauchy–Schwarz inequality). To prove Theorem 1.1, we use the robust vertex expansion to argue that each jump can be made longer ($m_{i+1} - m_i$ made larger) and this gives the improved bound. To prove Theorem 1.3, we argue that given an ordering of the vertices, if ϕ_k is large, then there are only a small number of indices in the jumping sequence whose corresponding level sets $[1, m_i]$ are of small expansion (see Lemma 2.3), and then we modify the induction hypothesis to obtain the result (see Lemma 2.4). The inductive arguments and the use of ϕ_k in arguing about expansions of level sets are the new elements in the proofs that improve upon the average conductance bound of Lovász and Kannan.

The previous analyses of both the truncated random walk algorithm [ST13] and the personal pagerank algorithm [ACL06] are based on the combinatorial technique introduced by Lovász and Simonovits [LS90] in analyzing the mixing time of random walks. This technique is quite different from the analysis of spectral partitioning algorithms. It requires one to consider the random walk vectors for many different time steps, and it is difficult to incorporate the notions of ϕ_k or λ_k in the analyses as the ordering and level sets are changing in each time step.[5]

Our techniques provide two approaches to lift the analysis of the spectral graph partitioning algorithm for local graph partitioning algorithms, bringing closer the analyses of these two types of algorithms. For the personal pagerank algorithm, we use the Lovász–Kannan approach to directly analyze the vector so that we can use ϕ_k to reason about the level sets (Lemma 2.3). We note that this approach is considered by Andersen and Chung to give a simplified proof of the personal pagerank algorithm [AC07], and we will reuse some of their lemmas to obtain Theorem 1.5. For the truncated random walk algorithm, we use the spectral approach of Arora, Barak, and Steurer [ABS10] to directly obtain a vector with small Rayleigh quotient and small support, so that the improved Cheeger's inequalities can be applied to obtain results for approximating small-set expansions.[6]

---

[5]We still don't know how to do a better analysis for the evolving set process because of this difficulty.

[6]We thank David Steurer for suggesting this spectral approach.

Finally, we remark that this approach can be applied to analyze the solutions to other Laplacian systems. Consider the following algorithm for approximating edge expansion. For an unknown target set S, pick a random vertex $s \in S$, inject n units of current to s and extract one unit of current from every vertex in the graph, sort the vertices by the voltages,[7] and output the level set with the smallest expansion among all level sets of size up to $n/2$. Our approach implies that this algorithm always outputs a set S' with $\phi(S') = O(\sqrt{\phi(S)\log n})$. We believe that this approach draws more connections to the mixing time literature and will find further applications.

2. Spectral partitioning. Let $\lambda := \lambda_2$ and x be a second eigenvector such that $Lx = \lambda x$. We assume the vertices of the graph are labeled by 1 to n and sort the vertices so that $x_1 \geq x_2 \geq \cdots \geq x_n$. We use $[a, b]$ to denote the subset of vertices $\{a, a+1, \ldots, b\}$.

2.1. Vertex expansion. The proof of Theorem 1.1 consists of two steps. The first step is to prove the drop lemma and then define a jumping sequence to apply the lemma. The second step is to use an inductive argument to derive a contradiction if the expansion of all level sets are large.

2.1.1. Drop lemma and jumping sequence. The following lemma bounds the decrease of the values in x to the expansion of the level sets of x. Recall that $w(S, T)$ denotes the total weight of the edges with one vertex in S and another vertex in T.

LEMMA 2.1 (drop lemma). *For $1 \leq a < b \leq n$, we have*

$$x_a - x_b \leq \frac{\lambda \sum_{i=1}^{a} x_i}{w([1, a], [b, n])}.$$

Proof. For each i,

$$x_i - \sum_j w_{ij} x_j = \lambda x_i.$$

Summing this equation for $1 \leq i \leq a$, we have

$$\sum_{i=1}^{a} \sum_j x_i w_{ij} - \sum_{i=1}^{a} \sum_j x_j w_{ij} = \lambda \sum_{i=1}^{a} x_i.$$

Since $w_{ij} = w_{ji}$, this can be simplified to

$$\sum_{i \leq a} \sum_{j > a} x_i w_{ij} - \sum_{i \leq a} \sum_{j > a} x_j w_{ij} = \lambda \sum_{i=1}^{a} x_i.$$

Consider the edges from the set $[1, a]$ to the set $[b, n]$. Each edge contributes $(x_i - x_j)w_{ij}$ to the left-hand side, which is at least $(x_a - x_b)w_{ij}$. Therefore, we have

$$w([1, a], [b, n]) \cdot (x_a - x_b) \leq \lambda \sum_{i=1}^{a} x_i \quad \text{and thus} \quad x_a - x_b \leq \frac{\lambda \sum_{i=1}^{a} x_i}{w([1, a], [b, n])}. \qquad \square$$

We define a jumping sequence of indices to apply the drop lemma. Let $m_0 = 1$ and

$$m_{i+1} = \lceil m_i(1 + \phi^V(m_i)) \rceil,$$

---

[7]Or equivalently, sort the vertices based on the expected hitting time to s.

where $\phi^V(m_i), \phi(m_i), \Psi(m_i)$ are shorthand for $\phi^V([1, m_i]), \phi([1, m_i]), \Psi([1, m_i])$, respectively. Then by the definition of $\phi^V(m_i)$, we have

$$w([1, m_i], [m_{i+1}, n]) \geq \frac{1}{2} m_i \cdot \phi(m_i).$$

Putting it in the above inequality with $a = m_i$ and $b = m_{i+1}$, it follows that

$$(3) \qquad x_{m_i} - x_{m_{i+1}} \leq \frac{2\lambda \sum_{i=1}^{m_i} x_i}{m_i \cdot \phi(m_i)} = \frac{2\lambda \overline{x}_{m_i}}{\phi(m_i)}, \quad \text{where} \quad \overline{x}_l := \frac{1}{l} \sum_{i=1}^{l} x_i.$$

Note that $\overline{x}_l$ is nonincreasing over l.

2.1.2. Induction. We will prove the following lemma by induction.

LEMMA 2.2. *If $\Psi(m_i) \geq 32\lambda$ and $\phi(m_i) \geq 32\lambda$ for all $m_i \leq n/2$, then $\overline{x}_{m_{i+1}} \leq 2x_{m_{i+1}}$ for all $m_i \leq n/2$.*

First we see how it implies Theorem 1.1. Let m_j be the first term in the jumping sequence such that $m_j > n/2$. Note that the assumptions of Lemma 2.2 would imply that $x_{m_j} \geq \frac{1}{2}\overline{x}_{m_j} > 0$, where the last inequality follows because $\sum_{i=1}^{n} x_i = 0$ (as the second eigenvector is orthogonal to the all-one vector) and so all partial sums are positive. But this implies that $x_k > 0$ for all $1 \leq k \leq n/2$, and applying the same argument to $-x$ will give us a contradiction. Therefore, the assumptions of Lemma 2.2 must not hold, and thus there is an $m_i \leq n/2$ with $\Psi(m_i) \leq 32\lambda$ or $\phi(m_i) \leq 32\lambda$, proving Theorem 1.1.

Now we proceed to prove Lemma 2.2. It is clear that the inequality holds for m_0. Assuming that $\overline{x}_{m_i} \leq cx_{m_i}$ where $c = 2$,[8] we would like to prove that $\overline{x}_{m_{i+1}} \leq cx_{m_{i+1}}$. Note that

$$\sum_{i=1}^{m_{i+1}} x_i = \sum_{i=1}^{m_i} x_i + \sum_{i=m_i+1}^{m_{i+1}} x_i \leq m_i \overline{x}_{m_i} + (m_{i+1} - m_i)x_{m_i} \leq x_{m_i}(m_{i+1} + (c-1)m_i).$$

Dividing both sides of this inequality by m_{i+1}, we have

$$\overline{x}_{m_{i+1}} \leq x_{m_i}\left(1 + (c-1)\frac{m_i}{m_{i+1}}\right) \leq x_{m_i}\left(1 + \frac{(c-1)}{1 + \phi^V(m_i)}\right)$$

$$\leq x_{m_{i+1}}\left(\frac{c + \phi^V(m_i)}{1 + \phi^V(m_i)}\right)\left(\frac{\phi(m_i)}{\phi(m_i) - 2\lambda c}\right) \leq cx_{m_{i+1}},$$

where the second inequality follows from the definition of m_{i+1}, the third inequality is by (3) and $\overline{x}_{m_i} \leq cx_{m_i}$, and the last inequality follows from the following claim by plugging in $\phi^V(m_i)$ for h and $\phi(m_i)$ for φ. Note that the conditions of Claim 1 follow from the assumptions of Lemma 2.2, and this completes the proof.

Claim 1. If $2 \leq c \leq 4$, $32\lambda \leq h\varphi$ and $32\lambda \leq \varphi$, then we have

$$\left(\frac{c + h}{1 + h}\right)\left(\frac{\varphi}{\varphi - 2\lambda c}\right) \leq c.$$

Proof. The conclusion to check is

$$c \geq \left(\frac{c + h}{1 + h}\right)\left(\frac{\varphi}{\varphi - 2\lambda c}\right) = \left(c - \frac{(c-1)h}{1 + h}\right)\left(1 + \frac{2\lambda c}{\varphi - 2\lambda c}\right),$$

---

[8]The variable c is used so that we can reuse the calculation here for the proof of Theorem 1.3.

which is equivalent to

$$0 \geq \frac{-(c-1)h}{1+h} + \frac{2\lambda c^2}{\varphi - 2\lambda c} - \frac{(c-1)h(2\lambda c)}{(1+h)(\varphi - 2\lambda c)}.$$

Since $\varphi \geq 32\lambda > 2c\lambda$, this is equivalent to

$$0 \geq -(c-1)h(\varphi - 2\lambda c) + 2\lambda c^2(1+h) - (c-1)h(2\lambda c),$$

which can be simplified to

$$\frac{c-1}{c^2} \geq \frac{2\lambda(1+h)}{h\varphi} = 2\lambda \left(\frac{1}{h\varphi} + \frac{1}{\varphi} \right).$$

Since $2 \leq c \leq 4$, the left-hand side is at least $1/8$. We consider two cases. The first case is when $1/(h\varphi) \geq 1/\varphi$, and so the right-hand side is at most $4\lambda/(h\varphi)$. We have $1/8 \geq 4\lambda/(h\varphi)$, as long as $h\varphi \geq 32\lambda$, which is satisfied by our assumption. The second case is when $1/(h\varphi) \leq 1/\varphi$, and so the right-hand side is at most $4\lambda/\varphi$. We have $1/8 \geq 4\lambda/\varphi$, as long as $\varphi \geq 32\lambda$, which is also satisfied by our assumption.   □

2.2. Proof of Theorem 1.3. We follow the same approach to prove Theorem 1.3. The additional arguments are in Lemma 2.3 to bound the number of terms in the jumping sequence with small expansion using ϕ_k and in Claim 2 to control the inductive bound dynamically.

For Theorem 1.3, we define the jumping sequence as follows. Let $m_0 = 1$ and

$$m_{i+1} = \left\lceil m_i \left(1 + \frac{1}{2}\phi(m_i) \right) \right\rceil.$$

Then we have

$$w([1, m_i], [m_{i+1}, n]) \geq m_i\phi(m_i) - (m_{i+1} - m_i - 1) \geq \frac{1}{2}m_i \cdot \phi(m_i),$$

so that (3) still holds after applying the drop lemma.

2.2.1. k-way expansion. The assumption on ϕ_k allows us to bound the number of terms in the jumping sequence with small expansion. We note that the following lemma can be applied to any ordering of vertices (not just for the second eigenvector), and it will be applied to personal pagerank vectors later.

LEMMA 2.3. *For any $\theta < \phi_k/4$, there are at most $16k/\phi_k$ terms m_i in the jumping sequence with $\theta \leq \phi(m_i) \leq 2\theta$.*

Proof. Suppose by contradiction that there are at least $16k/\phi_k$ such terms. Let y_0 be the first such term and let y_i be the $(16i/\phi_k)$th such term. We claim that the sets $\{[1, y_0], [y_0 + 1, y_1], \ldots, [y_{k-1} + 1, y_k]\}$ are all of expansion less than ϕ_k, contradicting the definition of ϕ_k. Note that

$$y_{i+1} \geq y_i \left(1 + \frac{\theta}{2} \right)^{\frac{16}{\phi_k}} \geq y_i \left(1 + \frac{8\theta}{\phi_k} \right)$$

by our choice of m_{i+1}, and thus

$$y_{i+1} - y_i \geq \frac{8\theta y_i}{\phi_k}.$$

The expansion of the set $[y_i + 1, y_{i+1}]$ is

$$
\begin{aligned}
\phi([y_i + 1, y_{i+1}]) &= \frac{w([y_i + 1, y_{i+1}], [1, y_i] \cup [y_{i+1} + 1, n])}{y_{i+1} - y_i} \\
&\leq \frac{w([1, y_i], \overline{[1, y_i]}) + w([1, y_{i+1}], \overline{[1, y_{i+1}]})}{y_{i+1} - y_i} \\
&\leq \frac{2\theta y_{i+1} + 2\theta y_i}{y_{i+1} - y_i} \\
&= 2\theta \left(1 + \frac{2y_i}{y_{i+1} - y_i} \right).
\end{aligned}
$$

Using the lower bound on $y_{i+1} - y_i$, we have

$$
\phi([y_i, y_{i+1}]) \leq 2\theta + \frac{\phi_k}{2} < \phi_k,
$$

where the last inequality is by our assumption that $\theta < \phi_k/4$. $\square$

2.2.2. Induction. In the following, we assume that $\phi_k^2 \geq 1024\lambda$, as otherwise Theorem 1.3 holds trivially since $\phi_k \geq \phi$. We will prove the following lemma by induction.

LEMMA 2.4. *If $\phi_k^2 \geq 1024\lambda$ and $\phi(m_i) \geq 256k\lambda/\phi_k$ for all $m_i \leq n/2$, then $\overline{x}_{m_{i+1}} \leq 4x_{m_{i+1}}$ for all $m_i \leq n/2$.*

As argued before in applying Lemma 2.2, the assumptions of Lemma 2.4 would imply that $x_i > 0$ for all $1 \leq i \leq n/2$, leading to a contradiction. So, the assumptions of Lemma 2.4 must not hold, and thus there is an m_i with $\phi(m_i) \leq 256k\lambda/\phi_k$, proving Theorem 1.3.

To prove Lemma 2.4, we will prove by induction that $\overline{x}_{m_i} \leq c_i x_{m_i}$ where initially $c_0 = 2$ and

$$
c_{i+1} = \left\{
\begin{array}{ll}
c_i & \text{if } \phi(m_i) \geq \phi_k/4, \\
c_i/(1 - \epsilon_i c_i) & \text{if } \phi(m_i) < \phi_k/4, \text{where } \epsilon_i = 2\lambda/\phi(m_i).
\end{array}
\right.
$$

We first assume this induction step and show that $c_\infty \leq 4$ using Lemma 2.3. Then we will verify the induction step.

Claim 2. $c_\infty \leq 4$.

Proof. First, we prove by induction that

$$
c_i = \frac{c_0}{1 - \sum_{j=0}^{i-1} \epsilon_j c_0},
$$

where $\epsilon_j = 0$ if $\phi(m_j) \geq \phi_k/4$. Assume this is true for i. Then

$$
c_{i+1} = \frac{c_i}{1 - \epsilon_i c_i} = \left(\frac{c_0}{1 - \sum_{j=0}^{i-1} \epsilon_j c_0} \right) \left(\frac{1}{1 - \epsilon_i \left(\frac{c_0}{1 - \sum_{j=0}^{i-1} \epsilon_j c_0} \right)} \right) = \frac{c_0}{1 - \sum_{j=0}^{i} \epsilon_j c_0}.
$$

Next, we bound c_∞ using Lemma 2.3. Recall that $\epsilon_i = 2\lambda/\phi(m_i)$ and we can assume that $\phi(m_i) \geq 256k\lambda/\phi_k$. Let $\theta_0 = 256k\lambda/\phi_k$ and $\theta_{i+1} = 2\theta_i$. By Lemma 2.3, there

are at most $16k/\phi_k$ terms m_i in the jumping sequence with $\theta \leq \phi(m_i) < 2\theta$ when $\theta < \phi_k/4$. Therefore,

$$\sum \epsilon_j = \sum_{i \geq 0} \sum_{j:\theta_i \leq \phi(m_j) \leq 2\theta_i} \epsilon_j \leq \sum_{i \geq 0} \sum_{j:\theta_i \leq \phi(m_j) \leq 2\theta_i} \frac{2\lambda}{\theta_i}$$
$$\leq \sum_{i \geq 0} \frac{16k}{\phi_k} \frac{2\lambda}{\theta_i} = \sum_{i \geq 0} \frac{32k\lambda}{\phi_k} \frac{\phi_k}{256k\lambda 2^i} = \frac{1}{4}.$$

Therefore,

$$c_\infty = \frac{c_0}{1 - \sum_j \epsilon_j c_0} \leq \frac{c_0}{1 - \frac{c_0}{4}} = 4. \qquad \square$$

We prove the induction step. There are two cases, depending on whether $\phi(m_i) < \phi_k/4$. We first consider the case when $\phi(m_i) < \phi_k/4$. In this case, just apply (3) and we have

$$x_{m_{i+1}} \geq x_{m_i} - \frac{2\lambda c_i x_{m_i}}{\phi(m_i)} = x_{m_i}(1 - \epsilon_i c_i) \geq \overline{x}_{m_i}\left(\frac{1 - \epsilon_i c_i}{c_i}\right) \geq \frac{\overline{x}_{m_{i+1}}}{c_{i+1}}$$

by the definition of ϵ_i and c_{i+1} and we are done in this case.

It remains to consider the case when $\phi(m_i) \geq \phi_k/4$. By induction, we assume that $\overline{x}_{m_i} \leq c_i x_{m_i}$, and we claim that $\overline{x}_{m_{i+1}} \leq c_i x_{m_{i+1}}$. By the same calculation as in the induction for Theorem 1.1, we have

$$\sum_{i=1}^{m_{i+1}} x_i = \sum_{i=1}^{m_i} x_i + \sum_{i=m_i+1}^{m_{i+1}} x_i \leq x_{m_i}(m_{i+1} + (c_i - 1)m_i).$$

Similarly, dividing both sides of this inequality by m_{i+1}, we have

$$\overline{x}_{m_{i+1}} \leq x_{m_{i+1}}\left(\frac{c_i + \frac{1}{2}\phi(m_i)}{1 + \frac{1}{2}\phi(m_i)}\right)\left(\frac{\phi(m_i)}{\phi(m_i) - 2\lambda c_i}\right) \leq c_i x_{m_{i+1}},$$

where the first inequality is by $m_i/m_{i+1} \leq 1/(1 + \phi(m_i)/2)$ and (3), and the last inequality follows from Claim 1 by plugging in $h = \phi(m_i)/2$, $\varphi = \phi(m_i)$, $c = c_i$ and checking that the conditions $2 \leq c \leq 4$ (Claim 2), $h\varphi \geq \varphi^2/2 \geq \phi_k^2/32 \geq 32\lambda$, and $\varphi \geq 32\lambda$ are satisfied by our assumptions. This completes the induction step and thus the proof of Lemma 2.4.

3. Personal pagerank. We show that a similar and simpler analysis applies to the personal pagerank vector and prove Theorem 1.5. In subsections 3.1 and 3.2, we will show some basic properties of the pagerank vector. After that, we will prove the first part of Theorem 1.5 in subsection 3.3 and the second part in subsection 3.4. Finally in subsection 3.5 we will show how to compute the required set in time almost linear to the size of the target set.

Given a parameter $\alpha \in (0,1]$ and a vertex s, the personal pagerank vector $r_{s,\alpha} \in \mathbb{R}^n$ is the unique solution to the equation $r_{s,\alpha} = \alpha\chi_s + (1-\alpha)Wr_{s,\alpha}$, where W is the transition matrix of the lazy random walks. Note that $r_{s,\alpha}$ is a probability distribution vector. In the following, we assume S is an unknown target set with $3|S|H(|S|) \leq n$, where $H(n) = \sum_{i=1}^n 1/i$ is the harmonic series.

3.1. Drop lemma. Let $x := r_{s,\alpha}$ be the personal pagerank vector and assume $x_1 \geq x_2 \geq \cdots \geq x_n$. Andersen and Chung proved a drop lemma for pagerank vectors (see Lemma 1 of [AC07] and compare to our Lemma 2.1), for $1 \leq a < b \leq n$,

$$(4) \qquad x_a - x_b \leq \frac{\alpha}{w([1,a],[b,n])}.$$

3.2. Escaping probability. Let S be an unknown target set. Using a bound on the escaping probability of random walks [ST13],[9] Andersen and Chung proved that for half of the vertices s in S, the personal pagerank vector $x := r_{s,\alpha}$ will have the property that (see Lemma 5 of [AC07])

$$(5) \qquad \sum_{i \in S} x_i \geq 1 - \frac{\phi(S)}{\alpha}.$$

Setting $\alpha = 3\phi(S)$ makes sure that $\sum_{i \in S} x_i \geq 2/3$ and it follows that (see Lemma 2 of [AC07]) there exists an $a \leq |S|$ with

$$x_a \geq \frac{2}{3aH(|S|)},$$

since otherwise

$$\sum_{i=1}^{|S|} x_a < \frac{2}{3} \sum_{i=1}^{|S|} \frac{H(|S|)}{a} = \frac{2}{3},$$

which leads to a contradiction. In the following we assume a satisfies $x_a \geq 2/(3aH(|S|))$.

3.3. Vertex expansion. For vertex expansion, we start our jumping sequence by setting $m_0 = a$ and then define

$$m_{i+1} = \lceil m_i(1 + \phi^V(G)) \rceil$$

whenever $m_i < 3|S|H(|S|)$. We denote by m_∞ the last defined term. By this definition, we have $w([1,m_i],[m_{i+1},n]) \geq \frac{1}{2}m_i \cdot \phi(m_i)$, and it follows that

$$x_{m_{i+1}} \geq x_{m_i} - \frac{2\alpha}{m_i \cdot \phi(m_i)} \quad \text{and} \quad x_{m_\infty} \geq x_a - \sum_{i \geq 0} \frac{2\alpha}{m_i \cdot \phi(m_i)}.$$

Suppose by contradiction that $\phi(m_i) > 36\phi(S)H(|S|)/\phi^V(G)$ for all $m_i \leq 3|S|H(|S|)$, where S is the unknown target set. Then

$$\sum_{i \geq 0} \frac{2\alpha}{m_i \cdot \phi(m_i)} < \sum_{i \geq 0} \frac{2\alpha\phi^V(G)}{36a(1 + \phi^V(G))^i\phi(S)H(|S|)} \leq \frac{1}{3aH(|S|)},$$

where the last inequality uses the bound that $\sum_{i \geq 0} 1/(1 + \phi^V(G))^i \leq (1 + \phi^V(G))/\phi^V(G) \leq 2/\phi^V(G)$ and our choice that $\alpha = 3\phi(S)$. This implies that

$$x_{3|S|H(|S|)} \geq x_{m_\infty} \geq x_a - \sum_{i \geq 0} \frac{2\alpha}{m_i \cdot \phi(m_i)} > \frac{2}{3aH(|S|)} - \frac{1}{3aH(|S|)} = \frac{1}{3aH(|S|)},$$

---

[9]Actually, using a stronger result by Oveis Gharan and Trevisan [OT12], one can show that $x(S) \geq \frac{\phi(S)(1+\alpha)}{\alpha+\phi(S)(1-\alpha)}$, but it does not change the results in the following subsections.

and thus

$$\sum_{j \geq 0} x_j > \sum_{1 \leq j \leq 3|S|H(|S|)} \frac{1}{3aH(|S|)} \geq 1,$$

since $a \leq |S|$, contradicting that x is a probability distribution vector. Therefore, there must exist an $m_i \leq 3|S|H(|S|)$ with $\phi(m_i) \leq 36\phi(S)H(|S|)/\phi^V(G)$. Since $H(|S|) = \Theta(\log |S|)$, this proves the first part of Theorem 1.5.

3.4. k-way expansion. For k-way expansion, we define the jumping sequence by setting $m_0 = a$ and

$$m_{i+1} = \left\lceil m_i \left(1 + \frac{1}{2}\phi(m_i) \right) \right\rceil.$$

As before, we have $w([1, m_i], [m_{i+1}, n]) \geq \frac{1}{2}m_i \cdot \phi(m_i)$, and it follows that

$$x_{m_{i+1}} \geq x_{m_i} - \frac{2\alpha}{m_i \cdot \phi(m_i)}$$

and hence

(6) $$x_{m_\infty} \geq x_a - \sum_{i \geq 0} \frac{2\alpha}{m_i \cdot \phi(m_i)}.$$

We divide the summation into two parts:

(7) $$\sum_{i:\phi(m_i)<\phi_k/4} \frac{2\alpha}{m_i \cdot \phi(m_i)} + \sum_{i:\phi(m_i)\geq\phi_k/4} \frac{2\alpha}{m_i \cdot \phi(m_i)}.$$

The second part is at most

$$\sum_{i:\phi(m_i)\geq\phi_k/4} \frac{2\alpha}{m_i\phi_k/4} \leq \sum_{i \geq 0} \frac{8\alpha}{a\phi_k(1+\phi_k/2)^i} \leq \frac{64\alpha}{a\phi_k^2}.$$

The first summation in (7) can be bounded by Lemma 2.3 as follows. Suppose by contradiction that $\phi(m_i) > 1152k\phi(S)H(|S|)/\phi_k$ for all $m_i \leq 3|S|H(|S|)$. Let $\theta_0 = 1152k\phi(S)H(|S|)/\phi_k$ and $\theta_i = 2\theta_{i-1}$ for $i \geq 1$. By Lemma 2.3, there are at most $16k/\phi_k$ terms in the jumping sequence having conductance between θ and 2θ when $\theta < \phi_k/4$. Therefore, the first part in (7) is at most

$$\sum_{j \geq 0} \sum_{i:\theta_j \leq \phi(m_i) \leq 2\theta_j} \frac{2\alpha}{a\phi(m_i)} \leq \sum_{j \geq 0} \frac{32k}{\phi_k} \frac{\alpha}{a\theta_j} = \sum_{j \geq 0} \frac{32k}{\phi_k} \frac{\alpha}{a2^j\theta_0} = \frac{64k\alpha}{a\phi_k\theta_0}.$$

Putting these back into (6), we have

$$x_{m_\infty} \geq x_a - \frac{64\alpha}{a\phi_k^2} - \frac{64k\alpha}{a\phi_k\theta_0} \geq \frac{2}{3aH(|S|)} - \frac{192\phi(S)}{a\phi_k^2} - \frac{192\phi(S)k}{a\phi_k\theta_0},$$

where the second inequality is by the lower bound of x_a and the choice of $\alpha = 3\phi(S)$. Note that we can assume $\phi_k^2 \geq 1152\phi(S)H(|S|) = \theta_0\phi_k/k$, since otherwise Theorem 1.5 can be obtained by [OT12] (and [KL12]), which finds a set S' with $|S'| = O(|S|\log|S|)$ and $\phi(S') = O(\sqrt{\phi(S)\log|S|})$. With this assumption, we have

$$x_{m_\infty} \geq \frac{2}{3aH(|S|)} - \frac{384\phi(S)k}{a\phi_k\theta_0} > \frac{1}{3aH(|S|)},$$

where the last inequality is by our choice of θ_0. This implies that

$$x_{3|S|H(|S|)} > \frac{1}{3aH(|S|)} \quad \text{and thus} \quad \sum_{j \geq 1} x_j > \sum_{1 \leq j \leq 3|S|H(|S|)} \frac{1}{3aH(|S|)} = 1$$

since $a \leq |S|$, contradicting that x is a probability distribution vector. Therefore, there must exist an $m_i \leq 3|S|H(|S|)$ with $\phi(m_i) \leq 1152k\phi(S)H(|S|)/\phi_k$, proving the second part of Theorem 1.5 as $H(|S|) = \Theta(\log |S|)$.

3.5. Local algorithm. Andersen and Chung [AC07] show that the drop lemma (equation (4)) still holds even for approximate personal pagerank vectors, which can be computed efficiently in unweighted graphs. In the following, we assume the graphs are unweighted d-regular (in our setting, the edge weights are either $1/d$ or 0). An ϵ-approximate vector for $r_{s,\alpha}$ is a vector $r'_{s,\alpha}$ that satisfies $r'_{s,\alpha} = \alpha(\chi_s - q) + (1-\alpha)Wr'_{s,\alpha}$ where the vector q is nonnegative and satisfies $q(i) \leq \epsilon$ for every vertex i in the graph.

LEMMA 3.1 (see [AC07]). *There is an algorithm that computes an ϵ-approximate vector $r'_{s,\alpha}$. The running time of the algorithm is $O(d/(\epsilon\alpha))$. Assume $r'_{s,\alpha}(1) \geq r'_{s,\alpha}(2) \geq \cdots \geq r'_{s,\alpha}(n)$. The approximate vector $r'_{s,\alpha}$ satisfies for any $1 \leq a < b \leq n$,*

$$r'_{s,\alpha}(a) - r'_{s,\alpha}(b) \leq \frac{\alpha}{w([1,a],[b,n])}.$$

Note that $0 \leq q \leq \epsilon\vec{1}$ implies

$$r_{s,\alpha} - r'_{s,\alpha} = \alpha(I - (1-\alpha)W)^{-1}(\chi_s - (\chi_s - q)) = \alpha(I - (1-\alpha)W)^{-1}q$$
$$\leq \epsilon\alpha(I - (1-\alpha)W)^{-1}\vec{1} = \epsilon\vec{1},$$

where the last equality holds since $\vec{1}$ is an eigenvector of both I and W with eigenvalue 1. Hence for any vertex i, we have $r'_{s,\alpha}(i) \geq r_{s,\alpha}(i) - \epsilon$. We set $\epsilon = 1/(6|S|)$ and $\alpha = 3\phi(S)$ so that

$$\sum_{i \in S} r'_{s,\alpha}(i) \geq \sum_{i \in S} r_{s,\alpha}(i) - \epsilon|S| \geq 1 - \frac{\phi(S)}{\alpha} - \frac{1}{6} = \frac{1}{2}$$

for those vertices s that satisfy (5). Hence there exists an $a \leq |S|$ with

$$r'_{s,\alpha}(a) \geq \frac{1}{2aH(|S|)},$$

as otherwise

$$\sum_{i=1}^{|S|} r'_{s,\alpha}(a) < \sum_{i=1}^{|S|} \frac{1}{2aH(|S|)} = 1,$$

contradicting to $\sum_{i \in S} r'_{s,\alpha}(i) \geq 1/2$.

Since $r'_{s,\alpha}$ satisfies the drop lemma (equation (4)) and contains good initial value, both arguments in vertex expansion and k-way expansion follow (with the assumption $3|S|H(|S|) \leq n$ replaced by $6|S|H(|S|) \leq n$). The runtime of this algorithm is dominated by the runtime for computing the approximate vector $p'_{s,t}$ and sorting at most $O(|S|\log(|S|))$ vertices after, and hence the total complexity is $O(d/(\epsilon\alpha) + |S|\log^2(|S|)) = O(d|S|/\phi(S) + |S|\log^2(|S|))$.

4. Random walks. In this section, we present a spectral analysis of the random walk local graph partitioning algorithm [ST13, KL12]. The proof consists of three steps. The first step is to show that the Rayleigh quotient of the random walk vector $p_{s,t} = W^t \chi_s$ is small for $t \approx \log(|S|)/\phi(S)$, by using the analysis in the power method. The second step is to show that the $\|p_{s,t}\|_2$ is large for many vertices s in the unknown target set, by using the bound on escaping probability (or the staying probability). This allows us to apply the argument in [ABS10] to $p_{s,t}$ to obtain a vector with small Rayleigh quotient and small support. Then we apply the improved Cheeger's inequality to prove Theorem 1.6. Finally, we show that the truncated random walk vectors would also work, thereby proving a local implementation of the algorithm.

4.1. Rayleigh quotient. Recall that the Rayleigh quotient of a vector x is defined as $\mathcal{R}(x) = x^T L x / \|x\|^2$. The following lemma shows that the Rayleigh quotient of the vector $p_{s,t} := W^t \chi_s$ becomes smaller when t becomes larger. The proof follows the analysis of the power method in computing the largest eigenvector.

LEMMA 4.1. *For any starting vertex s,*

$$\mathcal{R}(p_{s,t}) \leq 2 - 2\|p_{s,t}\|_2^{1/t}.$$

Proof. Let $\chi_s = \sum_{i=1}^n c_i v_i$, where v_i are unit eigenvectors of L. Note that the lazy random walk matrix is $W = I - L/2$, and thus the vector $p_{s,t} = W^t \chi_s = \sum_{i=1}^n c_i (1 - \lambda_i/2)^t v_i$. Hence, the Rayleigh quotient of $p_{s,t}$ is

$$\mathcal{R}(p_{s,t}) = \frac{p_{s,t}^T L p_{s,t}}{\|p_{s,t}\|^2} = \frac{\sum_{i=1}^n c_i^2 (1 - \lambda_i/2)^{2t} \lambda_i}{\sum_{i=1}^n c_i^2 (1 - \lambda_i/2)^{2t}} = 2 - 2 \frac{\sum_{i=1}^n c_i^2 (1 - \lambda_i/2)^{2t+1}}{\sum_{i=1}^n c_i^2 (1 - \lambda_i/2)^{2t}}.$$

Note that $\sum_i c_i^2 = \|\chi_s\|_2^2 = 1$, and thus c_i^2 can be viewed as a probability distribution. Let X be the random variable having value $1 - \lambda_i/2$ with probability c_i^2. Then we can write $\mathcal{R}(p_{s,t}) = 2 - 2\mathbb{E}[X^{2t+1}]/\mathbb{E}[X^{2t}]$. By the power mean inequality and the nonnegativity of X (or Jensen's inequality on the function $x^{1+1/(2t)}$), we have

$$\mathbb{E}[X^{2t+1}]^{1/(2t+1)} \geq \mathbb{E}[X^{2t}]^{1/(2t)}.$$

Hence

$$\mathcal{R}(p_{s,t}) \leq 2 - 2\mathbb{E}[X^{2t}]^{1/2t} = 2 - 2\left(\sum_{i=1}^n c_i^2 (1 - \lambda_i/2)^{2t}\right)^{1/2t} = 2 - 2\|p_{s,t}\|_2^{1/t}. \qquad \square$$

4.2. Small support vector with small Rayleigh quotient. A vector x is called spectrally δ-sparse if $\|x\|_1^2 \leq \delta n \|x\|_2^2$, and the support of the vector x is $\text{supp}(x) := \{i \mid x_i \neq 0\}$. First, by using a result by Oveis Gharan and Trevisan on escaping probability (or staying probability), we bound the spectral sparsity of the random walk vector. Then, we use a result used by Arora, Barak, and Steurer to turn a spectrally sparse vector into a small support vector with similar Rayleigh quotient.

The following lemma by Oveis Gharan and Trevisan shows that if $\phi(S)$ is small, there is a large subset $U \subseteq S$, such that the random walk starting at any vertex $s \in U$ stays entirely inside S with good probability. In particular, the probability that the walk ends inside S is large.

THEOREM 4.2 (see [OT12]). *For any subset $S \subseteq V$, there is a subset $U \subseteq S$ such that $|U| \geq |S|/2$, and for any $s \in U$ we have*

$$\sum_{v \in S} p_{s,t}(v) \geq \frac{1}{200}\left(1 - \frac{3\phi(S)}{2}\right)^t.$$

The following lemma provides a bound on the spectral sparsity of $p_{s,t}$.

LEMMA 4.3. *For any subset $S \subseteq V$, there is a subset $U \subseteq S$ such that $|U| \geq |S|/2$, and for any $s \in U$ we have*

$$\|p_{s,t}\|_1^2 \leq \frac{40000|S|}{(1 - 3\phi(S)/2)^{2t}} \|p_{s,t}\|_2^2.$$

Proof. By Cauchy–Schwarz and Theorem 4.2, we have

$$\|p_{s,t}\|_2^2 \geq \sum_{v \in S} p_{s,t}(v)^2 \geq \frac{1}{|S|} \left(\sum_{v \in S} p_{s,t}(v) \right)^2$$

$$\geq \frac{1}{|S|} \left(\frac{1}{200} \left(1 - \frac{3\phi(S)}{2} \right)^t \right)^2 = \frac{1}{40000|S|} \left(1 - \frac{3\phi(S)}{2} \right)^{2t}.$$

Since $\|\chi_s\|_1 = 1$ and W preserves the 1-norm, we have $\|p_{s,t}\|_1^2 = \|W^t \chi_s\|^2 = 1$, and the result follows. $\qquad \square$

The following lemma in [ABS10] shows how to obtain a vector y with small support and similar Rayleigh quotient from a spectrally δ-sparse vector x. The proof is by choosing an appropriate threshold t and set $y = \max(x - t, 0)$.

LEMMA 4.4 (see [ABS10]). *Let $x \in \mathbb{R}_{\geq 0}^{|V|}$ be a nonnegative vector with $\|x\|_1^2 \leq \delta n \|x\|_2^2$. Then there exists a vector y with $|\operatorname{supp}(y)| = O(\delta n)$ and $\mathcal{R}(y) = O(\mathcal{R}(x))$.*

We will apply Lemma 4.4 on $p_{s,t}$ and obtain a vector with small Rayleigh quotient (Lemma 4.1) and small support (Lemma 4.3).

4.3. Improved Cheeger's guarantees. We are ready to prove Theorem 1.6. We will prove the approximation guarantee in this subsection and show how to compute the required set fast in the following subsection. In the following we assume $\phi(S) \leq 1/4$ and $|S| \geq 2$. We set $t = \epsilon \log |S|/(6\phi(S))$ so that

$$(8) \quad \left(1 - \frac{3\phi(S)}{2} \right)^{2t} \geq \exp(-3\phi(S))^{2t} = \exp(-6t\phi(S)) = \exp(-\epsilon \log |S|) = |S|^{-\epsilon}.$$

By Lemma 4.3, we have

$$\|p_{s,t}\|_1^2 \leq 40000|S|^{1+\epsilon} \|p_{s,t}\|_2^2.$$

On the other hand, since $|S| \geq 2$,

$$(40000|S|)^{-1/(2t)} \geq \exp\left(-\frac{17 \log |S|}{2t} \right) = \exp\left(-\frac{51\phi(S)}{\epsilon} \right) \geq 1 - \frac{51\phi(S)}{\epsilon}.$$

Therefore, by Lemma 4.1, we have

$$(9) \qquad \mathcal{R}(p_{s,t}) \leq 2 - 2(1 - \frac{3\phi(S)}{2})(1 - \frac{51\phi(S)}{2\epsilon}) = O\left(\frac{\phi(S)}{\epsilon} \right).$$

Now, we apply Lemma 4.4 by plugging the vector $p_{s,t}$ for x and obtain a vector y with $|\operatorname{supp}(y)| \leq O(|S|^{1+\epsilon})$ and $\mathcal{R}(y) = O(\phi(S)/\epsilon)$. Finally, by the proof of the improved Cheeger's inequality (2) (see Appendix A), for any nonnegative vector x, we can find a level set S' with $|S'| \leq |\operatorname{supp}(x)|$ and $\phi(S') = O(k\mathcal{R}(x)/\phi_k)$ or

$\phi(S') = O(k\mathcal{R}(x)/\sqrt{\lambda_k})$. Applying this to the vector y, we find $|S'| \leq |\operatorname{supp}(y)| = O(|S|^{1+\epsilon})$ and

$$\phi(S') = O\left(\frac{k\phi(S)}{\epsilon\phi_k}\right) \quad \text{or} \quad \phi(S') = O\left(\frac{k\phi(S)}{\epsilon\sqrt{\lambda_k}}\right).$$

Since a level set of y is a level set of $p_{s,t}$, this proves the approximation guarantee of Theorem 1.6.

4.4. Local algorithm. Computing the vector $p_{s,t} = W^t\chi_s$ exactly requires at least linear time. In the following, we assume the graph is an unweighted d-regular graph (in our setting, the edge weight is either $1/d$ or 0). To obtain a local algorithm, we can compute a good approximation to $p_{s,t}$ by repeatedly applying the operator W (initially we compute $W\chi_s$) and truncating the small values to zero.

LEMMA 4.5 (see [ST13, KL12]). *Let $p_{s,t} = W^t\chi_s$ be the exact random walk vector starting at vertex s. There is an algorithm that computes a vector $p'_{s,t}$ such that $p_{s,t} \geq p'_{s,t} \geq p_{s,t} - \alpha\vec{1}$ and $p'_{s,t} \geq 0$ in time $O(dt^2/\alpha)$.*

We set $t = \epsilon \log |S|/\phi(S)$ and $\alpha = \phi(S)/(160000|S|^{1+\epsilon})$, so that the time complexity of our local algorithm is $O(d\epsilon^2|S|^{1+\epsilon}\log^2 |S|/\phi(S)^3)$. It remains to show that $p'_{s,t}$ is still spectrally sparse and has small Rayleigh quotient.

LEMMA 4.6. *For $p_{s,t}$ that satisfies the conclusion in Lemma* 4.3, *we have*

$$\|p'_{s,t}\|_1^2 \leq \frac{1}{80000|S|^{1+\epsilon}}\|p'_{s,t}\|_2^2.$$

Proof. In the proof, we let $x := p_{s,t}$ and $y := p'_{s,t}$. By Lemma 4.5, we have $y(i)^2 \geq x(i)^2 - 2\alpha x(i)$ since $y(i) \geq \max(x(i) - \alpha, 0)$. Therefore,

$$\|y\|_2^2 = \sum_i y(i)^2 \geq \sum_i x(i)^2 - 2\alpha \sum_i x(i) = \|x\|_2^2 - 2\alpha.$$

By Lemma 4.3 and (8), we have $\|x\|_2^2 \geq 1/(40000|S|^{1+\epsilon})$. From our choice of α, we have $2\alpha = \phi(S)/(80000|S|^{1+\epsilon}) \leq \phi(S)\|x\|_2^2/2$. Therefore,

$$\|x\|_2^2 - 2\alpha \geq \|x\|_2^2\left(1 - \frac{\phi(S)}{2}\right) \geq \frac{1}{80000|S|^{1+\epsilon}} \geq \frac{1}{80000|S|^{1+\epsilon}}\|y\|_1^2,$$

where the last inequality holds as $\|y\|_1^2 \leq \|x\|_1^2 = 1$. $\qquad\square$

LEMMA 4.7.

$$\mathcal{R}(p'_{s,t}) \leq O\left(\frac{\phi(S)}{\epsilon}\right).$$

Proof. Again, we let $x := p_{s,t}$ and $y := p'_{s,t}$ in the proof. Let $r = x - y \geq 0$. Then we have

$$\mathcal{R}(y) = \frac{y^T L y}{y^T y} = \frac{(x-r)^T L(x-r)}{y^T y} = \frac{x^T L x + r^T L r - 2x^T L r}{y^T y} \leq \frac{2x^T L x + 2r^T L r}{y^T y}.$$

By the calculation in Lemma 4.6, $\|y\|_2^2 \geq (1 - \phi(S)/2)\|x\|_2^2$. Hence, using $r \geq 0$ and $y \geq 0$, we have $\|r\|_2^2 \leq \|x\|_2^2 - \|y\|_2^2 \leq \phi(S)\|x\|_2^2/2$ and $r^T L r \leq 2r^T r \leq \phi(S)\|x\|_2^2$. So, we have

$$\mathcal{R}(y) = O\left(\frac{x^T L x}{y^T y} + \frac{r^T L r}{y^T y}\right) = O\left(\frac{x^T L x}{x^T x} + \frac{r^T L r}{x^T x}\right) = O(\mathcal{R}(x) + \phi(S)) = O\left(\frac{\phi(S)}{\epsilon}\right),$$

where the last inequality is by (9). $\qquad\square$

With Lemmas 4.6 and 4.7, we can use the same proof in section 4.3 to prove Theorem 1.6 with the time complexity claimed.

To prove Theorem 1.6.2, we only need to set $\epsilon = 1/\log(|S|)$ so that $|S|^{1+\epsilon} = O(|S|)$.

4.5. Local eigenvalue. We remark that if we do not care about local implementations, we can find a particular good starting vertex u such that the random walk algorithm starting at u gives a better performance guarantee $\phi(S') = O(k\lambda_S/(\epsilon\phi_k))$, where λ_S is the smallest eigenvalue of the matrix L_S which is the restriction of L on the subset S. Chung [Chu07] shows the following local Cheeger's inequality:

$$\lambda_S \leq \min_{T \subseteq S} \phi(T) \leq \sqrt{2\lambda_S}.$$

Hence λ_S is at most $\phi(S)$ and could be much smaller, for instance, when a subset of S has very small expansion. The idea is similar to that in [KL12] and we just give a quick sketch. Let v_S be the corresponding eigenvector with eigenvalue λ_S. We choose our starting vertex to be $u = \text{argmax}_i |v_S(i)|$. Then we show that the spectral sparsity of the t-steps random walk is at most $|S|/(1-\lambda_S)^{2t} < |S|/(1-O(\phi(S)))^{2t}$. This allows us to set t to be larger so as to improve the Rayleigh quotient of the random walk vector.

Appendix A. Relations between improved Cheeger's inequality and Theorem 1.3. First, although it is unclear how to derive Theorem 1.3 from the statement of improved Cheeger's inequality, we can show that Theorem 1.3 can be derived from the proof in [KLLOT13], as pointed out to us by Luca Trevisan. Then, we show that the improved Cheeger's inequality can be derived from Theorem 1.3, using a graph powering trick as described in [KL14].

A.1. Improved Cheeger's inequality implies Theorem 1.3. The following stronger statement was shown in [KLLOT13].

THEOREM A.1 (Theorem 3.5 of [KLLOT13], restated). *For any nonnegative vector x with $|\text{supp}(x)| \leq n/2$, let $\phi_{sweep}(x)$ be the minimum expansion of the level sets of x. At least one of the following holds:*
  1. *$\phi_{sweep}(x) \leq O(k)\mathcal{R}(x)$.*
  2. *There exist k disjointly supported vectors $x_1, \ldots, x_k$ such that for all $1 \leq i \leq k$, $\text{supp}(x_i) \subseteq \text{supp}(x)$ and $\mathcal{R}(x_i) \leq O(k^2\mathcal{R}(x)^2/\phi_{sweep}(x)^2)$.*

We apply the theorem with $x = \max(v_2, 0)$ or $x = \max(-v_2, 0)$, whichever is of smaller support. Note that $\mathcal{R}(x) \leq \lambda_2$ by standard argument [HLW06]. When the first case of Theorem A.1 holds, it is clear that

$$\phi(G) \leq \phi_{\text{sweep}}(x) \leq O(k)\mathcal{R}(x) \leq O(k\lambda_2) \leq O\left(\frac{k\lambda_2}{\phi_k}\right).$$

Otherwise, there exist k disjointly supported vectors, each with Rayleigh quotient not larger than $O(k^2\lambda_2^2/\phi_{\text{sweep}}^2(x))$. Applying (the original) Cheeger's arguments on these vectors, we can find k disjoint subsets S_i, each of them satisfying $\phi(S_i) \leq O(k\lambda_2/\phi_{\text{sweep}}(x))$. This implies that

$$\phi_k \leq O\left(\frac{k\lambda_2}{\phi_{\text{sweep}}(x)}\right), \quad \text{or} \quad \lambda_2 = \Omega\left(\frac{\phi_k\phi_{\text{sweep}}(x)}{k}\right) = \Omega\left(\frac{\phi_k\phi(G)}{k}\right),$$

and Theorem 1.3 follows.

A.2. Theorem 1.3 implies improved Cheeger's inequality. In [KL14], the authors proved a lower bound on the expansion of graph powers and used it to show some reductions on Cheeger's inequalities. We show that the same approach can be used to prove improved Cheeger's inequality by Theorem 1.3.

THEOREM A.2 (Theorem 1 of [KL14], restated). *Let H denote the graph with adjacency matrix W^t where W is the lazy random walk matrix of G. Then we have*

$$\phi(H) \geq \frac{1}{20}\left(1 - \left(1 - \frac{\phi(G)}{2}\right)^{\sqrt{t}}\right).$$

The following corollary is a generalization of Corollary 12 of [KL14], which shows that general cases of improved Cheeger's inequality can be reduced to the cases where λ_k is constant.

COROLLARY A.3. *Suppose one could prove that $\phi(H) \leq C\lambda_2(H)$ for some $C \geq 1/10$ for all graphs H with $\lambda_k(H) \geq 1/4$; then it implies that $\phi(G) \leq 40C\lambda_2(G)/\sqrt{\lambda_k(G)}$ for any G and any $\lambda_k(G)$.*

Proof. We assume that $\phi(G) \leq \sqrt{\lambda_k}$, as otherwise, by Cheeger's inequality, $2\lambda_2(G) \geq \phi(G)^2 \geq \phi(G)\sqrt{\lambda_k}$ and the statement is true. Consider H with adjacency matrix $W^{1/\lambda_k(G)}$. Then

$$\lambda_k(H) = 1 - \left(1 - \frac{\lambda_k(G)}{2}\right)^{1/\lambda_k} \geq 1 - \exp\left(-\frac{1}{2}\right) \geq 1/4.$$

Therefore, if one could prove that $\phi(H) \leq C\lambda_2(H)$, then

$$C\lambda_2(H) \geq \phi(H) \geq \frac{1}{20}\left(1 - \left(1 - \frac{\phi(G)}{2}\right)^{\sqrt{1/\lambda_k(G)}}\right)$$

$$\geq \frac{1}{20}\left(1 - \exp\left(-\frac{\phi(G)}{2\sqrt{\lambda_k(G)}}\right)\right) \geq \frac{\phi(G)}{80\sqrt{\lambda_k(G)}},$$

where the second inequality is by Theorem A.2. On the other hand,

$$\lambda_2(H) = 1 - \left(1 - \frac{\lambda_2(G)}{2}\right)^{1/\lambda_k(G)} \leq \frac{\lambda_2(G)}{2\lambda_k(G)},$$

and the corollary follows by combining the two inequalities. □

Now we show the improved Cheeger's inequality in [KLLOT13] follows from Corollary A.3 and Theorem 1.3. By the easy side of the higher order Cheeger's inequality, we have $\phi_k \geq \lambda_k/2$. Hence, for any graph G with $\lambda_k \geq 1/4$, we have $\phi_k \geq 1/8$ and Theorem 1.3 gives $\phi(G) = O(k\lambda_2(G))$. Therefore, we can apply Corollary A.3 (with $C = O(k)$) and conclude that $\phi(G) = O(k\lambda_2(G)/\sqrt{\lambda_k(G)})$ is true for any graph G and any λ_k, and the improved Cheeger's inequality in [KLLOT13] follows.

REFERENCES

[Alo86]   N. Alon, *Eigenvalues and expanders*, Combinatorica, 6 (1986), pp. 83–96.

[AM85]   N. Alon and V. Milman, l_1, *isoperimetric inequalities for graphs, and superconcentrators*, J. Combin. Theory Ser. B, 38 (1985), pp. 73–88.

[AC07]   R. Andersen and F. R. K. Chung, *Detecting sharp drops in pagerank and a simplified local partitioning algorithm*, in Proceedings of the 4th International Conference on Theory and Applications of Models of Computation, 2007, pp. 1–12.

[ACL06]   R. Andersen, F. R. K. Chung, and K. J. Lang, *Local graph partitioning using PageRank vectors*, in Proceedings of the 47th Annual IEEE Symposium on Foundations of Computer Science, 2006, pp. 475–486.

[AP09]   R. Andersen and Y. Peres, *Finding sparse cuts locally using evolving sets*, in Proceedings of the 41st Annual ACM Symposium on Theory of Computing, 2009, pp. 235–244.

[ABS10]   S. Arora, B. Barak, and D. Steurer, *Subexponential algorithms for unique games and related problems*, in Proceedings of the 51st Annual IEEE Symposium on Foundations of Computer Science, 2010, pp. 563–572.

[Bop87]   R. Boppana, *Eigenvalues and graph bisection: An average-case analysis*, in Proceedings of the 28th Annual Symposium on Foundations of Computer Science, 1987, pp. 280–285.

[Che70]   J. Cheeger, *A lower bound for the smallest eigenvalue of the Laplacian*, in Problems in Analysis, Princeton University Press, Princeton, NJ, 1970, pp. 195–199.

[Chu07]   F. R. K. Chung, *Random walks and local cuts in graphs*, Linear Algebra Appl., 423 (2007), pp. 22–32.

[Har66]   K. H. Harper, *Optimal numberings and isoperimetric problems on graphs*, J. Combin. Theory 1, (1966), pp. 385–393.

[HLW06]   S. Hoory, N. Linial, and A. Wigderson, *Expander graphs and their applications*, Bull. Amer. Math. Soc., 43 (2006), pp. 439–561.

[KLM06]   R. Kannan, L. Lovász, and R. Montenegro, *Blocking conductance and mixing in random walks*, Combin. Probab. Comput., 15 (2006), pp. 541–570.

[KL12]   T. C. Kwok and L. C. Lau, *Finding small sparse cuts by random walk*, in Proceedings of the 16th International Workshop on Randomization and Computation, 2012, pp. 615–626.

[KL14]   T. C. Kwok and L. C. Lau, *Lower bounds on expansions of graph powers*, in Proceedings of the 17th Annual International Workshop on Approximation Algorithms for Combinatorial Optimization Problems, 2014, pp. 313–324.

[KLLOT13] T. C. Kwok, L. C. Lau, Y. T. Lee, S. Oveis Gharan, and L. Trevisan, *Improved Cheeger's inequality: Analysis of spectral partitioning algorithms through higher order spectral gap*, in Proceedings of the 45th Annual Symposium on Theory of Computing, 2013, pp. 11–20.

[LOT12]   J. R. Lee, S. Oveis Gharan, and L. Trevisan, *Multi-way spectral partitioning and higher-order Cheeger inequalities*, in Proceedings of the 44th Annual Symposium on Theory of Computing, 2012, pp. 1117–1130.

[LRTV12]   A. Louis, P. Raghavendra, P. Tetali, and S. Vempala, *Many sparse cuts via higher eigenvalues*, in Proceedings of the 44th Annual ACM Symposium on Theory of Computing, 2012, pp. 1131–1140.

[LK99]   L. Lovász and R. Kannan, *Faster mixing via average conductance*, in Proceedings of the 31st Annual ACM Symposium on Theory of Computing, 1999, pp. 282–287.

[LS90]   L. Lovász and M. Simonovits, *The mixing time of Markov chains, an isoperimetric inequality, and computing the volume*, in Proceedings of the 31st Annual IEEE Symposium on Foundations of Computer Science, 1990, pp. 346–354.

[Mar74]   G. A. Margulis, *Probabilistic characteristics of graphs with large connectivity*, Problemy Peredachi Informatsii, 10 (1974), pp. 101–108.

[McS01]   F. McSherry, *Spectral partitioning of random graphs*, in Proceedings of the 42nd IEEE Symposium on Foundations of Computer Science, 2001, pp. 529–537.

[MP05]   B. Morris and Y. Peres, *Evolving sets, mixing and heat kernel bounds*, Probab. Theory Related Fields, 133 (2005), pp. 245–266.

[OT12]   S. Oveis Gharan and L. Trevisan, *Approximating the expansion profile and almost optimal local graph clustering*, in Proceedings of the 53rd Annual IEEE Symposium on Foundations of Computer Science, 2012, pp. 187–196.

[ST13]   D. A. SPIELMAN AND S.-H. TENG, *A local clustering algorithm for massive graphs and its applications to nearly-linear time graph partitioning*, SIAM J. Comput., 42 (2013), pp. 1–26.

[ZLM13]   Z. A. ZHU, S. LATTANZI, AND V. MIRROKNI, *Local graph clustering beyond Cheeger's inequality*, in Proceedings of the 30th International Conference on Machine Learning, 2013, pp. 396–404.