Improved Cheeger’s Inequality and Analysis of Local Graph Partitioning using Vertex Expansion and Expansion Profile

Tsz Chiu Kwok\textsuperscript{*} \quad Lap Chi Lau\textsuperscript{†} \quad Yin Tat Lee\textsuperscript{‡}

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 $G$,

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

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 $k \geq 2$,

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

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 non-expanding 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.

\textsuperscript{*}École polytechnique fédérale de Lausanne, tckwok0@gmail.com.
\textsuperscript{†}Simons Institute and University of Waterloo, lapchi@uwaterloo.ca. This material is based upon work supported by the National Science Foundation under Grant No. 1216642.
\textsuperscript{‡}Massachusetts Institute of Technology, yintat@mit.edu.
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\footnote{By standard arguments, the results can be extended to handle non-regular graphs using the notion of conductance and the normalized Laplacian matrix.}. 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

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

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 of $L$ being $0 = \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n \leq 2$. Cheeger’s inequality \cite{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 \cite{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 \cite{KLLOT13} of Cheeger’s inequality bounds the edge expansion of $G$ using the second and the $k$-th eigenvalues of $L$ for any $k \geq 2$,

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

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 analyze 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 following Kannan, Lovász and Montene-
gro [KLM06]. For $S \subseteq V$, let $N_{1/2}(S) := \min\{|T| \mid T \subseteq V - S \text{ and } w(S, T) \geq \frac{1}{2}w(S, \overline{S})\}$. 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.**

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

**Corollary 1.**

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

Note that $\phi^V(S) \geq \frac{1}{2}\phi(S)$ and so Corollary 1 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 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 $\delta$-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.

---

\(^2\) Note that the usual definition of vertex expansion, define 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 other 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.

\(^3\) 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 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 will give a bound of $\Omega(1/\log n)$ on the second eigenvalue, and Theorem 1 will give the correct bound of $\Omega(1/\log n)$, while Cheeger’s inequality only gives a bound of $\Omega(1/\log^2 n)$. 

---

2
Theorem 2. For all \( k \geq 2 \),
\[
\lambda_2 = \Omega\left(\frac{1}{k} \cdot \phi_k(G) \cdot \phi(G)\right).
\]

Corollary 2. For all \( \delta \leq 1/2 \),
\[
\lambda_2 = \Omega\left(\delta \cdot \phi_\delta(G) \cdot \phi(G)\right).
\]

Both Theorem 1 and Theorem 2 are tight up to a constant factor. Both proofs of Theorem 1 and Theorem 2 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 a \( O(1/\phi(V(G))) \)-approximation and a \( O(k/\phi_k(G)) \)-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 \( r_{s,\alpha} = \alpha \chi_s + (1 - \alpha) W r_{s,\alpha} \), where \( W \) is the transition matrix of the lazy random walks.

Theorem 3. For any (unknown target) set \( S \subseteq V \), there is a polynomial time randomized algorithm to find a set \( S' \) with
\[
1. \ \phi(S') = O(\phi(S) \log(|S|)/\phi(V(G))) \text{ and } |S'| = O(|S| \log |S|),
2. \ \phi(S') = O(k\phi(S) \log(|S|)/\phi_k(G)) \text{ and } |S'| = O(|S| \log |S|),
\]
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 3 implies that the personal pagerank algorithm is a \( O(\log(|S|)/\phi^V(G)) \)-approximation and a \( O(k \log(|S|)/\phi_k(G)) \)-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 4. For any (unknown target) set \( S \subseteq V \), there is a polynomial time randomized algorithm to find a set \( S' \) with
\[
1. \ \phi(S') = O(k\phi(S)/(\epsilon \phi_k(G))) \text{ and } |S'| = O(|S|^{1+\epsilon}),
2. \ \phi(S') = O(k\phi(S) \log(|S|)/\phi_k(G)) \text{ and } |S'| = O(|S|),
\]
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 \( O(\varepsilon^2|S|^{1+\varepsilon} \log^2(|S|)/\phi(S)^3) \).

Theorem 4 implies that the truncated random walk algorithm is a \( O(k/\phi_k(G)) \)-approximation or a \( O(k \log(|S|)/\phi_k(G)) \)-approximation for the small-set expansion problem, with different tradeoffs 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 Theorem 1 and Theorem 2 (within at most a \( O(\log(|S|)) \)-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:

\[
\phi_{O(k^{1/100})}(G) = O(\sqrt{\lambda_k \log n}),
\]

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

\[
\frac{1}{2} \lambda_k \leq \phi_k(G) \leq O(\sqrt{\lambda_{2k} \log k}),
\]

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:

\[
\phi(G) \leq O(k) \frac{\lambda_2}{\sqrt{\lambda_k}}
\]

which shows that the spectral partitioning algorithm performs better in instances with \( \lambda_k \) 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 2 (see Section A in the Appendix), but our approach is quite different and could be used to prove Theorem 1 and to extend Theorem 2 to analyze personal pagerank vectors. We also note that our proof of Theorem 2 can be used to prove (1.2) using a graph powering trick as described in [KL14] (see Section A in the Appendix).
1.2.2 Local Graph Partitioning Algorithms

Local graph partitioning algorithms are useful in finding a small non-expanding set in a large graph, as their running times are only weakly dependent on the graph size and could be sub-linear 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 $\phi(S') = O(\sqrt{\phi(S) \log^3 n})$ 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 improved 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 $O(|S'| \text{polylog}(n)/\sqrt{\phi(S)})$.

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 et al. [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 $\int \frac{dx}{\Phi(x)}$ where $\Phi(x) = \min_{0 \leq \delta \leq x} \phi_\delta$, 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 (will be 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 $\lambda_2 \geq \Omega(\Psi(G)/\log(n))$. 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 $\phi_k$ is considerably stronger. Using $\phi_\delta$, the average conductance bound only gives $1/\lambda_2 \leq O(\log(\delta n)/\phi^2_\delta + \log(1/\delta)/\phi^2)$, not improving on Cheeger’s inequality even when $\phi_\delta = \Omega(1)$ for constant $\delta$, while Corollary 2 gives a $O(1)$-approximation when $\phi_\delta = \Omega(1)$ for constant $\delta$. 

5
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 \( x_1 \geq x_2 \geq \ldots \geq x_n \), and relates 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 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), 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, 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 2, we argue that given an ordering of the vertices, if \( \phi_k \) 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 3), and then we modify the induction hypothesis to obtain the result (see Lemma 4). The inductive arguments and the use of \( \phi_k \) 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 to consider the random walk vectors for many different time steps, and it is difficult to incorporate the notions of \( \phi_k \) or \( \lambda_k \) in the analyses as the ordering and level sets are changing in each time step.

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 \( \phi_k \) to reason about the level sets (Lemma 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 3. For the truncated random walk algorithm, we use the spectral approach of Arora-Barak-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.

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 extracts one unit of current from every vertex in the graph, sort the vertices by the voltages, and output the level set with the smallest expansion among all level sets of size up to \( n/2 \). Our approach implies that this

\[ \text{We still don’t know how to do a better analysis for the evolving set process because of this difficulty.} \]

\[ \text{We thank David Steurer for suggesting this spectral approach.} \]

\[ \text{Or equivalently, sort the vertices based on the expected hitting time to } s. \]
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 $Lx = \lambda x$. Sort the vertices so that $x_1 \geq x_2 \geq \ldots \geq x_n$.

2.1 Vertex Expansion

The proof of Theorem 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 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 $i$,

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

Sum 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

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

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

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

\qed
We define a jumping sequence of indices to apply the drop lemma. Let \( m_0 = 1 \) and
\[
m_{i+1} = \lceil m_i (1 + \phi^V(m_i)) \rceil,
\]
where \( \phi^V(m_i), \phi(m_i), \Psi(m_i) \) are short hands for \( \phi^V([1, m_i]), \phi([1, m_i]), \Psi([1, m_i]) \) respectively. Then, for \( m_i \leq n/2 \), 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
\[
x_{m_i} - x_{m_{i+1}} \leq \frac{2\lambda\sum_{i=1}^{m_i} x_i}{m_i \cdot \phi(m_i)} = \frac{2\lambda \overline{x}_{m_i}}{\phi(m_i)}, \text{ where } \overline{x}_l := \frac{1}{l} \sum_{i=1}^{l} x_i. \tag{2.1}
\]
Note that \( \overline{x}_l \) is non-increasing over \( l \).

### 2.1.2 Induction

We will prove the following lemma by induction.

**Lemma 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. Let \( m_j \) be the first term in the jumping sequence such that \( m_j > n/2 \). Note that the assumptions of Lemma 2 would imply that \( x_j \geq \frac{1}{\phi^V} > 0 \), 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 \( x_i > 0 \) for all \( 1 \leq i \leq n/2 \), and applying the same argument to \(-x\) will give us a contradiction. Therefore, the assumptions of Lemma 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.

Now we proceed to prove Lemma 2. It is clear that the inequality holds for \( m_0 \). Assume that \( \overline{x}_{m_i} \leq cx_{m_i} \) where \( c = 2 \), 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
\[
\overline{x}_{m_{i+1}} \leq \frac{x_{m_i}(1 + (c - 1)m_i/m_{i+1})}{m_{i+1}} \leq x_{m_i}(1 + \frac{(c - 1)}{1 + \phi^V(m_i)}) \leq x_{m_{i+1}}(\frac{c + \phi^V(m_i)}{1 + \phi^V(m_i)} \cdot \frac{\phi(m_i)}{\phi(m_i) - 2\lambda c}) \leq cx_{m_{i+1}},
\]
where the second inequality follows from the definition of \( m_{i+1} \), the third inequality is by (2.1), and the last inequality follows from the following claim by plugging in \( \phi^V(m_i) \) for \( h \) and \( \phi(m_i) \) for \( \varphi \). Note that the conditions of Claim 1 follows from the assumptions of Lemma 2, and this completes the proof.

---

8The variable \( c \) is used so that we can reuse the calculation here for the proof of Theorem 2.
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
\[
c \geq \left(\frac{c+h}{1+h}\right)\left(\frac{\varphi}{\varphi-2\lambda c}\right) = (c - \frac{(c-1)h}{1+h})(1 + \frac{2\lambda c}{\varphi-2\lambda c}),
\]
which is equivalent to
\[
0 > -\frac{(c-1)h}{1+h} + \frac{2\lambda c^2}{\varphi-2\lambda c} - \frac{(c-1)h(2\lambda c)}{(1+h)(\varphi-2\lambda c)}.
\]
Since $\varphi \geq 32\lambda > 2c\lambda$, this is equivalent to
\[
0 > -(c-1)h(\varphi-2\lambda c) + 2\lambda c^2(1+h) - (c-1)h(2\lambda c),
\]
which can be simplified to
\[
\frac{c-1}{c^2} > \frac{2\lambda(1+h)}{h\varphi} = 2\lambda \left(\frac{1}{h\varphi} + \frac{1}{\varphi}\right).
\]
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. \hfill \square

2.2 Proof of Theorem 2

We follow the same approach to prove Theorem 2. The additional arguments are in Lemma 3 to bound the number of terms in the jumping sequence with small expansion using $\phi_k$ and in Claim 2 to control the inductive bound dynamically.

For Theorem 2, we define the jumping sequence as follows. Let $m_0 = 1$ and
\[
m_{i+1} = \lceil m_i(1 + \frac{1}{2}\phi(m_i)) \rceil.
\]
Then, for $m_i \leq n/2$, 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 equation (2.1) still holds after applying the drop lemma.
2.2.1  

\textit{k-way Expansion}

The assumption on \( \phi_k \) 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 second eigenvector), and it will be applied to personal pagerank vectors later.

\begin{lemma}
For any \( \theta < \phi_k/4 \), there are at most \( 16k/\phi_k \) terms \( m_i \) in the jumping sequence with \( \theta \leq \phi(m_i) \leq 2\theta \).
\end{lemma}

\begin{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 \( \{[1, y_0], [y_0, y_1], \ldots, [y_{k-1}, y_k]\} \) are all of expansion less than \( \phi_k \), contradicting the definition of \( \phi_k \). Note that

\[ y_{i+1} \geq y_i \left(1 + \frac{8 \theta}{\phi_k}\right), \]

and thus \( y_{i+1} - y_i \geq \frac{8 \theta y_i}{\phi_k} \).

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

\[ \phi([y_i, y_{i+1}]) = w([y_i, y_{i+1}], [1, y_i] \cup [y_{i+1}, n]) \]

\[ \leq \frac{y_{i+1} - y_i}{y_{i+1} - y_i} \]

\[ \leq \frac{2 \theta y_{i+1} + 2 \theta y_i}{y_{i+1} - y_i} \]

\[ = 2 \theta \left(1 + \frac{2y_i}{y_{i+1} - y_i}\right). \]

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 \( \theta < \phi_k/4 \).
\end{proof}

2.2.2  

\textit{Induction}

In the following, we assume that \( \phi_k^2 \geq 1024 \lambda \), as otherwise Theorem 2 holds trivially. We will prove the following lemma by induction.

\begin{lemma}
If \( \phi_k^2 \geq 1024 \lambda \) and \( \phi(m_i) \geq 256k \lambda / \phi_k \) for all \( m_i \leq n/2 \), then \( x_{m_{i+1}} \leq 4x_{m_{i+1}} \) for all \( m_i \leq n/2 \).
\end{lemma}

As argued before, the assumptions of Lemma 4 would imply that \( x_i > 0 \) for all \( 1 \leq i \leq n/2 \), leading to a contradiction. So, the assumptions of Lemma 4 must not hold, and thus there is an \( m_i \) with \( \phi(m_i) \leq 256k \lambda / \phi_k \), proving Theorem 2.

To prove Lemma 4, we will prove by induction that \( x_{m_i} \leq c_i x_{m_i} \), where initially \( c_0 = 2 \) and

\[ c_{i+1} = \begin{cases} 
    c_i & \text{if } \phi(m_i) \geq \phi_k/4, \\
    c_i(1 - \epsilon_i c_i) & \text{if } \phi(m_i) < \phi_k/4, \text{where } \epsilon_i = 2\lambda/\phi(m_i). 
\end{cases} \]

10
We first assume this induction step and show that \( c_\infty \leq 4 \) using Lemma 3. Then we will verify the induction step.

**Claim 2.** \( c_\infty \leq 4 \).

**Proof.** First, we prove by induction that
\[
c_i = \frac{c_0}{1 - \sum_{j=0}^{i-1} \epsilon_j c_0}.
\]

Assume this is true for \( i \). Then
\[
c_{i+1} = \frac{c_i}{1 - \epsilon_i c_i} = \frac{c_0}{1 - \sum_{j=0}^{i-1} \epsilon_j c_0} \left( \frac{1}{1 - \epsilon_i \left( \frac{c_0}{1 - \sum_{j=0}^{i-1} \epsilon_j c_0} \right)} \right) = \frac{c_0}{1 - \sum_{j=0}^{i} \epsilon_j c_0}.
\]

Next, we bound \( c_\infty \) using Lemma 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 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,
\[
\sum_{i \geq 0} \sum_{j: \theta_i \leq \phi(m_j) \leq 2\theta_i} \epsilon_j \leq \sum_{i \geq 0} \sum_{j: \theta_i \leq \phi(m_j) \leq 2\theta_i} \frac{2\lambda}{\theta_i} \leq \sum_{i \geq 0} \frac{16k}{\phi_k} \frac{2\lambda}{\theta_i} = \sum_{i \geq 0} \frac{32k\lambda}{\phi_k} \frac{2\lambda}{256k\lambda} = \frac{1}{4}.
\]

Therefore,
\[
c_\infty = \frac{c_0}{1 - \sum_j \epsilon_j c_0} \leq \frac{c_0}{1 - \frac{c_\infty}{4}} = 4.
\]

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 equation (2.1) and we have
\[
x_{m_i+1} \geq x_{m_i} - \frac{2\lambda c_i x_{m_i}}{\phi(m_i)} = x_{m_i} (1 - \epsilon_i c_i) \geq x_{m_i} \left( \frac{1 - \epsilon_i c_i}{c_i} \right) \geq \overline{x}_{m_i+1},
\]
by the definition of \( \epsilon_i \) 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 \( \overline{x}_{m_i} \leq c_i x_{m_i+1} \), 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, we have
\[
\sum_{i=1}^{m_i} x_i = \sum_{i=1}^{m_i} x_i + \sum_{i=m_i+1}^{m_i+1} x_i \leq x_{m_i} (m_i + (c_i - 1)m_i).
\]

Similarly, dividing both sides of this inequality by \( m_i+1 \), we have
\[
\overline{x}_{m_i+1} \leq x_{m_i+1} (c_i + \frac{\phi(m_i)}{\phi(m_i)} \left( \frac{\phi(m_i)}{\phi(m_i)} - 2\lambda c_i \right) \leq c_i x_{m_i+1},
\]
where the last inequality follows from Claim 1 by plugging in \( k = \phi(m_i)/2 \), \( \varphi = \phi(m_i) \), \( c = c_i \) 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 4.
3 Personal Pagerank

We show that a similar and simpler analysis applies to the personal pagerank vector. 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 $r_{s,\alpha} = \alpha \chi_s + (1 - \alpha)W r_{s,\alpha}$, 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 $3 |S| \log(|S|) \leq n$.

3.1 Drop Lemma

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

$$x_a - x_b \leq \frac{\alpha}{w([1, a], [b, n])}. \quad (3.1)$$

3.2 Escaping Probability

Let $S$ be an unknown target set. Using a bound on the escaping probability of random walks [ST13]$, 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])

$$\sum_{i \in S} x_i \geq 1 - \frac{\phi(S)}{\alpha}. \quad (3.2)$$

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

$$x_a \geq \frac{2}{3a \log(|S|)}.$$

3.3 Vertex Expansion

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

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

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 $\phi(m_i) \geq 36 \phi(S) \log(|S|)/\phi^V(G)$ for all $m_i \leq 3 |S| \log(|S|)$. Then

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

$^9$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.
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
\[
x_m \geq \phi^V(S) \frac{1}{3\alpha \log(|S|)} \quad \text{and thus} \quad \sum_{j \geq 0} x_j \geq \sum_{0 \leq j \leq 3\alpha \log(|S|)} \phi^V(S) \frac{1}{3\alpha \log(|S|)} > 1,
\]
since \( a \leq |S| \), contradicting that \( x \) is a probability distribution vector. Therefore, there must exist an \( m_i \leq 3|S|\log(|S|) \) with \( \phi(m_i) \leq 36\phi(S)\log(|S|)/\phi^V(G) \), proving the first part of Theorem 3.

### 3.4 \( k \)-way Expansion

For \( k \)-way expansion, we define the jumping sequence by setting \( m_0 = a \) and
\[
m_{i+1} = \lceil m_i (1 + \phi(m_i)) \rceil.
\]
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
\[
x_{m_{i+1}} \geq x_{m_i} - \frac{2\alpha}{m_i \cdot \phi(m_i)} \quad \text{and} \quad x_m \geq x_a - \sum_{i \geq 0} \frac{2\alpha}{m_i \cdot \phi(m_i)}.
\]
We divide the summation into two parts
\[
\sum_{i : \phi(m_i) < \phi_k/4} \frac{2\alpha}{m_i \cdot \phi(m_i)} + \sum_{i : \phi(m_i) \geq \phi_k/4} \frac{2\alpha}{m_i \cdot \phi(m_i)}.
\]
The second part is at most
\[
\sum_{i : \phi(m_i) \geq \phi_k/4} \frac{2\alpha}{m_i \cdot \phi(m_i)} \leq \sum_{i : \phi(m_i) \geq \phi_k/4} \frac{8\alpha}{a \phi_k (1 + \phi_k/4)^i} \leq \frac{64\alpha}{a \phi_k^2}.
\]
The first part can be bounded by Lemma 3 as follows. Suppose by contradiction that \( \phi(m_i) \geq 1152k\phi(S)\log(|S|)/\phi_k \) for all \( m_i \leq 3|S|\log(|S|) \). Let \( \theta_0 = 1152k\phi(S)\log(|S|)/\phi_k \) and \( \theta_i = 2\theta_{i-1} \) for \( i \geq 1 \). By Lemma 3, there are at most \( 16k/\phi_k \) terms in the jumping sequence having conductance between \( \theta \) and \( 2\theta \) when \( \theta < \phi_k/4 \). Therefore, the first part is at most
\[
\sum_j \sum_{i : \theta_j \leq \phi(m_i) \leq 2\theta_j} \frac{2\alpha}{a \phi(m_i)} \leq \sum_j \frac{32k}{a \phi_k} \frac{\alpha}{a \theta_j} = \sum_j \frac{32k}{a \phi_k} \frac{\alpha}{a \theta_0} = \frac{64k\alpha}{a \phi_k \theta_0}.
\]
Putting these back into the first inequality, we have
\[
x_m \geq x_a - \frac{64\alpha}{a \phi_k^2} - \frac{64k\alpha}{a \phi_k \theta_0} \geq \frac{2}{3a \log(|S|)} - \frac{192\phi(S)}{a \phi_k} - \frac{192\phi(S)k}{a \phi_k \theta_0} \geq \frac{1}{3a \log(|S|)} \quad \text{and thus} \quad x_0 \geq \frac{1}{3a \log(|S|)} > 1,
\]
where the second inequality is by the lower bound of \( x_a \) and the choice of \( \alpha = 3\phi(S) \), and the last inequality is by our choice of \( \theta_0 \). This implies that
\[
x_0 \geq \frac{1}{3a \log(|S|)} \quad \text{and thus} \quad \sum_{j \geq 0} x_j \geq \sum_{0 \leq j \leq 3a \log(|S|)} \frac{1}{3a \log(|S|)} > 1,
\]
since \( a \leq |S| \), contradicting that \( x \) is a probability distribution vector. Therefore, there must exist an \( m_i \leq 3|S|\log(|S|) \) with \( \phi(m_i) \leq 1152k\phi(S)\log(|S|)/\phi_k \), proving the second part of Theorem 3.
3.5 Local Algorithm

Andersen and Chung [AC07] show that the drop lemma (equation (3.1)) 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 $\epsilon$-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 non-negative and satisfies $q(u) \leq \epsilon$ for every vertex $u$ in the graph.

**Lemma 5 ([AC07]).** There is an algorithm that computes an $\epsilon$-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 $0 \leq q \leq \epsilon I$ implies

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

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

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

for those vertices $s$ that satisfy equation (3.2). Hence there exists an $a \leq |S|$ with

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

Since $r'_{s,\alpha}$ satisfies the drop lemma (equation (3.1)) and contains good initial value, both arguments in vertex expansion and $k$-way expansion follow (with the assumption $3|S|\log(|S|) \leq n$ replaced by $6|S|\log(|S|) \leq n$). 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, 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 4. 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 6.** 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 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
\[
\mathcal{R}(p_{s,t}) = \frac{p_{s,t}^T L p_{s,t}}{\| p_{s,t} \|^2} = \frac{\sum_{i=1}^{n} c_i^2 (1 - \lambda_i/2)^{2t} \lambda_i}{\sum_{i=1}^{n} c_i^2 (1 - \lambda_i/2)^{2t}} = 2 - \frac{\sum_{i=1}^{n} c_i^2 (1 - \lambda_i/2)^{2t+1}}{\sum_{i=1}^{n} c_i^2 (1 - \lambda_i/2)^{2t}}.
\]

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 \( \mathcal{R}(p_{s,t}) = 2 - 2E[X^{2t+1}] / E[X^{2t}] \). By the power mean inequality and the non-negativity of \( X \), we have
\[
E[X^{2t+1}]^{1/(2t+1)} \geq E[X^{2t}]^{1/(2t)}.
\]

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

4.2 Small Support Vector with Small Rayleigh Quotient

A vector \( x \) is called spectrally \( \delta \)-sparse if \( \| x \|_2 \leq \delta n \| x \|_2^2 \). 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 5 ([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
\[
\sum_{v \in S} p_{s,t}(v) \geq \frac{1}{200} \left( 1 - \frac{3\phi(S)}{2} \right)^t.
\]

This provides a bound on the spectral sparsity of \( p_{s,t} \).
Lemma 7. 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
\[ \|p_{s,t}\|^2_1 \leq \frac{40000|S|}{(1 - 3\phi(S))/2} \|p_{s,t}\|^2_2. \]

Proof. By Cauchy-Schwarz and Theorem 5, we have
\[ \|p_{s,t}\|^2_2 \geq \sum_{v \in S} p_{s,t}(v)^2 \geq \frac{1}{|S|} \left( \sum_{v \in S} p_{s,t}(v) \right)^2 \geq \frac{1}{|S|} \left( \frac{1}{200} \left( 1 - \frac{3\phi(S)}{2} \right) \right)^2 \geq \frac{1}{40000|S|} \left( 1 - \frac{3\phi(S)}{2} \right)^2. \]

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

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

Lemma 8 ([ABS10]). Let $x \in \mathbb{R}^{|V|}_{\geq 0}$ be a non-negative vector with $\|x\|^2_1 \leq \delta n \|x\|^2_2$. Then there exists a vector $y$ with $\text{supp}(y) = O(\delta n)$ and $\mathcal{R}(y) = O(\mathcal{R}(x))$.

We will apply Lemma 8 on $p_{s,t}$ and obtain a vector with small Rayleigh quotient (Lemma 6) and small support (Lemma 7).

### 4.3 Improved Cheeger’s Guarantees

We are ready to prove Theorem 4. In the following we assume $\phi(S) \leq \frac{1}{4}$ and $|S| \geq 2$. We set $t = \epsilon \log |S|/(6\phi(S))$ so that
\[ \left( 1 - \frac{3\phi(S)}{2} \right)^{2t} \geq \exp(-3\phi(S))^{2t} = \exp(-6t\phi(S)) = \exp(-\epsilon \log |S|) = |S|^{-\epsilon}. \]

By Lemma 7, we have
\[ \|p_{s,t}\|^2_1 \leq 40000|S|^{1+\epsilon} \|p_{s,t}\|^2_2. \]

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 6, we have
\[ \mathcal{R}(p_{s,t}) \leq 2 - 2(1 - \frac{3\phi(S)}{2})(1 - \frac{51\phi(S)}{2\epsilon}) = O \left( \frac{\phi(S)}{\epsilon} \right). \]

Now, we apply Lemma 8 by plugging the vector $p_{s,t}$ for $x$ and obtain a vector $y$ with $\text{supp}(y) \leq O(|S|^{1+\epsilon})$ and $\mathcal{R}(y) = O(\phi(S)/\epsilon)$. Finally, by the proof of the improved Cheeger’s inequality (1.2) (see Section A in Appendix), we find a level set $S'$ with $|S'| \leq |\text{supp}(y)| = O(|S|^{1+\epsilon})$ 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 4.
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 9 ([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 compute a vector \( p'_{s,t} \) such that \( p_{s,t} \geq p'_{s,t} \geq p_{s,t} - \alpha \mathbf{1} \) 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 \( O(d\epsilon^2|S|^{1+\epsilon}) \), so that the time complexity of our local algorithm is \( O(d\epsilon^2|S|^{1+\epsilon}) \), so that the time complexity of our local algorithm is \( O(d\epsilon^2|S|^{1+\epsilon}) \), so that the time complexity of our local algorithm is \( O(d\epsilon^2|S|^{1+\epsilon}) \). It remains to show that \( p'_{s,t} \) is still spectrally sparse and has small Rayleigh quotient.

**Lemma 10.** For \( p_{s,t} \) that satisfies the conclusion in Lemma 7, we have

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

**Proof.** In the proof, we let \( x := p_{s,t} \) and \( y := p'_{s,t} \). By Lemma 9, 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 7, we have \( \|x\|_2^2 \geq 1/(40000|S|^{1+\epsilon}) \). From our choice of \( \alpha \), we have \( 2\alpha = \phi(S)/(80000|S|^{1+\epsilon}) \leq \phi(S)/|x|_2^2 \). 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 \( \|y\|_1^2 \leq \|x\|_1^2 = 1 \).

**Lemma 11.**

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

**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

\[
\mathcal{R}(y) = \frac{y^T Ly}{y^T y} = \left( x - r \right)^T L (x - r) = \frac{x^T L x + r^T L r - 2x^T L r}{y^T y} \leq \frac{2x^T L x + 2r^T L r}{y^T y}.
\]

By the calculation in Lemma 10, \( \|r\|_2^2 \geq (1 - \phi(S)/2) \|x\|_2^2 \). Hence, using \( r \geq 0 \) and \( y \geq 0 \), we have \( \|r\|_2^2 \leq \|x\|_2^2 \leq \phi(S)\|x\|_2^2 \). By the calculation in Lemma 10, \( \|r\|_2^2 \geq (1 - \phi(S)/2) \|x\|_2^2 \). Hence, using \( r \geq 0 \) and \( y \geq 0 \), we have \( \|r\|_2^2 \leq \|x\|_2^2 \leq \phi(S)\|x\|_2^2 /2 \) and \( r^T L r \leq 2r^T r \leq \phi(S)\|x\|_2^2 \). So, we have

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

\[\square\]
With Lemma 10 and Lemma 11, we can use the same proof in Section 4.3 to prove Theorem 4 with the time complexity claimed.

To prove Theorem 4(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 \( \lambda_S \) 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 \( \lambda_S \) 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 \( \lambda_S \). 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 \(|S|/(1 - \lambda_S)^{2t} \leq |S|/(1 - O(\phi(S)))^{2t} \). This allows us to set \( t \) to be larger so as to improve the Rayleigh quotient of the random walk vector.

### Acknowledgement

Part of the work was done while we were long-term participants in the Algorithmic Spectral Graph Theory program at the Simons Institute for the Theory of Computing in Fall 2014. We thank the organizers for the support and the wonderful research environment. We also thank David Steurer for suggesting the spectral approach to analyze random walks, and Luca Trevisan for pointing out that Theorem 2 can be derived from the proof of the improved Cheeger’s inequality in [KLLOT13].

### References


A Relations between Improved Cheeger’s Inequality and Theorem 2

First, we show that Theorem 2 can be derived from the proof of the improved Cheeger’s inequality, as pointed out to us by Luca Trevisan. Then, we show that the improved Cheeger’s inequality can be derived from Theorem 2, using a graph powering trick as described in [KL14].

A.1 Improved Cheeger’s inequality implies Theorem 2

The following stronger statement was shown in [KLLOT13].

Theorem 6 (Theorem 3.5 of [KLLOT13], restated). For any non-negative vector $x$ with $\text{supp}(x) \leq n/2$, let $\phi_{\text{sweep}}(x)$ be the minimum expansion of the level sets of $x$. At least one of the following holds:

1. $\phi_{\text{sweep}}(x) \leq O(k)R(x)$.

2. There exists $k$ 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 $R(x_i) \leq O(k^2R(x)^2/\phi_{\text{sweep}}(x)^2)$.

We apply the theorem with $x = \max(v_2, 0)$ or $x = \max(-v_2, 0)$, whichever of smaller support. Note that $R(x) \leq \lambda_2$ by standard argument [HLW06]. When the first case of Theorem 6 holds, it is clear that

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

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

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

and Theorem 2 follows.
A.2 Theorem 2 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 2.

**Theorem 7** (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

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

The following corollary is a generalization of Corollary 12 of [KL14], which shows that general cases of improved Cheeger’s inequality can be reduce to the cases where $\lambda_k$ is constant.

**Corollary 3.** Suppose one could prove that $\phi(H) \leq C\lambda_2(H)$ for some $C \geq 1/10$ whenever $\lambda_k(H) \geq 1/4$, 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

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

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

$$C\lambda_2(H) \geq \phi(H) \geq \frac{1}{20}(1 - (1 - \frac{\phi(G)}{2})^{\sqrt{1/\lambda_k(G)}}) \geq \frac{1}{20}(1 - \exp(-\frac{\phi(G)}{2\sqrt{\lambda_k(G)}})) \geq \frac{\phi(G)}{80\sqrt{\lambda_k(G)}},$$

where the second inequality is by Theorem 7. On the other hand,

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

and the corollary follows by combining the two inequalities. □

Now we show the improved Cheeger’s inequality in [KLLOT13] follows from Corollary 3 and Theorem 2. 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 2 gives $\phi(G) = O(k\lambda_2(G))$. Therefore, we can apply Corollary 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 $\lambda_k$, and the improved Cheeger’s inequality in [KLLOT13] follows.