


Cheeger Inequalities for Vertex Expansion and Reweighted Eigenvalues

Tsz Chiu Kwok

Institute for Theoretical Computer Science
Shanghai University of Finance and Economics
Shanghai, China
kwok@mail.shufe.edu.cn

Lap Chi Lau

School of Computer Science
University of Waterloo
Waterloo, Canada
lapchi@uwaterloo.ca

Kam Chuen Tung 

School of Computer Science
University of Waterloo
Waterloo, Canada
kctung@uwaterloo.ca

Abstract—The classical Cheeger’s inequality relates the edge conductance of a graph and the second smallest eigenvalue of the Laplacian matrix. Recently, Olesker-Taylor and Zanetti discovered a Cheeger-type inequality connecting the vertex expansion of a graph and the maximum reweighted second smallest eigenvalue of the Laplacian matrix.

In this work, we first improve their result to a logarithmic dependence on the maximum degree in the graph, which is optimal up to a constant factor. Also, the improved result holds for weighted vertex expansion, answering an open question by Olesker-Taylor and Zanetti.

Building on this connection, we then develop a new spectral theory for vertex expansion. We discover that several interesting generalizations of Cheeger inequalities relating edge conductances and eigenvalues have a close analog in relating vertex expansions and reweighted eigenvalues. These include an analog of Trevisan’s result on bipartiteness, an analog of higher order Cheeger’s inequality, and an analog of improved Cheeger’s inequality.

Finally, inspired by this connection, we present negative evidence to the 0/1-polytope edge expansion conjecture by Mihail and Vazirani. We construct 0/1-polytopes whose graphs have very poor vertex expansion. This implies that the fastest mixing time to the uniform distribution on the vertices of these 0/1-polytopes is almost linear in the graph size.

Index Terms—Cheeger inequalities, vertex expansion, reweighted eigenvalues, mixing time, spectral analysis

I. INTRODUCTION

The connection between vertex expansion and reweighted eigenvalue is discovered through the study of the fastest mixing time problem introduced by Boyd, Diaconis and Xiao [11]. In the fastest mixing time problem, we are given an undirected graph $G = (V, E)$ and a target probability distribution $\pi : V \rightarrow \mathbb{R}$. The task is to find a time-reversible transition matrix $P \in \mathbb{R}^{|V| \times |V|}$ supported on the edges of G , so that the stationary distribution of random walks with transition matrix P is π . The objective is to find such a transition matrix that minimizes the mixing time to the stationary distribution π . It is well-known that the mixing time to the stationary distribution is approximately inversely proportional to the spectral gap $1 - \alpha_2(P)$ of the time-reversible transition matrix P , where

The first author is supported by Science and Technology Innovation 2030 - “New Generation of Artificial Intelligence” Major Project No.(2018AAA0100903), NSFC grant 61932002, Program for Innovative Research Team of Shanghai University of Finance and Economics and the Fundamental Research Funds for the Central Universities. The second and third authors are supported by NSERC Discovery Grant.

$1 = \alpha_1(P) \geq \alpha_2(P) \geq \dots \geq \alpha_{|V|}(P) \geq -1$ are the eigenvalues of P . The fastest mixing time problem is thus formulated as follows in [11] by the maximum spectral gap achievable through such a “reweighting” P of the input graph G .

Definition I.1 (Maximum Reweighted Spectral Gap [11]). *Given an undirected graph $G = (V, E)$ and a probability distribution π on V , the maximum reweighted spectral gap is defined as*

$$\begin{aligned} \lambda_2^*(G) &:= \max_{P \geq 0} 1 - \alpha_2(P) \\ \text{subject to} & \quad P(u, v) = P(v, u) = 0 \quad \forall uv \notin E \\ & \quad \sum_{v \in V} P(u, v) = 1 \quad \forall u \in V \\ & \quad \pi(u)P(u, v) = \pi(v)P(v, u) \quad \forall uv \in E. \end{aligned}$$

The graph is assumed to have a self-loop on each vertex, to ensure that the optimization problem for $\lambda_2^*(G)$ is always feasible. In the context of Markov chains, this corresponds to allowing a non-negative holding probability on each vertex.

The last constraint is the time reversible condition to ensure that the transition matrix P corresponds to random walks on an undirected graph (where the edge weight of uv is $\pi(u)P(u, v)$) and that the stationary distribution of P is π . Note that $\lambda_2^*(G) = \max_{P \geq 0} (1 - \alpha_2(P)) = \max_{P \geq 0} \lambda_2(I - P)$, which is the maximum reweighted second smallest eigenvalue of the normalized Laplacian matrix of G (where the edge weight of uv is $\pi(u)P(u, v)$) subject to the above constraints.

Boyd, Diaconis and Xiao showed that this optimization problem can be written as a semidefinite program and thus $\lambda_2^*(G)$ can be computed in polynomial time. Subsequently, the fastest mixing time problem has been studied in various work (see [9], [10], [15], [20], [37] and more references in [35]), but no general characterization was known. Roch [37] showed that the vertex expansion $\psi(G)$ is an upper bound on the optimal spectral gap $\lambda_2^*(G)$.

Definition I.2 (Weighted Vertex Expansion). *Let $G = (V, E)$ be an undirected graph and π be a probability distribution on V . For a subset $S \subseteq V$, let $\partial S := \{v \notin S \mid \exists u \in S \text{ with } uv \in E\}$ be the vertex boundary of S , and $\pi(S) := \sum_{v \in S} \pi(v)$ be*

the weight of S . The weighted vertex expansion of a set $S \subseteq V$ and of a graph G are defined as $\psi(S) := \frac{\pi(\partial S)}{\pi(S)}$ and

$$\psi(G) := \min \left\{ 1, \min_{S \subseteq V: 0 < \pi(S) \leq 1/2} \psi(S) \right\}.$$

When π is the uniform distribution, $\psi(S)$ is the usual vertex expansion $|\partial S|/|S|$.

Recently, Olesker-Taylor and Zanetti [35] discovered an elegant Cheeger-type inequality for vertex expansion and the maximum reweighted spectral gap, showing that small vertex expansion is qualitatively the only obstruction for the fastest mixing time to be small. Note that their result only holds when π is the uniform distribution.

Theorem I.3 (Cheeger Inequality for Vertex Expansion [35]). *For any undirected graph $G = (V, E)$ and the uniform distribution $\pi = \bar{1}/|V|$,*

$$\frac{\psi(G)^2}{\log |V|} \lesssim \lambda_2^*(G) \lesssim \psi(G).$$

In terms of the fastest mixing time $\tau^(G)$ to the uniform distribution, $\frac{1}{\psi(G)} \lesssim \tau^*(G) \lesssim \frac{\log^2 |V|}{\psi^2(G)}$. (See section II for definitions for random walks and mixing time.)*

Unlike Cheeger's inequality for edge conductance where $\phi(G)^2 \lesssim \lambda_2(G) \lesssim \phi(G)$, it is noted in [35] that the $\log |V|$ term might not be completely removed: Louis, Raghavendra and Vempala [34] proved that it is NP-hard to distinguish between $\psi(G) \leq \epsilon$ and $\psi(G) \gtrsim \sqrt{\epsilon \log d}$ for every $\epsilon > 0$ where d is the maximum degree of the graph G , assuming the small-set expansion conjecture of Raghavendra and Steurer [36].

Besides the fastest mixing time problem, we note that these "reweighting problems" relating vertex expansion and reweighted eigenvalues are also well motivated in the study of approximation algorithms. One example is a conjecture of Arora and Ge [5, Conjecture 12], which roughly states that, if a graph G has almost perfect vertex expansion for every set, then there exists a reweighted doubly stochastic matrix P of the adjacency matrix of G so that P has few eigenvalues less than $-\frac{1}{16}$. They proved that if the conjecture was true, then there is an improved subexponential time algorithm for coloring 3-colorable graphs. Another example is a conjecture of Steurer [39, Conjecture 9.2], which is also known to be related to a reweighting problem between vertex expansion and the graph spectrum, that if true would imply an improved subexponential time approximation algorithm for the sparsest cut problem.

A. Our Results

First we improve and generalize the result of Olesker-Taylor and Zanetti. Then we build on this new connection to develop

¹When π is the uniform distribution, $\min_{S \subseteq V: 0 < \pi(S) \leq 1/2} \psi(S)$ is always at most 1. For general π , however, this could be arbitrarily large. Consider for example a star graph where the center has most of the π -weight. Therefore, we need to put an upper bound of 1 on $\psi(G)$, as otherwise $\psi(G)$ cannot be bounded by eigenvalues of the normalized Laplacian matrix which are always upper bounded by 2.

a spectral theory for vertex expansion. Finally we present 0/1-polytopes with poor vertex expansion and discuss the implications to the 0/1-polytope expansion conjecture.

1) *Optimal Cheeger Inequality for Vertex Expansion:* Olesker-Taylor and Zanetti [35] posed the problem of reducing the $\log |V|$ factor in Theorem I.3 to $\log d$, and also the problem of generalizing their result to weighted vertex expansion. Our first result provides a positive answer to these two questions.

Theorem I.4 (Cheeger Inequality for Weighted Vertex Expansion). *For any undirected graph $G = (V, E)$ with maximum degree d and any probability distribution π on V ,*

$$\frac{\psi(G)^2}{\log d} \lesssim \lambda_2^*(G) \lesssim \psi(G).$$

In terms of the fastest mixing time $\tau^(G)$ to the stationary distribution, $\frac{1}{\psi(G)} \lesssim \tau^*(G) \lesssim \frac{\log d \cdot \log \pi_{\min}^{-1}}{\psi^2(G)}$.*

We show that the $\log d$ factor in Theorem I.4 is optimal, by exhibiting graphs G with $\lambda_2^*(G) \asymp \frac{\psi(G)^2}{\log d}$. Note that the tightness result does not rely on the small-set expansion hypothesis.

We note that Louis, Raghavendra and Vempala [34] gave an SDP approximation algorithm for vertex expansion with the same approximation guarantee, but their SDP is different from and stronger than that in Definition I.1 (see Lemma III.10), and so it does not have the natural interpretation as the reweighted second eigenvalue and does not imply the result on fastest mixing time. The proof of Theorem I.4 is based on the techniques in [8], [34], which we will discuss in detail in section III-A2.

2) *Maximum Reweighted Lower Spectral Gap and Bipartite Vertex Expansion:* Trevisan [42] proved that the lower spectral gap $1 + \alpha_{\min}(G)$ of the normalized adjacency matrix of $G = (V, E)$ is small if and only if there is a subset $S \subseteq V$ which is an almost bipartite component in G with small edge conductance $\phi(S)$. We define the analogous notions for vertex expansion and for reweighted lower spectral gap.

Definition I.5 (Bipartite Vertex Expansion). *Given an undirected graph $G = (V, E)$, the bipartite vertex expansion of G is defined as*

$$\psi_B(G) := \min \left\{ 1, \min_{\emptyset \neq S \subseteq V} \left\{ \psi(S) \mid G[S] \text{ is an induced bipartite graph} \right\} \right\}.$$

Definition I.6 (Maximum Reweighted Lower Spectral Gap). *Given an undirected graph $G = (V, E)$ and a probability distribution π on V , the maximum reweighted lower spectral gap is defined as*

$$\begin{aligned} \zeta^*(G) &:= \max_{P \geq 0} \lambda_{\min}(D_P + P) \\ \text{subject to} & \quad P(u, v) = P(v, u) = 0 \quad \forall uv \notin E \\ & \quad \sum_{v \in V} P(u, v) \leq 1 \quad \forall u \in V \\ & \quad \pi(u)P(u, v) = \pi(v)P(v, u) \quad \forall uv \in E. \end{aligned}$$

where D_P is the diagonal matrix of row sums of P such that $D_P(u, u) = \sum_{v \in V} P(u, v)$ for $u \in V$. We note that this program is slightly different from that in Definition I.1, and the main reason is that self-loops should not be allowed in this problem.

We prove an analog of Trevisan's result that the maximum reweighted lower spectral gap is small if and only if there is an induced bipartite subgraph on S with small vertex expansion $\psi(S)$.

Theorem I.7 (Cheeger Inequality for Bipartite Vertex Expansion). *For any undirected graph $G = (V, E)$ with maximum degree d and any probability distribution π on V ,*

$$\frac{\psi_B(G)^2}{\log d} \lesssim \zeta^*(G) \lesssim \psi_B(G).$$

This is the first approximation algorithm for bipartite vertex expansion to our knowledge. Finding a two-colorable set with small vertex expansion is one of the three ways in Blum's coloring tools [7] to make progress in designing approximation algorithms for coloring 3-colorable graphs. Indeed, it is in this context that Arora and Ge [5] made the reweighting conjecture mentioned in the introduction. Theorem I.7 does not imply anything new about approximating graph coloring, but we hope that it is a step towards answering Arora and Ge's conjecture.

3) *Higher-Order Cheeger Inequality for Vertex Expansion:* Lee, Oveis Gharan and Trevisan [29] and Louis, Raghavendra, Tetali and Vempala [33] proved the higher-order Cheeger inequalities, which state that the k -th smallest eigenvalue $\lambda_k(G)$ of the normalized Laplacian matrix of $G = (V, E)$ is small if and only if the k -way edge conductance $\phi_k(G)$ is small. More precisely, they proved that $\lambda_k(G) \lesssim \phi_k(G) \lesssim k^2 \sqrt{\lambda_k}$ and $\lambda_{\frac{k}{2}}(G) \lesssim \sqrt{\lambda_k \log k}$. We consider the analogous notion of k -way vertex expansion.

Definition I.8 (k -Way Vertex Expansion). *Given an undirected graph $G = (V, E)$ and a probability distribution π on V , the k -way vertex expansion of G is defined as*

$$\psi_k(G) := \min \left\{ 1, \min_{S_1, \dots, S_k \subseteq V} \max_{1 \leq i \leq k} \psi(S_i) \right\},$$

where the minimum is taken over pairwise disjoint subsets S_1, \dots, S_k of V .

Definition I.9 (Maximum Reweighted k -th Smallest Eigenvalue). *Given an undirected graph $G = (V, E)$ and a probability distribution π on V , the maximum reweighted k -th smallest eigenvalue of the normalized Laplacian matrix of G is defined as $\lambda_k^*(G) := \max_{P \geq 0} \lambda_k(I - P)$, where P is subject to the same constraints stated in Definition I.1.*

We prove an analog of higher-order Cheeger inequalities that the maximum reweighted k -th smallest eigenvalue is small if and only if the k -way vertex expansion is small. As in previous work [29], [33], there is a better approximation guarantee if we consider only $\frac{k}{2}$ -way vertex expansion.

Theorem I.10 (Higher-Order Cheeger Inequality for Vertex Expansion). *For any undirected graph $G = (V, E)$ with maximum degree d and any probability distribution π on V ,*

$$\lambda_k^*(G) \lesssim \psi_k(G) \lesssim k^{\frac{3}{2}} \log k \sqrt{\log d \cdot \lambda_k^*(G)}$$

and

$$\psi_{\frac{k}{2}}(G) \lesssim \sqrt{k} \log k \sqrt{\log d \cdot \lambda_k^*(G)}.$$

Chan, Louis, Tang and Zhang [12] developed a spectral theory for hypergraphs and proved a higher-order Cheeger inequality for hypergraph (edge) expansion. Through a reduction from vertex expansion to hypergraph expansion, they proved that $\psi_{\frac{k}{2}}(G) \lesssim k^{\frac{5}{2}} \log k \log \log k \cdot \log d \cdot \sqrt{\xi_k}$ for graphs with bounded ratio between the maximum degree and the minimum degree, where $\xi_k \lesssim \psi_k(G)$ is a relaxation for k -way vertex expansion. Compared to their result, Theorem I.10 does not require the assumption about the maximum degree and the minimum degree of G , and has a better approximation ratio for $\frac{k}{2}$ -way vertex expansion. Furthermore, Theorem I.10 provides the first true approximation algorithm for k -way vertex expansion $\psi_k(G)$ to our knowledge.

4) *Improved Cheeger Inequality for Vertex Expansion:* Kwok, Lau, Lee, Oveis Gharan, and Trevisan [27] proved an improved Cheeger inequality that $\phi(G) \lesssim k \lambda_2(G) / \sqrt{\lambda_k(G)}$ for any $k \geq 2$. This shows that $\lambda_2(G)$ is a tighter approximation to $\phi(G)$ when $\lambda_k(G)$ is large for a small k . The result provides an explanation for the good empirical performance of the spectral partitioning algorithm.

We prove an analogous result that if the $\lambda_k^*(G)$ is large for a small k , then $\lambda_2^*(G)$ is a tighter approximation to the vertex expansion $\psi(G)$. The following result is close to the tight result in [27] for edge conductance.

Theorem I.11 (Improved Cheeger Inequality for Vertex Expansion). *For any undirected graph $G = (V, E)$ with maximum degree d , and for any probability distribution π on V and any $k \geq 2$,*

$$\lambda_2^*(G) \lesssim \psi(G) \lesssim \frac{k^{\frac{3}{2}} \cdot \lambda_2^*(G) \cdot \log d}{\sqrt{\lambda_k^*(G)}}.$$

We remark that the reweighting used in $\lambda_2^*(G)$ and $\lambda_k^*(G)$ may be different. Through Theorem I.10, we obtain the following corollary that only depends on the graph structure: If the k -way vertex expansion $\psi_k(G)$ is large for a small k , then $\lambda_2^*(G)$ is a tighter approximation to $\psi(G)$.

5) *Vertex Expansion of 0/1-Polytopes:* Mihail and Vazirani (see [18]) conjectured that the graph $G = (V, E)$ (i.e. 1-skeleton) of any 0/1-polytope is an edge expander, such that $|\delta(S)|/|S| \geq 1$ for every subset $S \subseteq V$ with $|S| \leq |V|/2$, where $\delta(S)$ denotes the set of edges between S and $V \setminus S$. This conjecture would imply fast mixing time of random walks to the stationary distribution, with applications in designing fast sampling algorithms for many classes of combinatorial objects. The conjecture is proved to be correct in several cases [3], [18], [25], most notably the recent resolution of the matroid

expansion conjecture [3] by Anari, Liu, Oveis Gharan and Vintzant.

In all these positive results, the Markov chain can be set up so that the stationary distribution is the uniform distribution, with the mixing time to the stationary distribution poly-logarithmic in the graph size. Then the fast sampling algorithms can also be used to obtain an approximate counting algorithm on the number of vertices in the given 0/1-polytope, with poly-logarithmic runtime in the graph size. Therefore, sampling from the uniform distribution is usually the setting of interest.

Inspired by the connection between fastest mixing time and vertex expansion, we consider a variant of Mihail and Vazirani’s conjecture: Is the graph of every 0/1-polytope a vertex expander? Perhaps surprisingly, we show that there are 0/1-polytopes whose graphs are very poor vertex expanders.

Theorem I.12 (0/1-Polytopes with Poor Vertex Expansion). *Let π be the uniform distribution. For any $k > 2$ and any $n > 2k$ sufficiently large, there is a 0/1-polytope $Q = Q_{n,k} \subseteq \{0, 1\}^n$ with $O(n^k)$ vertices and*

$$\psi(Q) \lesssim \frac{(4k)^k}{n^{k-2}}.$$

Theorem I.12 and Theorem I.3 together imply that even the fastest mixing time of the reversible random walks on some 0/1-polytopes is almost linear in the graph size.

Corollary I.13 (Trepid Mixing to Uniform Distribution). *For any constant $k > 2$, there exists a 0/1-polytope Q such that any reversible Markov chain on its graph $G_Q = (V, E)$ with stationary distribution $\vec{1}/|V|$ has mixing time $\Omega(|V|^{1-\frac{2}{k}})$.*

While Theorem I.12 does not provide a counterexample to the conjecture of Mihail and Vazirani, it shows that even if the conjecture is true, there are 0/1-polytopes for which random walks cannot be used for efficient uniform sampling and for efficient approximate counting.

Remark I.14. *After posting the first version of this paper on arXiv, we recently found out that Gillmann [22, Chapter 3.2] has already constructed examples of 0/1-polytopes whose graphs have poor vertex expansion. The polytopes $Q \subseteq \{0, 1\}^n$ constructed have $2^{(h(c)+o(1))n}$ vertices and satisfy*

$$\psi(Q) \lesssim 2^{-(h(c)-2c)n},$$

where $h(x) := -x \log x - (1-x) \log(1-x)$ is the binary entropy function and $c := 1/5$ (correspondingly $h(c) = 0.7219\dots$). Applying Theorem I.3, this would imply a fastest mixing time bound of $\Omega(|V|^{0.4459\dots})$. By choosing smaller values of c , an almost linear fastest mixing time bound can be obtained as in Corollary I.13.

B. Related Work

In this subsection, we review previous spectral approaches for vertex expansion and compare them to the current approach using reweighted eigenvalues. For previous results about Cheeger’s inequalities for edge conductances mentioned

in the introduction, they will be discussed in the corresponding technical sections.

Second Eigenvalue and Vertex Expansion: There are classical results in spectral graph theory relating vertex expansions and (ordinary) eigenvalues. For any graph $G = (V, E)$ with maximum degree d , let $\lambda'_2(G)$ be the second smallest eigenvalue of the (unnormalized) Laplacian matrix, it is known that

$$\psi(G) \geq \frac{2\lambda'_2(G)}{d + 2\lambda'_2(G)} \quad \text{and} \quad \lambda'_2(G) \geq \frac{\psi(G)^2}{4 + 2\psi(G)^2},$$

where the first inequality is the “easy” direction proved by Tanner [40] and Alon and Milman [2], and the second inequality is the “hard” direction proved by Alon [1]. These imply that $\lambda'_2(G)$ can be used to give an $O(\sqrt{d \cdot \psi(G)})$ -approximation algorithm to $\psi(G)$. Compared to Cheeger’s inequality for edge conductance that $\phi(G)^2/2 \leq \lambda_2(G) \leq 2\phi(G)$ where $\lambda_2(G)$ is the second smallest eigenvalue of the normalized Laplacian matrix, there is an extra factor d between the upper and lower bounds.

Spectral Formulation: Bobkov, Houdré and Tetali [8] defined an interesting “spectral” quantity called λ_∞ (see Definition III.5), which satisfies an exact analog of Cheeger’s inequality for symmetric vertex expansion:

$$\frac{1}{2}\psi_{\text{sym}}(G)^2 \leq \lambda_\infty \leq 2\psi_{\text{sym}}(G),$$

where the symmetric vertex boundary of a set $S \subseteq V$ is defined as $\partial_{\text{sym}}(S) := \partial(S) \cup \partial(V - S)$ and the symmetric vertex expansion of S is defined as $\psi_{\text{sym}}(S) := |\partial_{\text{sym}}(S)|/|S|$, and the symmetric vertex expansion of a graph G is defined as $\psi_{\text{sym}}(G) := \min_{S: |S| \leq |V|/2} \psi_{\text{sym}}(S)$. However, it is not known how to compute λ_∞ efficiently, and it is recently shown to be NP-hard to compute λ_∞ by Farhadi, Louis and Tetali [17].

Semidefinite Programming Relaxations: Louis, Raghavendra and Vempala [34] gave a semidefinite programming relaxation sdp_∞ for λ_∞ , and proved that for any graph $G = (V, E)$ with maximum degree d ,

$$\frac{\psi_{\text{sym}}(G)^2}{\log d} \lesssim \text{sdp}_\infty \lesssim \psi_{\text{sym}}(G).$$

Then, by constructing a graph H such that $\psi_{\text{sym}}(H) = \Theta(\psi(G))$, they reduce vertex expansion to symmetric vertex expansion and obtain a Cheeger’s inequality for $\psi(G)$, one that is of the same form as in Theorem I.4 for $\lambda_2^*(G)$. We will show in Lemma III.10 that $\lambda_2^*(G)$ and sdp_∞ are different and sdp_∞ is a stronger relaxation such that $\lambda_2^*(G) \leq \text{sdp}_\infty$.

The current best known approximation algorithm for vertex expansion $\psi(G)$ is an $O(\sqrt{\log |V|})$ SDP-based approximation algorithm by Feige, Hajiaghayi and Lee [19]. This is an extension of the $O(\sqrt{\log |V|})$ SDP-based approximation algorithm for edge conductance $\phi(G)$ by Arora, Rao, and Vazirani [6]. The SDP formulation of [6] is known to be strictly more powerful than the spectral formulation by the second eigenvalue.

Even though $\lambda_2^*(G)$, sdp_∞ and the SDP in [19] are all semidefinite programming relaxations for $\psi(G)$ and satisfy

similar inequalities, we note that the approach of using reweighted eigenvalues has some additional features. One important feature is that $\lambda_2^*(G)$ is closely related to fastest mixing time. This allows one to develop a spectral theory for vertex expansion that relates (i) vertex expansion, (ii) reweighted eigenvalues and (iii) fastest mixing time, which parallels the classical spectral graph theory that relates (i) edge conductance, (ii) eigenvalues and (iii) mixing time. Another feature is that it allows one to extend known generalizations of Cheeger inequalities to the vertex expansion setting, and as a consequence to obtain approximation algorithms for bipartite vertex expansion and k -way vertex expansion.

Spectral Hypergraph Theory: Louis [32] and Chan, Louis, Tang, Zhang [12] developed a spectral theory for hypergraphs. They defined a continuous time diffusion process on a hypergraph $H = (V, E)$ and used it to define the Laplacian operator and its eigenvalues $\gamma_1 \leq \gamma_2 \leq \dots \leq \gamma_{|V|}$. The formulation is similar to the one in [8] for vertex expansion, and they proved that there is an exact analog of Cheeger’s inequality for hypergraphs:

$$\frac{1}{2}\phi(H) \leq \gamma_2 \leq \sqrt{2\phi(H)},$$

where $\phi(H)$ is the hypergraph edge conductance of H . As in [8], the quantity γ_2 is not polynomial time computable, and a semidefinite programming relaxation similar to that in [34] is used to design a $O(\sqrt{\phi(G)} \log r)$ -approximation algorithm for hypergraph edge conductance where r is the maximum size of a hyperedge. Using this spectral theory, they prove an analog of higher-order Cheeger inequality for hypergraph edge conductance, and also an approximation algorithm for small-set hypergraph edge conductance. Through a reduction from vertex expansion to hypergraph edge conductance, they obtain an analog of higher-order Cheeger inequality for vertex expansion as mentioned earlier after Theorem I.10 and also an approximation algorithm for small-set vertex expansion. This theory also relates (i) expansion, (ii) eigenvalues and (iii) mixing time, and so the work in [12], [32] is closest to the current work.

Compared to the theory in [12], [32] for hypergraphs and for vertex expansion through reduction, we note that the current approach using reweighted eigenvalues is more direct and effective for vertex expansion. The reduction in [12, Fact 3] from vertex expansion $\psi(G)$ of graph G with maximum degree d_{\max} and minimum degree d_{\min} to edge conductance $\phi(H)$ only satisfies

$$d_{\min} \cdot \phi(H) \leq \psi(G) \leq d_{\max} \cdot \phi(H),$$

and so the approximation ratio depends on the ratio between the maximum degree and the minimum degree. The current approach using reweighted eigenvalues does not have this dependency and also proves stronger bounds in k -way vertex expansion as discussed after Theorem I.10. Also, the definitions of the hypergraph diffusion process and its eigenvalues are quite technically involved and require considerable effort to make rigorous [13]. We believe that the definitions of

reweighted eigenvalues are more intuitive and more closely related to ordinary eigenvalues. Also, reweighted eigenvalues have close connections to other important problems such as fastest mixing time and the reweighting conjectures in approximation algorithms.

C. Techniques

Technically, the advantage of relating reweighted eigenvalues to vertex expansions is that many ideas relating eigenvalues to edge conductances can be carried over to the new setting. So, many steps in our proofs are natural extensions of previous arguments, and we focus our discussion here on the new elements.

Vertex Expansion: The proof of Theorem I.3 by Olesker-Taylor and Zanetti is based on the dual characterization of Definition I.1 in Proposition III.1, due to Roch [37], and it has two main steps. In the first step, they used the Johnson-Lindenstrass lemma to project the SDP solution into a $O(\log |V|)$ -dimension solution, and then further reduce it to a 1-dimensional “spectral” solution by taking the best coordinate. This is the step where the $\log |V|$ factor is lost. In the second step, they introduced an interesting new concept called the “matching conductance”, and used some combinatorial arguments about greedy matchings for the analysis of Cheeger rounding on Roch’s dual program.

In our proof of Theorem I.4, we also use Roch’s dual characterization and follow the same two steps. In the first step, we use the Gaussian projection method in [34] to reduce the SDP solution to a 1-dimensional solution directly, and adapt their analysis to show that only a factor of $\log d$ is lost. In the second step, we bypass the concept of matching conductance and do a more traditional analysis of Cheeger rounding as in Bobkov, Houdré and Tetali [8]. It turns out that this analysis works smoothly for weighted vertex conductance, while the approach using matching conductance faced some difficulty as described in [35]. A new element in our proof is the introduction of an intermediate dual program using graph orientation, which is important in the analysis of both steps. In section III, we will review the background from [8], [34], [35], [37] and give a more detailed comparison and overview.

Bipartite Vertex Expansion: The proof of Theorem I.7 for bipartite vertex expansion follows closely the proof of Theorem I.4 and Trevisan’s result [42], once the correct formulation in Definition I.6 is found.

Multiway Vertex Expansion: For the proof of higher-order Cheeger inequality for vertex expansion in Theorem I.10, one technical issue is that we do not know of a convex relaxation for the maximum reweighted k -th smallest eigenvalue in Definition I.9. Instead, we define a related quantity $\sigma_k^*(G)$ called the maximum reweighted sum of the k smallest eigenvalues, which can be written as a semidefinite program. We show that this quantity has a nice dual characterization that satisfies the sub-isotropy condition. This allows us to adapt the techniques in [29] to decompose the SDP solution into k disjointly supported SDP solutions with small objective values, so that

we can apply Theorem I.4 to find k disjoint sets with small vertex expansion.

Improved Cheeger Inequality: The proof of improved Cheeger inequality for vertex expansion is similar to that in [27], which has two main steps. The first step is to prove that if the 1-dimensional solution to Roch’s dual program is close to a k -step function, then Cheeger rounding performs well. The second step is to prove that if the 1-dimensional solution to Roch’s dual program is far from a k -step function, then we can construct an SDP solution to σ_k^* with small objective value, which proves that λ_k^* is small. Therefore, if λ_k^* is large, then the 1-dimensional solution must be close to a k -step function, and hence Cheeger rounding performs well. One interesting aspect in this proof is to relate the performance of a rounding algorithm of one SDP (in this case $\lambda_2^*(G)$) to the objective value of another SDP (in this case $\sigma_k^*(G)$).

Vertex Expansion of 0/1-Polytopes: The examples in Theorem I.12 for 0/1-polytope is obtained by a simple probabilistic construction. The graph of a 0/1-polytope is defined by the set of points chosen in $\{0, 1\}^n$. Let L be the set of points with k ones, and let R be the set of points with $(n - k)$ ones. We prove that if we choose a random set M of points with $n/2$ ones and set $|M| \asymp 4^k n^2$, then with high probability there are no edges between L and R in the resulting polytope, and so M is a small vertex separator of L and R where each has $\binom{n}{k}$ points. The proof is by elementary geometric arguments about the edges of a polytope, and a simple result bounding the number of linear threshold functions in the boolean hypercube $\{0, 1\}^n$.

D. Concurrent Work

Jain, Pham, and Vuong [24] independently published a proof of Theorem I.4 for the uniform distribution case. Their approach is based on a better analysis of dimension reduction for maximum matching, which is quite different from our approach as we bypassed the concept of matching conductance in [35].

E. Full Version of the Paper

This is an abridged version of the paper. The full version of the paper can be found at arXiv:2203.06168.

II. PRELIMINARIES

Notations: Given two functions f, g , we use $f \lesssim g$ to denote the existence of a positive constant $c > 0$, such that $f \leq c \cdot g$ always holds. We use $f \asymp g$ to denote $f \lesssim g$ and $g \lesssim f$. For positive integers k , we use $[k]$ to denote the set $\{1, 2, \dots, k\}$. For a function $f : X \rightarrow \mathbb{R}$, $\text{supp}(f)$ denotes the domain subset on which f is nonzero. For an event E , $\mathbb{1}[E]$ denotes the indicator function that is 1 when E is true and 0 otherwise.

Graphs: Let $G = (V, E)$ be an undirected graph. Throughout this paper, we use $n := |V|$ to denote the number of vertices and $m := |E|$ to denote the number of edges in the graph. If uv is an edge in G , we either write $uv \in E$ or use the notation $u \sim v$. The degree of a vertex v , denoted by

$\deg(v)$, is the number of edges incident to v . The maximum degree of a graph is defined as $\max_{v \in V} \deg(v)$. We usually associate G with a probability distribution $\pi : V \rightarrow \mathbb{R}$ on the set of vertices, and we write $\pi(S) := \sum_{v \in S} \pi(v)$ for a subset $S \subseteq V$. We assume without loss that $\pi(u) > 0$ for all $u \in V$.

Let $S \subseteq V$ be a subset of vertices. The edge boundary of S is defined as $\delta(S) := \{uv \in E \mid u \in S, v \notin S\}$. The volume of S is defined as $\text{vol}(S) := \sum_{v \in S} \deg(v)$. The edge conductance of S is defined as $\phi(S) := |\delta(S)|/|S|$. The vertex boundary of S is defined as $\partial S := \{v \in V \setminus S \mid \exists u \in S \text{ with } uv \in E\}$. The π -weighted vertex expansion of S is defined as $\psi(S) := \pi(\partial S)/\pi(S)$, and when π is the uniform distribution $\psi(S) = |\partial S|/|S|$ is the usual vertex expansion. The induced edge set of S is defined as $E[S] := \{uv \in E \mid u \in S \text{ and } v \in S\}$.

Let $G = (V, \vec{E})$ be a directed graph. If uv is a directed edge in G , we either write $uv \in \vec{E}$ or use the notation $u \rightarrow v$. The indegree of a vertex v is defined as $\deg^{\text{in}}(v) := |\{u \in V \mid u \rightarrow v\}|$. We will define a directed analog of vertex expansion in section III. The definition is not standard and hence deferred to the relevant section.

Linear Algebra: Let $M \in \mathbb{R}^{n \times n}$ be a matrix. When M is symmetric, the spectral theorem states that M admits an orthonormal eigendecomposition $M = UDU^{-1}$, where D is a diagonal matrix and U is a unitary matrix such that $U^{-1}U = I_n$ where I_n is the $n \times n$ identity matrix.

Two matrices $M, N \in \mathbb{R}^{n \times n}$ are said to be cospectral if they are both diagonalizable, and their eigenvalues are the same. There are two cases of cospectral matrices that we will use.

Fact II.1. *Let $M, N \in \mathbb{R}^{n \times n}$. Suppose that M is diagonalizable and that M and N are similar (i.e. $M = X^{-1}NX$ for some invertible matrix $X \in \mathbb{R}^{n \times n}$). Then, N is also diagonalizable, and M and N are cospectral.*

Fact II.2. *Let $M, N \in \mathbb{R}^{n \times n}$. Suppose that there exist $A, B \in \mathbb{R}^{n \times n}$ such that $M = AB$ and $N = BA$. If M is diagonalizable, then N is also diagonalizable, and M and N are cospectral.*

Given that M is symmetric, we say that M is positive semidefinite (PSD) if $v^T M v \geq 0$ for all $v \in \mathbb{R}^n$, and we write $M \succcurlyeq 0$. Equivalently, M is PSD if all its eigenvalues are nonnegative. Also equivalently, M is PSD if there exists X such that $M = X^T X$. Let $x_i \in \mathbb{R}^n$ be the i -th column of X . Then M is called the Gram matrix of $x_1, \dots, x_n \in \mathbb{R}^n$ as $M(i, j) = \langle x_i, x_j \rangle$ for all $i, j \in [n]$.

The trace of a matrix $M \in \mathbb{R}^{n \times n}$ is defined as $\text{tr}(M) := \sum_{i=1}^n M(i, i)$. We will often use the fact that $\text{tr}(AB) = \text{tr}(BA)$ for two matrices of compatible dimensions.

Random Walks: Given a finite state space X , a Markov chain on X is represented by a matrix $P \in \mathbb{R}^{X \times X}$, where $P(u, v)$ is the probability of traversing from state u to state v in one step. Thus, P has nonnegative entries and satisfies $\sum_{v \in X} P(u, v) = 1$ for all $u \in X$. A distribution $\pi : X \rightarrow \mathbb{R}$ is said to be a stationary distribution of P if $\pi^T P = \pi^T$.

A transition matrix P is said to be time-reversible with respect to π if $\pi(u)P(u, v) = \pi(v)P(v, u)$ for any $u, v \in X$. Note that this implies that π is a stationary distribution of P . The time reversibility condition can be written as $\Pi P = P^T \Pi$, where $\Pi := \text{diag}(\pi)$. Thus, $\Pi^{1/2} P \Pi^{-1/2}$ is symmetric, hence diagonalizable with eigenvalues $1 = \alpha_1(P) \geq \alpha_2(P) \geq \dots \geq \alpha_n(P) \geq -1$. As P is similar to $\Pi^{1/2} P \Pi^{-1/2}$, they have the same eigenvalues by Fact II.1. The spectral gap of P is defined as $1 - \alpha_2(P)$.

For $\epsilon \in (0, 1)$, we define the ϵ -mixing time $\tau_{\text{mix}}(P, \epsilon)$ of P to be the smallest $t \in \mathbb{N}$ such that $d_{TV}(\pi, \rho) \leq \epsilon$ for any initial distribution ρ . Here, $d_{TV}(\cdot, \cdot)$ is the total variation distance, defined as $d_{TV}(\rho_1, \rho_2) := \max_{S \subseteq V} |\rho_1(S) - \rho_2(S)|$ for any two distributions $\rho_1, \rho_2 : X \rightarrow \mathbb{R}_{\geq 0}$. The relaxation time $\tau_{\text{rel}}(P)$ of P is defined as the reciprocal of the spectral gap, so $\tau_{\text{rel}}(P) := \frac{1}{1 - \alpha_2(P)}$. Let $\pi_{\min} := \min_{u \in V} \pi(u)$. It is known that (see e.g. Chapter 12 of [30])

$$(\tau_{\text{rel}}(P) - 1) \cdot \log \frac{1}{2\epsilon} \leq \tau_{\text{mix}}(P, \epsilon) \leq \tau_{\text{rel}}(P) \cdot \log \frac{1}{\epsilon \cdot \pi_{\min}}.$$

Because of this connection between the spectral gap and the mixing time of P , the optimization problem of maximizing the spectral gap of the random walk matrix is referred to as “fastest mixing time” in [11].

Spectral Graph Theory: Given a graph $G = (V, E)$, its adjacency matrix $A = A(G)$ is a $n \times n$ matrix where the (u, v) -th entry is $\mathbb{1}_{uv \in E}$. The Laplacian matrix is defined as $L := D - A$, where $D := \text{diag}(\{\deg(v)\}_{v \in V})$ is the diagonal degree matrix. For a vector $x \in \mathbb{R}^n$, the Laplacian matrix has a useful quadratic form $x^T L x = \sum_{uv \in E} (x(u) - x(v))^2$.

The normalized adjacency matrix is defined as $\mathcal{A} = D^{-1/2} A D^{-1/2}$, and the normalized Laplacian matrix is defined as $\mathcal{L} := I - \mathcal{A}$. Observe that \mathcal{A} is similar to the simple random walk matrix on G , so it is diagonalizable with eigenvalues $1 = \alpha_1(\mathcal{A}) \geq \alpha_2(\mathcal{A}) \geq \dots \geq \alpha_n(\mathcal{A}) \geq -1$. Therefore, \mathcal{L} is diagonalizable, and its eigenvalues are $0 = \lambda_1(\mathcal{L}) \leq \lambda_2(\mathcal{L}) \leq \dots \leq \lambda_n(\mathcal{L}) \leq 2$. Note that we use α_i to denote the eigenvalues of the normalized adjacency matrix \mathcal{A} and random walk matrix P , and we use λ_i to denote the eigenvalues of the normalized Laplacian matrix \mathcal{L} .

Let $\phi(G) := \min_{S \subseteq V: 0 < \pi(S) \leq 1/2} \phi(S)$ be the edge conductance of the graph G . Cheeger’s inequality [1], [2], [14] states that

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

This theorem is important because it connects (i) the spectral gap of the normalized Laplacian matrix, (ii) the edge conductance of the graph and (iii) the mixing time of random walks.

III. OPTIMAL CHEEGER INEQUALITY FOR VERTEX EXPANSION

The goal of this section is to prove Theorem I.4. We will first review the proofs in [34], [35] in section III-A, and then present how to combine their proofs with a graph orientation idea to prove Theorem I.4 in section III-B.

A. Background

We will first review the proofs by Olesker-Taylor and Zanetti [35] in section III-A1, and then the proofs by Louis, Raghavendra and Vempala in [34] in section III-A2.

In this subsection, the stationary distribution π is assumed to be the uniform distribution. This will slightly simplify the presentation and was also the setting considered in previous works.

1) *Review of [35]:* Recall the fastest mixing time problem formulated in Definition I.1. When π is the uniform distribution, the problem is to find a doubly stochastic reweighted matrix P of G that minimizes the second largest eigenvalue of P .

The starting point is the following dual characterization of the primal program in Definition I.1 obtained by Roch [37], which is stated in the form for a general distribution π that we will use.

Proposition III.1 (Dual Program for Fastest Mixing [35], [37]). *Given an undirected graph $G = (V, E)$ and a probability distribution π on V , the following semidefinite program is dual to the primal program in Definition I.1 with strong duality $\lambda_2^*(G) = \gamma(G)$ where $\gamma(G)$ is defined as*

$$\begin{aligned} \gamma(G) := & \min_{\substack{f: V \rightarrow \mathbb{R}^n \\ g: V \rightarrow \mathbb{R}_{\geq 0}}} \sum_{v \in V} \pi(v) g(v) \\ \text{subject to } & \sum_{v \in V} \pi(v) \|f(v)\|^2 = 1 \\ & \sum_{v \in V} \pi(v) f(v) = \vec{0} \\ & g(u) + g(v) \geq \|f(u) - f(v)\|^2 \quad \forall uv \in E. \end{aligned}$$

We note that this is equivalent to the dual program given in [11], but Roch’s program is written in a vector program form that will be more convenient for rounding.

The proof of Theorem I.3 has two main steps. The first step is to project the above dual program to the following one-dimensional “spectral” program.

Definition III.2 (One-Dimensional Dual Program for Fastest Mixing [35]). *Given an undirected graph $G = (V, E)$ and a probability distribution π on V , $\gamma^{(1)}(G)$ is defined as*

$$\begin{aligned} \gamma^{(1)}(G) := & \min_{\substack{f: V \rightarrow \mathbb{R} \\ g: V \rightarrow \mathbb{R}_{\geq 0}}} \sum_{v \in V} \pi(v) g(v) \\ \text{subject to } & \sum_{v \in V} \pi(v) f(v)^2 = 1 \\ & \sum_{v \in V} \pi(v) f(v) = 0 \\ & g(u) + g(v) \geq (f(u) - f(v))^2 \quad \forall uv \in E. \end{aligned}$$

Olesker-Taylor and Zanetti use the Johnson-Lindenstrauss lemma to first project the solution in Proposition III.1 to $O(\log n)$ dimensions with constant distortion, and then take the best coordinate to obtain a 1-dimensional solution with the following guarantee. Note that this step works for any probability distribution π on V .

Proposition III.3 ([35], Proposition 2.9). *For any undirected graph $G = (V, E)$ and any probability distribution π on V ,*

$$\gamma(G) \leq \gamma^{(1)}(G) \lesssim \log |V| \cdot \gamma(G).$$

In the second step, Olesker-Taylor and Zanetti observed that the dual program in Definition III.2 is similar to the weighted vertex cover problem with edge weights $(f(u) - f(v))^2$ for each edge $uv \in E$, which is equivalent to the fractional matching problem by linear programming duality. To analyze Definition III.2, they introduced an interesting new concept called “matching conductance”, and used some combinatorial arguments about greedy matching as well as some spectral arguments to prove the following Cheeger-type inequality.

Theorem III.4 ([35], Theorem 2.10). *For any undirected graph $G = (V, E)$ and the uniform distribution $\pi = \bar{1}/|V|$,*

$$\psi(G)^2 \lesssim \gamma^{(1)}(G) \lesssim \psi(G).$$

Combining Proposition III.1 and Proposition III.3 and Theorem III.4 gives

$$\psi(G)^2 \lesssim \gamma^{(1)}(G) \lesssim \log |V| \cdot \gamma(G) = \log |V| \cdot \lambda_2^*(G)$$

and

$$\lambda_2^*(G) = \gamma(G) \leq \gamma^{(1)}(G) \lesssim \psi(G),$$

proving Theorem I.3.

Note that the proof of the second step only works when π is the uniform distribution. Olesker-Taylor and Zanetti discussed some difficulty in generalizing their combinatorial arguments to the weighted setting, and left it as an open question to prove Theorem III.4 for any probability distribution π .

2) *Review of [34]:* Our proof is based on the techniques in [34] which we review here. Their algorithm is based on the following “spectral” formulation λ_∞ by Bobkov, Houdré and Tetali [8], which is for the uniform distribution π .

Definition III.5 (λ_∞ in [8]). *Given an undirected graph $G = (V, E)$,*

$$\lambda_\infty(G) := \min_{x: V \rightarrow \mathbb{R}, x \perp \bar{1}} \frac{\sum_{u \in V} \max_{v: (v, u) \in E} (x(u) - x(v))^2}{\sum_{u \in V} x(u)^2}.$$

Bobkov, Houdré and Tetali [8] proved an exact analog of Cheeger’s inequality for symmetric vertex expansion that $\frac{1}{2}\psi_{\text{sym}}(G)^2 \leq \lambda_\infty(G) \leq 2\psi_{\text{sym}}(G)$. We will use some of their arguments to prove a similar statement in Theorem III.15 in section III-B.

The issue is that λ_∞ is not known to be efficiently computable, and indeed recently Farhadi, Louis and Tetali [17] proved that it is NP-hard to compute $\lambda_\infty(G)$ exactly. To design an approximation algorithm for $\psi(G)$, Louis, Raghavendra and Vempala [34] defined the following semidefinite programming relaxation for λ_∞ , which we denote by sdp_∞ .

Definition III.6 (sdp_∞ in [34]). *Given an undirected graph $G = (V, E)$, define the $\text{sdp}_\infty(G)$ program as*

$$\begin{aligned} & \min_{\substack{f: V \rightarrow \mathbb{R}^n \\ g: V \rightarrow \mathbb{R}}} \sum_{v \in V} g(v) \\ \text{subject to } & \sum_{v \in V} \|f(v)\|^2 = 1 \\ & \sum_{v \in V} f(v) = \vec{0} \\ & g(v) \geq \|f(u) - f(v)\|^2 \quad \forall u \in V \text{ with } uv \in E. \end{aligned}$$

The rounding algorithm in [34] is to project the solution to sdp_∞ into a 1-dimensional solution by setting $x(v) = \langle f(v), h \rangle$ where $h \sim N(0, 1)^n$ is a random Gaussian vector. They proved that the 1-dimensional solution is a $O(\log d)$ -approximation to sdp_∞ where d is the maximum degree of the graph.

Theorem III.7 ([34], Lemma 9.6). *For any undirected graph $G = (V, E)$ with maximum degree d ,*

$$\text{sdp}_\infty(G) \leq \lambda_\infty \lesssim \log d \cdot \text{sdp}_\infty(G).$$

For the analysis, they used the following properties of Gaussian random variables, for which we will also use in our proofs and so we state them here. The first fact is for the analysis of the numerator and the second fact is for the analysis of the denominator of λ_∞ .

Fact III.8 ([34], Fact 9.7). *Let Y_1, Y_2, \dots, Y_d be d Gaussian random variables with mean 0 and variance at most σ^2 . Let Y be the random variable defined as $Y := \max\{Y_i \mid i \in [d]\}$. Then*

$$\mathbb{E}[Y] \leq 2\sigma\sqrt{\log d}.$$

Fact III.9 ([34], Lemma 9.8). *Suppose z_1, \dots, z_m are Gaussian random variables (not necessarily independent) such that $\mathbb{E}[\sum_{i=1}^m z_i^2] = 1$. Then*

$$\Pr \left[\sum_{i=1}^m z_i^2 \geq \frac{1}{2} \right] \geq \frac{1}{12}.$$

B. Proof of Theorem I.4

We follow the same two-step plan as in [35]. We will prove in Proposition III.14 in section III-B2 that $\gamma^{(1)}(G) \lesssim \gamma(G) \cdot \log d$ for any probability distribution π . Note that this already improves Theorem I.3 to the optimal bound, when π is the uniform distribution. Then, we will prove in Theorem III.15 in section III-B3 that $\psi(G)^2 \lesssim \gamma^{(1)}(G) \lesssim \psi(G)$ for any probability distribution π on V . As in [35], combining Proposition III.1 and Proposition III.14 and Theorem III.15 gives Theorem I.4.

1) *Dual Program on Graph Orientation:* To extend the techniques in [8], [34] to prove the two steps, we will introduce a “directed” program $\vec{\gamma}(G)$ to bring $\gamma(G)$ in Proposition III.1 closer to $\text{sdp}_\infty(G)$ in Definition III.6.

Observe that the two SDP programs $\gamma(G)$ and $\text{sdp}_\infty(G)$ have very similar form. The only difference is that the last

constraint in Proposition III.1 only requires that $g(u) + g(v) \geq \|f(u) - f(v)\|^2$ for $uv \in E$, while the last constraint in Definition III.6 has a stronger requirement that $\min\{g(u), g(v)\} \geq \|f(u) - f(v)\|^2$ for $uv \in E$. So sdp_∞ is a stronger relaxation than $\gamma(G) = \lambda_2^*(G)$.

Lemma III.10. *For any undirected graph $G = (V, E)$ and any probability distribution π on V ,*

$$\lambda_2^*(G) \leq \text{sdp}_\infty(G).$$

For our analysis of $\lambda_2^*(G)$, we consider the following “directed” program $\vec{\gamma}(G)$ where the last constraint is $\max\{g(u), g(v)\} \geq \|f(u) - f(v)\|^2$ for $uv \in E$. We also state the corresponding 1-dimensional version as in Definition III.2 in the following definition.

Definition III.11 (Directed Dual Programs for $\gamma(G)$). *Given an undirected graph $G = (V, E)$ and a probability distribution π on V ,*

$$\begin{aligned} \vec{\gamma}(G) := & \min_{\substack{f: V \rightarrow \mathbb{R}^n \\ g: V \rightarrow \mathbb{R}_{\geq 0}}} \sum_{v \in V} \pi(v) g(v) \\ \text{subject to } & \sum_{v \in V} \pi(v) \|f(v)\|^2 = 1 \\ & \sum_{v \in V} \pi(v) f(v) = \vec{0} \\ & \max\{g(u), g(v)\} \geq \|f(u) - f(v)\|^2 \\ & \quad \forall uv \in E. \end{aligned}$$

$\vec{\gamma}^{(1)}(G)$ is defined as the 1-dimensional program of $\vec{\gamma}(G)$ where $f: V \rightarrow \mathbb{R}$ instead of $f: V \rightarrow \mathbb{R}^n$.

Note that $\vec{\gamma}(G)$ is not a semidefinite program because of the max constraint, but $\gamma(G)$ and $\vec{\gamma}(G)$ are closely related and $\vec{\gamma}(G)$ is only used in the analysis as a proxy for $\gamma(G)$.

Lemma III.12. *For any undirected graph $G = (V, E)$ and any probability distribution π on V ,*

$$\gamma(G) \leq \vec{\gamma}(G) \leq 2\gamma(G)$$

and

$$\gamma^{(1)}(G) \leq \vec{\gamma}^{(1)}(G) \leq 2\gamma^{(1)}(G).$$

Proof. As $g \geq 0$, any feasible solution f, g to $\vec{\gamma}(G)$ is a feasible solution to $\gamma(G)$ and so the first inequalities follow. On the other hand, for any feasible solution f, g to $\gamma(G)$, note that $f, 2g$ is a feasible solution to $\vec{\gamma}(G)$ and so the second inequalities follow. \square

The reason that we call $\vec{\gamma}(G)$ the “directed” program is as follows. For each edge $uv \in E$, the constraint in $\text{sdp}_\infty(G)$ requires both $g(u)$ and $g(v)$ to be at least $\|f(u) - f(v)\|^2$, while the constraint in $\vec{\gamma}(G)$ only requires at least one of $g(u)$ or $g(v)$ to be at least $\|f(u) - f(v)\|^2$. We think of $\vec{\gamma}(G)$ as assigning a direction to each edge and requiring that $g(v) \geq \|f(u) - f(v)\|^2$ for each directed edge $u \rightarrow v$. Then, we can rewrite the programs $\vec{\gamma}(G)$ and $\vec{\gamma}^{(1)}(G)$ by eliminating the

variables $g(v)$ for $v \in V$, by minimizing over all possible orientations of the edge set E .

Lemma III.13 (Directed Dual Programs Using Orientation for $\gamma(G)$). *Let $G = (V, E)$ be an undirected graph and π be a probability distribution on V . Let \vec{E} be an orientation of the undirected edges in E . Then*

$$\begin{aligned} \vec{\gamma}(G) = & \min_{f: V \rightarrow \mathbb{R}^n} \min_{\vec{E}} \sum_{v \in V} \pi(v) \max_{u: uv \in \vec{E}} \|f(u) - f(v)\|^2 \\ \text{subject to } & \sum_{v \in V} \pi(v) \|f(v)\|^2 = 1 \\ & \sum_{v \in V} \pi(v) f(v) = \vec{0}. \end{aligned}$$

Similarly, $\vec{\gamma}^{(1)}(G)$ can be written in the same form with $f: V \rightarrow \mathbb{R}$ instead of $f: V \rightarrow \mathbb{R}^n$.

Proof. In one direction, given an orientation \vec{E} , we can define $g(v) := \max_{u: uv \in \vec{E}} \|f(u) - f(v)\|^2$, so that f, g is a feasible solution to $\vec{\gamma}(G)$ as stated in Definition III.11 with the same objective value.

In the other direction, given a solution f, g in Definition III.11, we can define an orientation \vec{E} of E so that each directed edge uv satisfies $g(v) \geq \|f(u) - f(v)\|^2$. Note that $g(v) \geq \max_{u: uv \in \vec{E}} \|f(u) - f(v)\|^2$, and setting it to be an equality would satisfy all the constraints and not increase the objective value as $g \geq 0$. \square

This formulation will be useful in both the Gaussian projection step for Proposition III.14 and the threshold rounding step for Theorem III.15.

2) *Gaussian Projection:* The following proposition is an improvement of Proposition III.3 in [35]. The formulation in Lemma III.13 allows us to use the expected maximum of Gaussian random variables in Fact III.8 to analyze the projection as was done in [34].

Proposition III.14 (Gaussian Projection for $\gamma(G)$). *For any undirected graph $G = (V, E)$ with maximum degree d and any probability distribution π on V ,*

$$\gamma(G) \leq \gamma^{(1)}(G) \lesssim \gamma(G) \cdot \log d.$$

Proof. We will prove that $\vec{\gamma}(G) \leq \vec{\gamma}^{(1)}(G) \lesssim \log d \cdot \vec{\gamma}(G)$, and the proposition will follow from Lemma III.12. The first inequality is immediate as $\vec{\gamma}^{(1)}(G)$ is a restriction of $\vec{\gamma}(G)$, so we focus on proving the second inequality.

Let $f: V \rightarrow \mathbb{R}^n$ and \vec{E} be a solution to $\vec{\gamma}(G)$ as stated in Lemma III.13. As in [34], we construct a 1-dimensional solution $y \in \mathbb{R}^n$ to $\vec{\gamma}^{(1)}(G)$ by setting $y(v) = \langle f(v), h \rangle$, where $h \sim N(0, 1)^n$ is a Gaussian random vector with independent entries.

First, consider the expected objective value of y to $\vec{\gamma}^{(1)}(G)$. For each max term in the summand,

$$\begin{aligned} \mathbb{E} \left[\max_{u: u \rightarrow v} (y(u) - y(v))^2 \right] &= \mathbb{E} \left[\max_{u: u \rightarrow v} \langle f(u) - f(v), h \rangle^2 \right] \\ &\leq 2 \max_{u: u \rightarrow v} \|f(u) - f(v)\|^2 \cdot \log d, \end{aligned}$$

where the last inequality is by applying Fact III.8 on normal random variable $\langle f(u) - f(v), h \rangle$ with variance $\|f(u) - f(v)\|^2$ for each of the at most d terms. By linearity of expectation, the expected objective value of $\vec{\gamma}^{(1)}(G)$ is

$$\begin{aligned} & \mathbb{E} \left[\sum_{v \in V} \pi(v) \max_{u: u \rightarrow v} (y(u) - y(v))^2 \right] \\ & \leq 2 \log d \cdot \sum_{v \in V} \pi(v) \max_{u: u \rightarrow v} \|f(u) - f(v)\|^2 \\ & = 2 \log d \cdot \vec{\gamma}(G). \end{aligned}$$

Therefore, by Markov's inequality,

$$\Pr \left[\sum_{v \in V} \pi(v) \max_{u: u \rightarrow v} (y(u) - y(v))^2 \geq 48 \log d \cdot \vec{\gamma}(G) \right] \leq \frac{1}{24}.$$

Next, by applying Fact III.9 with $z_v = \sqrt{\pi(v)} \cdot y(v)$, it follows that

$$\begin{aligned} & \mathbb{E} \left[\sum_{v \in V} \pi(v) y(v)^2 \right] = \sum_{v \in V} \pi(v) \|f(v)\|^2 = 1 \\ \implies & \Pr \left[\sum_{v \in V} \pi(v) y(v)^2 \geq \frac{1}{2} \right] \geq \frac{1}{12}. \end{aligned}$$

Finally, since $\sum_{v \in V} \pi(v) f(v) = \vec{0}$, it holds that

$$\sum_{v \in V} \pi(v) y(v) = \sum_{v \in V} \pi(v) \langle f(v), h \rangle = \left\langle \sum_{v \in V} \pi(v) f(v), h \right\rangle = 0.$$

Therefore, with probability at least $\frac{1}{24}$, all of these events hold simultaneously. The second event

$$\sum_{v \in V} \pi(v) y(v)^2 \geq \frac{1}{2}$$

means that we can rescale y by a factor of at most $\sqrt{2}$, so that the constraint $\sum_{v \in V} \pi(v) y(v)^2 = 1$ is satisfied and the objective value is at most $96 \log d \cdot \vec{\gamma}(G)$. Hence we conclude that $\vec{\gamma}^{(1)}(G) \lesssim \vec{\gamma}(G) \cdot \log d$. \square

3) *Cheeger Rounding for Vertex Expansion:* We generalize Theorem III.4 to weighted vertex expansion. Our proof does not use the concept of matching conductance in [35], rather it is based on a more traditional analysis as in [8] using the directed program $\vec{\gamma}^{(1)}(G)$ in Lemma III.13.

Theorem III.15 (Cheeger Inequality for Weighted Vertex Expansion). *For any undirected graph $G = (V, E)$ and any probability distribution π on V ,*

$$\psi(G)^2 \lesssim \gamma^{(1)}(G) \lesssim \psi(G).$$

The organization is as follows. The proof of the easy direction is omitted and can be found in the full version of the paper. For the hard direction, we will work on $\vec{\gamma}^{(1)}(G)$ instead. First we do the standard preprocessing step to truncate the solution to have π -weight at most $1/2$. Then the main step is to define a modified vertex boundary condition for directed graphs and use it for the analysis of the standard threshold

rounding. Finally we clean up the solution obtained from threshold rounding to find a set with small vertex expansion in the underlying undirected graph.

Lemma III.16 (Easy Direction). *For any undirected graph $G = (V, E)$ and any probability distribution π on V ,*

$$\gamma^{(1)}(G) \leq 2\psi(G).$$

We now turn to proving the hard direction. Given a solution $y : V \rightarrow \mathbb{R}$ to $\vec{\gamma}^{(1)}(G)$ in Lemma III.13 satisfying $y \perp \pi$, we do the standard preprocessing step to truncate y to obtain a non-negative solution x with $\pi(\text{supp}(x)) \leq 1/2$ and comparable objective value. Note that we no longer require that $x \perp \pi$. The proof of the following lemma is standard and can be found in the full version of the paper.

Lemma III.17 (Truncation). *Let $G = (V, E)$ be an undirected graph and π be a probability distribution on V . Given a solution y and \vec{E} to $\vec{\gamma}^{(1)}(G)$ as stated in Lemma III.13, there is a solution x and \vec{E} with $x \geq 0$ and $\pi(\text{supp}(x)) \leq 1/2$ and*

$$\frac{\sum_{v \in V} \pi(v) \max_{u: uv \in \vec{E}} (x(u) - x(v))^2}{\sum_{v \in V} \pi(v) x(v)^2} \leq 4 \vec{\gamma}^{(1)}(G).$$

For the standard threshold rounding, we define the appropriate vertex boundary $\vec{\partial} S$ for the analysis of the directed program $\vec{\gamma}^{(1)}(G)$. Note that, unlike ∂S , $\vec{\partial} S$ may contain vertices in S . A good interpretation is to think of $\vec{\partial} S$ as a vertex cover of the edge boundary $\delta(S)$ in the undirected sense.

Definition III.18 (Directed Vertex Boundary and Expansion). *Let $G = (V, \vec{E})$ be a directed graph. For $S \subseteq V$, define the directed vertex boundary and the directed vertex expansion as*

$$\begin{aligned} \vec{\partial} S &:= \{v \in S \mid \exists u \notin S \text{ with } uv \in \vec{E}\} \\ &\cup \{v \notin S \mid \exists u \in S \text{ with } uv \in \vec{E}\} \end{aligned}$$

and

$$\vec{\psi}(S) := \frac{\pi(\vec{\partial} S)}{\pi(S)}.$$

The main step is to prove that the standard threshold rounding will find a set S with small directed vertex expansion $\vec{\psi}(S)$.

Proposition III.19 (Threshold Rounding for $\gamma(G)$). *Let $G = (V, E)$ be an undirected graph and π be a probability distribution on V . Given a solution x and \vec{E} with $x \geq 0$ and*

$$\frac{\sum_{v \in V} \pi(v) \max_{u: u \rightarrow v} (x(u) - x(v))^2}{\sum_{v \in V} \pi(v) x(v)^2} \leq \gamma(G),$$

there is a set $S \subseteq \text{supp}(x)$ with $\vec{\psi}(S) \lesssim \sqrt{\gamma(G)}$.

Proof. For any $t \geq 0$, define $S_t := \{v \in V \mid x(v)^2 > t\}$. By a standard averaging argument,

$$\min_t \vec{\psi}(S_t) \leq \frac{\int_0^\infty \pi(\vec{\partial} S_t) dt}{\int_0^\infty \pi(S_t) dt}.$$

The denominator is

$$\int_0^\infty \pi(S_t) dt = \sum_{v \in V} \pi(v) \int_0^\infty \mathbb{1}[v \in S_t] dt = \sum_{v \in V} \pi(v) x(v)^2.$$

For the numerator, note that a vertex v is in $\vec{\partial} S_t$ if and only if $\min\{x(u)^2 \mid uv \in \vec{E}\} \leq t \leq \max\{x(u)^2 \mid uv \in \vec{E}\}$, where we recall the assumption that every vertex has a self loop, and so $vv \in \vec{E}$ and thus $\min\{x(u)^2 \mid uv \in \vec{E}\} \leq x(v)^2 \leq \max\{x(u)^2 \mid uv \in \vec{E}\}$. Hence the numerator is

$$\begin{aligned} & \int_0^\infty \pi(\vec{\partial} S_t) dt \\ &= \sum_{v \in V} \pi(v) \cdot \int_0^\infty \mathbb{1}[v \in \vec{\partial} S_t] dt \\ &= \sum_{v \in V} \pi(v) \cdot \int_0^\infty \mathbb{1}\left[\min\{x(u)^2 \mid uv \in \vec{E}\} \leq t \wedge \right. \\ & \quad \left. t \leq \max\{x(u)^2 \mid uv \in \vec{E}\}\right] dt \\ &= \sum_{v \in V} \pi(v) \left[\max_{\substack{u:u \rightarrow v \\ x(u) > x(v)}} \{x(u)^2 - x(v)^2\} \right. \\ & \quad \left. + \max_{\substack{u:u \rightarrow v \\ x(u) < x(v)}} \{x(v)^2 - x(u)^2\} \right] \\ &\leq 2 \sum_{v \in V} \pi(v) \left[\max_{u:u \rightarrow v} \{|x(u)^2 - x(v)^2|\} \right] \\ &\leq 2 \sum_{v \in V} \pi(v) \left[\max_{u:u \rightarrow v} \{(x(u) - x(v))^2\} \right. \\ & \quad \left. + 2x(v) \cdot |x(u) - x(v)| \right] \\ &\leq 2 \sum_{v \in V} \pi(v) \max_{u:u \rightarrow v} (x(u) - x(v))^2 \\ & \quad + 4 \sqrt{\sum_{v \in V} \pi(v) x(v)^2 \cdot \sum_{v \in V} \pi(v) \max_{u:u \rightarrow v} (x(u) - x(v))^2}, \end{aligned}$$

where the second-last inequality is by $|x(u)^2 - x(v)^2| \leq |x(u) - x(v)| \cdot (|x(u) - x(v)| + 2|x(v)|)$, and the last inequality is by the Cauchy-Schwarz inequality.

Combining the numerator and the denominator bounds,

$$\begin{aligned} & \frac{\int_0^\infty \pi(\vec{\partial} S_t) dt}{\int_0^\infty \pi(S_t) dt} \\ &\leq \frac{\sum_{v \in V} \pi(v) \max_{u:u \rightarrow v} (x(u) - x(v))^2}{\sum_{v \in V} \pi(v) x(v)^2} \\ & \quad + 2 \sqrt{\frac{\sum_{v \in V} \pi(v) \max_{u:u \rightarrow v} (x(u) - x(v))^2}{\sum_{v \in V} \pi(v) x(v)^2}} \\ &= \gamma + 2\sqrt{\gamma} \lesssim \sqrt{\gamma}, \end{aligned}$$

where the last inequality is by $\gamma \leq 2$ as was shown in the proof of the easy direction in Lemma III.16. Therefore, $\min_t \vec{\psi}(S_t) \lesssim \sqrt{\gamma}$ and $S_t \subseteq \text{supp}(x)$ by construction. \square

Finally, given a set S with small directed vertex expansion $\vec{\psi}(S)$, we show how to find a set $S' \subseteq S$ with small

vertex expansion $\psi(S')$. This step is similar to the step in [35, Proposition 2.2] from matching conductance to vertex expansion.

Lemma III.20 (Postprocessing for Vertex Expansion). *Let $G = (V, \vec{E})$ be a directed graph. Given a set S with $\vec{\psi}(S) < 1/2$, there is a set $S' \subseteq S$ with $\psi(S') \leq 2\vec{\psi}(S)$ in the underlying undirected graph of G .*

Proof. From Definition III.18, the observation is that all undirected edges in $\delta(S)$ are incident to at least one vertex in $\vec{\partial} S$. Define $S' := S - \vec{\partial} S$. Then observe that $\partial S' \subseteq \vec{\partial} S$, as there are no incoming edges to S' from $V - (S' \cup \vec{\partial} S)$ and all outgoing edges from S' go to $\vec{\partial} S$. This implies that

$$\pi(\partial S') \leq \pi(\vec{\partial} S) = \vec{\psi}(S) \cdot \pi(S) \leq 2\vec{\psi}(S) \cdot \pi(S'),$$

where the last inequality uses the assumption that $\vec{\psi}(S) = \pi(\vec{\partial} S)/\pi(S) < 1/2$ and so $\pi(S') \geq \pi(S) - \pi(\vec{\partial} S) \geq \pi(S)/2$. We conclude that $\psi(S') \leq 2\vec{\psi}(S)$. \square

We now have all the necessary components to complete the proof of Theorem III.15. To see how to put together the steps, refer to the full version of the paper.

IV. CONCLUDING REMARKS

We present a new spectral theory which relates (i) reweighted eigenvalues, (ii) vertex expansion and (iii) fastest mixing time. This is analogous to the classical spectral theory which relates (i) eigenvalues, (ii) edge conductance and (iii) mixing time. This spectral approach for vertex expansion has the advantage that most existing results and proofs for edge conductances and eigenvalues have a close analog for vertex expansion and reweighted eigenvalues with almost tight bounds. We do not intend to be exhaustive in this paper, and we fully expect that other results relating eigenvalues and edge conductances also have an analog for vertex expansion using reweighted eigenvalues.

To conclude, we believe that our work provides an interesting spectral theory for vertex expansion, as the formulations have the natural interpretation as reweighted eigenvalues and also have close connections to other important problems such as fastest mixing time and the reweighting conjectures in approximation algorithms. We also believe that this approach can be extended further for hypergraph edge expansion.

ACKNOWLEDGMENT

We thank Shayan Oveis Gharan for suggesting the tight example for Theorem I.4, Robert Wang for suggesting the connection to the 0/1-polytope expansion conjecture, and Sam Olesker-Taylor for providing insightful comments that improved the presentation of the paper.

REFERENCES

- [1] Noga Alon. *Eigenvalues and expanders*. *Combinatorica*, 6, 83–96, 1986.
- [2] Noga Alon, Vitali Milman. λ_1 , *isoperimetric inequalities for graphs, and superconcentrators*. *Journal of Combinatorial Theory, Series B*, 38(1), 73–88, 1985.

- [3] Nima Anari, Kuikui Liu, Shayan Oveis Gharan, Cynthia Vinzant. *Log-concave polynomials II: high-dimensional walks and an FPRAS for counting bases of a matroid*. In Proceedings of the 51th Annual Symposium on Theory of Computing (STOC), 1–12, 2019.
- [4] Sanjeev Arora, Boaz Barak, David Steurer. *Subexponential algorithms for unique games and related problems*. In Proceedings of the 51st Annual IEEE Symposium on Foundations of Computer Science (FOCS), 563–572, 2010.
- [5] Sanjeev Arora, Rong Ge. *New tools for graph coloring*. In Proceedings of the 14th International Workshop on Approximation Algorithms for Combinatorial Optimization (APPROX), 1–12, 2011.
- [6] Sanjeev Arora, Satish Rao, Umesh Vazirani. *Expander flows, geometric embeddings and graph partitioning*. Journal of the ACM, 56(2):1–37, 2009.
- [7] Avrim Blum. *New approximation algorithms for graph coloring*. Journal of the ACM, 41:470–516, 1994.
- [8] S. Bobkov, C. Houdré, Prasad Tetali. λ_∞ , vertex isoperimetry and concentration. *Combinatorica* 20(2):153–172, 2000.
- [9] Stephen Boyd, Persi Diaconis, Pablo Parrilo, Lin Xiao. *Fastest mixing time on graphs with symmetries*. *SIAM Journal on Optimization*, 20(2):792–819, 2009.
- [10] Stephen Boyd, Persi Diaconis, J. Sun, Lin Xiao. *Fastest mixing Markov chain on a path*. *American Mathematical Monthly*, 113(1):70–74, 2006.
- [11] Stephen Boyd, Persi Diaconis, Lin Xiao. *Fastest mixing markov chain on a graph*. *SIAM review*, 46(4):667–689, 2004.
- [12] T.-H. Hubert Chan, Anand Louis, Zhihao Gavin Tang, Chenzi Zhang. *Spectral properties of hypergraph Laplacian and approximation algorithms*. *Journal of the ACM*, 65(3):1–48, 2018.
- [13] T.-H. Hubert Chan, Zhihao Gavin Tang, Xiaowei Wu, Chenzi Zhang. *Diffusion operator and spectral analysis for directed hypergraph Laplacian*. *CoRR abs/1711.01560*, 2017.
- [14] Jeff Cheeger. *A lower bound for the smallest eigenvalue of the Laplacian*. *Problems in Analysis*, Princeton University Press, 195–199, 1970.
- [15] O. Cihan, M. Akar. *Fastest mixing reversible Markov chains on graphs with degree proportional stationary distributions*. *Transactions on Automatic Control*, 60(1):227–232, 2015.
- [16] Thomas M. Cover. *Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition*. *IEEE Transactions on Electronic Computers*, 3:326–334, 1965.
- [17] Majid Farhadi, Anand Louis, Prasad Tetali. *On the complexity of λ_∞ , vertex expansion, and spread constant of trees*. In arXiv:2003.05582, 2020.
- [18] Tomás Feder, Milena Mihail. *Balanced matroids*. In Proceedings of the 24th Annual Symposium on Theory of Computing (STOC), 26–38, 1992.
- [19] Uriel Feige, MohammadTaghi Hajiaghayi, James R. Lee. *Improved approximation algorithms for minimum weight vertex separators*. *SIAM Journal on Computing*, 38(2):629–657, 2008.
- [20] J.A. Fill, Jeff Kahn. *Comparison inequalities and fastest mixing Markov chains*. *Annals of Applied Probability*, 23(5):1778–1816, 2013.
- [21] Bernd Gärtner and Jiri Matousek. *Approximation algorithms and semidefinite programming*. Springer, 2012.
- [22] Rafael Gillmann. *0/1-polytopes: typical and extremal properties*. Ph.D. thesis, Technische Universität Berlin, 2007.
- [23] Anupam Gupta, Robert Krauthgamer, James R. Lee. *Bounded geometries, fractals, and low-distortion embeddings*. In Proceedings of the 44th Annual Symposium on Foundations of Computer Science (FOCS), 534–543, 2003.
- [24] Vishesh Jain, Huy Tuan Pham, and Thuy-Duong Vuong. *Dimension reduction for maximum matchings and the Fastest Mixing Markov Chain*. In arXiv:2203.03858, 2022.
- [25] Volker Kaibel. *On the expansion of graphs of 0/1-polytopes. The Sharpest Cut*. *MPS-SAIM series on optimization* 4, 199–216, 2004.
- [26] Volker Kaibel, Anja Remshagen. *On the graph-density of random 0/1-polytopes*. In Proceedings of the 6th International Workshop on Randomization and Approximation Techniques in Computer Science (RANDOM), 318–328, 2003.
- [27] Tsz Chiu Kwok, Lap Chi Lau, Yin Tat Lee, Shayan Oveis Gharan, Luca 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 (STOC), 11–20, 2013.
- [28] Beatrice Laurent, Pascal Massart. *Adaptive estimation of a quadratic functional by model selection*. *Annals of Statistics*, 1302–1338, 2000.
- [29] James R. Lee, Shayan Oveis Gharan, Luca Trevisan. *Multi-way spectral partitioning and higher-order Cheeger inequalities*. In Proceedings of the 44th Annual Symposium on Theory of Computing (STOC), 1117–1130, 2012.
- [30] David Levin and Yuval Peres. *Markov chains and mixing times*. American Mathematical Society, Volume 107, 2017.
- [31] Shengqiao Li. *Concise formulas for the area and volume of a hyper-spherical cap*. *Asian Journal of Mathematics and Statistics*, Volume 4, No. 1, 66–70, 2011.
- [32] Anand Louis. *Hypergraph Markov operators, eigenvalues and approximation algorithms*. In Proceedings of the 47th Annual Symposium on Theory of Computing (STOC), 713–722, 2015.
- [33] Anand Louis, Prasad Raghavendra, Prasad Tetali, Santosh Vempala. *Many sparse cuts via higher eigenvalues*. In Proceedings of the 44th Annual ACM Symposium on Theory of Computing (STOC), 1131–1140, 2012.
- [34] Anand Louis, Prasad Raghavendra, Santosh Vempala. *The complexity of approximating vertex expansion*. In Proceedings of the 54th IEEE Annual Symposium on Foundations of Computer Science (FOCS), 360–369, 2013.
- [35] Sam Olesker-Taylor, Luca Zanetti. *Geometric bounds on the fastest mixing Markov chain*. In the 13th Innovations in Theoretical Computer Science Conference (ITCS 2022).
- [36] Prasad Raghavendra, David Steurer. *Graph expansion and the unique games conjecture*. In Proceedings of the 42nd Annual ACM Symposium on Theory of Computing (STOC), 755–764, 2010.
- [37] Sébastien Roch. *Bounding fastest mixing*. *Electron. Commun. Probab.* 10:282–296, 2005.
- [38] Stephen Simons. *Minimax theorems and their proofs*. *Minimax and applications*, 1–23, 1995.
- [39] David Steurer. *On the complexity of unique games and graph expansion*. Ph.D. thesis, Princeton University, 2010.
- [40] R.M. Tanner. *Explicit construction of concentrators from generalized N -gons*. *SIAM J. Algebraic Discrete Methods* 5: 287–294, 1984.
- [41] Tomasz Tkocz. *An upper bound for spherical caps*. *The American Mathematical Monthly*, Volume 119, No. 7, 606–607, 2012.
- [42] Luca Trevisan. *Max cut and the smallest eigenvalue*. In Proceedings of the 41st Annual ACM Symposium on Theory of Computing (STOC), 263–272, 2009.