Multi-way dual Cheeger constants and spectral bounds of graphs

CrossMark

Shiping Liu

*Department of Mathematical Sciences, Durham University, DH1 3LE Durham,
United Kingdom*

A R T I C L E   I N F O

A B S T R A C T

We introduce a set of multi-way dual Cheeger constants and prove universal higher-order dual Cheeger inequalities for eigenvalues of normalized Laplace operators on weighted finite graphs. Our proof proposes a new spectral clustering phenomenon deduced from metrics on real projective spaces. We further extend those results to a general reversible Markov operator and find applications in characterizing its essential spectrum.

1. Introduction and main ideas

Cheeger constant, encoding global connectivity properties of the underlying space, was invented by Cheeger [7] and related to the first non-zero eigenvalue of the Laplace–Beltrami operator on a compact Riemannian manifold, which is now well-known as Cheeger inequality. Afterwards, it was extended to discrete settings by several authors in spectral graph theory or Markov chain theory, see e.g. [17,2,1,24,31,37,16,8]. Intriguingly, this stimulated research in many unexpected theoretical or practical areas, such as the explicit construction of expander graphs, see e.g. [1,28,38,34], graph coloring, image

E-mail address: shiping.liu@durham.ac.uk.

segmentation, web search, approximate counting, for which we refer to [26,23] for detailed references.

Recently, Miclo [29] (see also [14]) introduced a set of multi-way Cheeger constants (alternatively called higher-order isoperimetric constants), $h(k)$, $k = 1, 2, \ldots$, in discrete setting and conjectured related higher-order Cheeger inequalities universal for any weighted graph. This conjecture was solved by Lee, Oveis Gharan and Trevisan [26] by bringing in the powerful tool of random metric partitions developed originally in theoretical computer science. Moreover, their approach justifies the empirical spectral clustering algorithms in [33] which are very popular and powerful tools in many practical fields (see e.g. [40,35]). Amazingly, this new progress of spectral graph theory provides feedback to the setting of Riemannian manifold. Funano [18] and Miclo [30] (with different strategies) extended the higher-order Cheeger inequalities to weighted Riemannian manifolds and found very important applications there.

In contrast to a Riemannian manifold, on a graph, the spectrum of the normalized Laplace operator is bounded from above by 2. Explicitly, one can list them as

$$0 = \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_N \leq 2, \tag{1}$$

where N is the size of the graph G. Therefore, a graph has its own particular spectral gaps $2 - \lambda_k$ which have no counterparts in the Riemannian setting. In order to investigate the spectral gap $2 - \lambda_N$, Bauer and Jost [4] introduced a dual Cheeger constant, $\bar{h}(1)$ in our notation below, encoding the bipartiteness property of the underlying graph. (Independently, the constant $1 - \bar{h}(1)$ under the name bipartiteness ratio was studied by Trevisan [39].) Explicitly, it holds that

$$\text{connected } G \text{ is bipartite} \quad \Leftrightarrow \quad \bar{h}(1) = 1. \tag{2}$$

(A graph is bipartite if its vertex set can be divided into two classes and edges are only permitted between two vertices from opposite classes.) They then proved a dual Cheeger inequality, providing a strong quantitative version of the fact that $2 - \lambda_N$ vanishes if and only if the underlying graph is bipartite. This has already found important applications for the convergence of random walks on graphs, synchronization for coupled map lattices [4] and characterizing behaviors of the essential spectrum of infinite graphs [3].

In this paper, we introduce a set of multi-way dual Cheeger constants, $\bar{h}(k)$, $k = 1, 2, \ldots, N$, encoding more detailed information about how far/close a graph is from being a bipartite one. The dual relations between $\bar{h}(k)$ and $h(k)$ are manifested by the fact that

$$G \text{ is bipartite} \quad \Leftrightarrow \quad h(k) + \bar{h}(k) = 1, \quad 1 \leq k \leq N. \tag{3}$$

In fact, if a graph can satisfy $h(k) + \bar{h}(k) = 1$ for a small number k, then, roughly speaking, it actually has a large size bipartite subgraph in a reasonable sense (Proposition 3.1(iii)).

For example, it holds that for an odd cycle $\mathcal{C}_N$ (Proposition 7.2),

$$h(k) + \bar{h}(k) = 1, \quad 2 \leq k \leq N.$$

Recall that an odd cycle is not bipartite but very close to be bipartite. Moreover, this framework provides a new viewpoint about the previously defined constants. We see that the dual Cheeger constant of Bauer and Jost, $\bar{h}(1)$, is actually dual to $h(1) = 0$ and the Cheeger constant $h(2)$ is dual to $\bar{h}(2)$.

We prove higher-order dual Cheeger inequalities, i.e., we derive estimates for the spectral gaps $2 - \lambda_{N-k+1}$ in terms of $\bar{h}(k)$, which hold universally for any weighted finite graph (see Theorem 1.2). This completes the picture about graph spectra and (dual) isoperimetric constants. Interestingly, our proof proposes a new type of spectral clustering via the top k eigenfunctions employing metrics on real projective spaces. As in [26], the proof is in principle algorithmic and hence we anticipate the practical applications of this new spectral clustering.

The deep relations between higher eigenvalues and geometry of graphs have been explored in the works of Chung, Grigor'yan and Yau [12,10,11]. For discussions about the spectral gap $2 - \lambda_N$ and curvature of graphs, we refer the readers to [5]. In Markov chain theory, there is the fundamental work of Diaconis and Stroock [16] about various geometric bounds of eigenvalues, in particular, of $2 - \lambda_N$. Note that the language of Markov chains and that of normalized graph Laplacian we use here can be translated into each other. For example, a chain is aperiodic if and only if its associated graph is not bipartite.

In this spirit, it turns out that our results can be applied to a very general setting. We extend the multi-way dual Cheeger constants and higher-order dual Cheeger inequalities to a reversible Markov operator P on a probability space $(X, \mathcal{F}, \mu)$, following recent works of Miclo [30] and F.-Y. Wang [41]. Let us denote the infimum (supremum, resp.) of the essential spectrum of P by $\bar{\lambda}_{\mathrm{ess}}(P)$ ($\lambda_{\mathrm{ess}}(P)$, resp.). We obtain a characterization for $\bar{\lambda}_{\mathrm{ess}}$ in terms of extended multi-way dual Cheeger constants $\bar{h}_P(k)$,

$$\bar{\lambda}_{\mathrm{ess}}(P) > -1 \quad \Leftrightarrow \quad \inf_{k \geq 1} \bar{h}_P(k) < 1.$$

It can be considered as the counterpart of F.-Y. Wang's new criterion for λ_{ess} in terms of multi-way Cheeger constants $h_P(k)$,

$$\lambda_{\mathrm{ess}}(P) < 1 \quad \Leftrightarrow \quad \sup_{k \geq 1} h_P(k) > 0.$$

Both arguments employ an approximation procedure developed by Miclo, by which he solves the conjecture of Simon and Høegh-Krohn [36] in a semi-group context. A further discussion about the relations between $h_P(k)$ and $\bar{h}_P(k)$ enables us to arrive at

$$\sup_{k \geq 1} h_P(k) > 0 \quad \Leftrightarrow \quad -1 < \bar{\lambda}_{\mathrm{ess}}(P) \leq \lambda_{\mathrm{ess}}(P) < 1.$$

1.1. Statements of main results

In order to put our results into perspective, we start with recalling the (higher-order) Cheeger inequalities. Let $G = (V, E, w)$ be an undirected, weighted finite graph without self-loops. V and E stand for the set of vertices and edges, respectively. We denote by w_{uv} the positive symmetric weight associated to $u, v \in V$, where $e = \{u, v\} \in E$ (sometimes we also write $u \sim v$). For convenience, we may put $w_{uv} = 0$ if u, v are not connected by an edge. The degree d_u of a vertex u is then defined as $d_u := \sum_{v, v \sim u} w_{uv}$.

The expansion (or conductance) of any non-empty subset $S \subseteq V$ is defined as

$$\phi(S) = \frac{|E(S, \overline{S})|}{\mathrm{vol}(S)},$$

where $\overline{S}$ represents the complement of S in V, and $|E(S, \overline{S})| := \sum_{u \in S, v \in \overline{S}} w_{uv}$, $\mathrm{vol}(S) := \sum_{u \in S} d_u = |E(S, S)| + |E(S, \overline{S})|$.

Then, for every $k \in \mathbb{N}$, the k-way Cheeger constant is defined as

$$h(k) = \min_{S_1, S_2, \ldots, S_k} \max_{1 \le i \le k} \phi(S_i), \tag{4}$$

where the minimum is taken over all collections of k non-empty, mutually disjoint subsets $S_1, S_2, \ldots, S_k \subseteq V$. We call such kind of k subsets a k-subpartition of V, following [14]. Note by definition, we have the monotonicity $h(k) \le h(k+1)$. The classical Cheeger inequality asserts that

$$\frac{h(2)^2}{2} \le \lambda_2 \le 2h(2). \tag{5}$$

Resolving a conjecture of Miclo [29] (see also [14]), Lee, Oveis Gharan and Trevisan [26] prove the following higher-order Cheeger inequality.

Theorem 1.1 *(Lee–Oveis Gharan–Trevisan). For every graph G, and each natural number $1 \le k \le N$, we have*

$$\frac{\lambda_k}{2} \le h(k) \le Ck^2 \sqrt{\lambda_k}, \tag{6}$$

or in another form

$$\frac{1}{C^2 k^4} h(k)^2 \le \lambda_k \le 2h(k), \tag{7}$$

where C is a universal constant.

Observe that when $k > \frac{N}{2}$, at least one of k disjoint non-empty subsets must contain a single vertex, hence $h(k) = 1$. Therefore (7) is more useful for the first half part of the spectrum.

We will study the corresponding phenomena for the remaining part of the spectrum. Define the following quantity for a pair of disjoint subsets $V_1, V_2 \subseteq V$, for which $V_1 \cup V_2 \neq \emptyset$,

$$\bar{\phi}(V_1, V_2) = \frac{2|E(V_1, V_2)|}{\mathrm{vol}(V_1 \cup V_2)}.$$

Then, for every $k \in \mathbb{N}$, we can define a k-way dual Cheeger constant as follows.

$$\bar{h}(k) = \max_{(V_1, V_2), \ldots, (V_{2k-1}, V_{2k})} \min_{1 \leq i \leq k} \bar{\phi}(V_{2i-1}, V_{2i}), \tag{8}$$

where the maximum is taken over all collections of k pairs of subsets

$$(V_1, V_2), (V_3, V_4), \ldots, (V_{2k-1}, V_{2k})$$

which satisfy

$$V_p \cap V_q = \emptyset, \quad \forall 1 \leq p \neq q \leq 2k, \qquad V_{2i-1} \cup V_{2i} \neq \emptyset, \quad \forall 1 \leq i \leq k.$$

For notational simplicity, we will denote the space of all k pairs of subsets described as above by $\mathrm{Pair}(k)$ and call every element of $\mathrm{Pair}(k)$ a k-sub-bipartition of V. Here we have the monotonicity $\bar{h}(k) \geq \bar{h}(k+1)$.

Bauer and Jost [4] and independently Trevisan [39] with different methods proved a dual Cheeger inequality

$$\frac{(1 - \bar{h}(1))^2}{2} \leq 2 - \lambda_N \leq 2(1 - \bar{h}(1)). \tag{9}$$

Our main result in this paper is the following higher-order dual Cheeger inequality.

Theorem 1.2. *For every graph G, and each natural number $1 \leq k \leq N$, we have*

$$\frac{2 - \lambda_{N-k+1}}{2} \leq 1 - \bar{h}(k) \leq Ck^3 \sqrt{2 - \lambda_{N-k+1}}, \tag{10}$$

or in another form,

$$\frac{1}{C^2 k^6} \left(1 - \bar{h}(k)\right)^2 \leq 2 - \lambda_{N-k+1} \leq 2(1 - \bar{h}(k)), \tag{11}$$

where C is a universal constant.

This can be considered as a strong quantitative version of the fact that $\lambda_{N-k+1} = 2$ if and only if G has at least k bipartite connected components (see Proposition 3.1(i)).

Dually, when $k > \frac{N}{2}$, at least one of the subset pairs $\{(V_{2i-1}, V_{2i})\}_{i=1}^k \in \mathrm{Pair}(k)$ has to contain an empty subset, hence $\bar{h}(k) = 0$. Therefore (11) is more useful for the second half part of the spectrum.

1.2. Clustering on real projective spaces

The lower bound estimate of $2 - \lambda_{N-k+1}$ in (11) is the essential part of Theorem 1.2. For the proof, we will follow the route in Lee, Oveis Gharan and Trevisan [26] which justifies the spectral clustering algorithms using the bottom k eigenfunctions of [33]. (Note that by this route, one can also only get an order k^3 in (6), Lee, Oveis Gharan and Trevisan used other strong techniques to derive k^2 for the price of a much larger C.) The novel point of our proof is to explore a new type of spectral clustering.

For an orthogonal system of eigenfunctions $f_1, f_2, \ldots, f_k$ of the normalized Laplace operator Δ, one can construct the mapping

$$F : V \to \mathbb{R}^k, \quad v \mapsto \big(f_1(v), f_2(v), \ldots, f_k(v)\big). \tag{12}$$

For illustration, we ignore those vertices on which F vanishes and consider the induced mapping to a unit sphere,

$$\widetilde{F} : V \to \mathbb{S}^{k-1}, \quad v \mapsto \frac{F(v)}{\|F(v)\|},$$

where $\| \cdot \|$ is the Euclidean norm in $\mathbb{R}^k$. We will also use $\langle \cdot, \cdot \rangle$ for the inner product of vectors in $\mathbb{R}^k$.

The spectral clustering algorithms using the bottom k eigenfunctions aim at obtaining k subsets of V with smaller expansions, i.e. clustering those groups of vertices which are closely connected inside the group and loosely connected with outside vertices. Roughly speaking, [26] used the sphere distance to cluster vertices in V via their image on the unit sphere under $\widetilde{F}$.

We explore the clustering phenomenon using the top k eigenfunctions $f_{N-k+1}, \ldots, f_N$. Now use these functions in the definition of F given in (12). We first observe that for any $u \in V$

$$\frac{1}{d_u} \sum_{v, v \sim u} \big\langle F(u), F(v) \big\rangle w_{uv} = \sum_{j=N-k+1}^{N} (1 - \lambda_j) f_j(u)^2 \leq (1 - \lambda_{N-k+1}) \big\| F(u) \big\|^2.$$

Therefore if $\lambda_{N-K+1} > 1$ is large, there exists at least one neighbor v_0 of u such that $\langle F(u), F(v_0) \rangle < 0$. That is, every vertex has always at least one neighbor far away from it under the sphere distance. This indicates that the aim of a proper clustering in this case should be different. In fact, instead of pursuing small expansion subsets, we aim at finding k subsets, each of which has a bipartition such that the quantity $1 - \overline{\phi}$ is small. Roughly speaking, we hope to find k subsets whose induced subgraphs are all close to bipartite ones.

Let us explain how real projective spaces come into the situation by the following extremal but inspiring example. Consider a disconnected graph G which has two bipartite
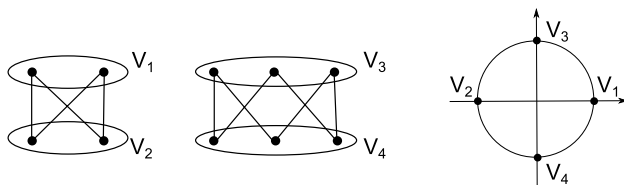
Fig. 1. The graph G and its embedding into the sphere.

connected components. Then the embedding of its vertices into the sphere $\mathbb{S}^1$ via its top two eigenfunctions are shown in Fig. 1. If we use the sphere distance, we will obtain two clusters, e.g. $V_1 \cup V_3$ and $V_2 \cup V_4$. But we actually hope to find the clusters $V_1 \cup V_2$ and $V_3 \cup V_4$. A solution for this problem is to identify the antipodal points of the sphere and to obtain two clusters $V_1 = V_2$ and $V_3 = V_4$. Afterwards, we "unfold" each cluster to get two pairs of subsets which we desire. Therefore, we should use the metric on the real projective space instead of the sphere.

To understand the above clustering via top k eigenfunctions more intuitively, we can think of the edges in E as "hostile" relations. Vertices are clustered because they share common enemies. In contrast, the traditional clustering via the bottom k eigenfunctions treat edges as "friendly" relations. We anticipate applications of this kind of hostile spectral clustering methods in practical fields, e.g. the research of social relationship networks. This hostile clustering is technically quite crucial for our purpose of proving Theorem 1.2, as discussed in Lemmas 5.2 and 5.3.

1.3. Organization of the paper

In Section 2 we collect necessary results from spectral graph theory and random partition theory of doubling metric spaces. Section 3 is devoted to various interesting relations between $\bar{h}(k)$ and $h(k)$. We discuss the lower bound estimates of λ_{N-k+1} in Section 4. In Sections 5 and 6 we present the proof of the lower bound estimate of $2 - \lambda_{N-k+1}$ of (11). In Section 7, we prove for cycles a slightly "shifted" version of higher-order dual Cheeger inequalities with an absolute constant which is even independent of k, based on the results of [14]. We also analyze the example of unweighted cycles in detail. In Section 8, we explore an application of higher-order dual Cheeger inequalities to the essential spectrum of a general reversible Markov operator.

We comment that the results about weighted graphs in this paper (except for Proposition 3.5) can be extended to graphs permitting self-loops, or in the language of Markov chains, lazy random walks. One just needs to be careful about the fact $\mu(u) \geq \sum_{v,v \sim u, v \neq u} w_{uv}$ (see below for μ) in that case.

2. Preliminaries

2.1. Spectral theory for normalized graph Laplacian

We assign a natural measure μ to V that $\mu(u) = d_u$, for every $u \in V$. The inner product of two functions $f, g : V \to \mathbb{R}$ is given by

$$(f, g)_\mu = \sum_{u \in V} \mu(u) f(u) g(u).$$

We denote by $l^2(V, \mu)$ the Hilbert space of functions on V with the above inner product.

The normalized graph Laplacian Δ is defined as follows. For any $f \in l^2(V, \mu)$, and $u \in V$

$$\Delta f(u) := \frac{1}{d_u} \sum_{v, v \sim u} w_{uv} \big(f(u) - f(v) \big). \tag{13}$$

In matrix form, $\Delta = I - P$, where I is the identity matrix, and $P := D^{-1}A$, $D^{-1}f(u) := d_u^{-1}f(u)$ and $Af(u) := \sum_{v, v \sim u} w_{uv} f(v)$.

For a map $F : V \to \mathbb{R}^k$, we denote the Rayleigh quotient of F by

$$\mathcal{R}(F) := \frac{\sum_{e=\{u,v\} \in E} \|F(u) - F(v)\|^2 w_{uv}}{\sum_{u \in V} \|F(u)\|^2 \mu(u)}, \tag{14}$$

and a dual version of the Rayleigh quotient of F

$$\overline{\mathcal{R}}(F) := \frac{\sum_{e=\{u,v\} \in E} \|F(u) + F(v)\|^2 w_{uv}}{\sum_{u \in V} \|F(u)\|^2 \mu(u)}. \tag{15}$$

The support of a map F is defined as

$$\mathrm{supp}(F) := \big\{ v \in V : F(v) \neq 0 \big\}.$$

We call λ an eigenvalue of Δ if there exists some $f \not\equiv 0$ with $\Delta f = \lambda f$. Let $0 = \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_N$ be all the eigenvalues of Δ. The Courant–Fischer–Weyl min-max principle tells us

$$\lambda_k = \min_{\substack{f_1, f_2, \ldots, f_k \not\equiv 0 \\ (f_i, f_j)_\mu = 0, \, \forall i \neq j}} \max_{\substack{f \not\equiv 0 \\ f \in \mathrm{span}\{f_1, f_2, \ldots, f_k\}}} \mathcal{R}(f), \tag{16}$$

and, dually,

$$\lambda_{N-k+1} = \max_{\substack{f_1, f_2, \ldots, f_k \not\equiv 0 \\ (f_i, f_j)_\mu = 0, \, \forall i \neq j}} \min_{\substack{f \not\equiv 0 \\ f \in \mathrm{span}\{f_1, f_2, \ldots, f_k\}}} \mathcal{R}(f). \tag{17}$$

The next lemma can be found in Bauer and Jost [4] (Lemma 3.1 there). For its various variants, see e.g. [9,26].

Lemma 2.1. *For any nonnegative function g with $\mathrm{supp}(g) \neq \emptyset$, there exists a subset $\emptyset \neq S \subseteq \mathrm{supp}(g)$ such that*

$$\mathcal{R}(g) \geq 1 - \sqrt{1 - \phi(S)^2}. \tag{18}$$

The next lemma is basically contained in the proof of Theorem 3.2 in Bauer and Jost [4].

Lemma 2.2. *For any function f with $\mathrm{supp}(f) \neq \emptyset$, there exist two subsets $V_1, V_2 \subseteq \mathrm{supp}(f)$ such that $V_1 \cap V_2 = \emptyset$, $V_1 \cup V_2 \neq \emptyset$ and*

$$\overline{\mathcal{R}}(f) \geq 1 - \sqrt{1 - \left(1 - \overline{\phi}(V_1, V_2)\right)^2}. \tag{19}$$

We remark that here we do not require each of V_1, V_2 to be non-empty, but only their union. This lemma is derived from the combination of Lemma 2.1 and a construction in Bauer and Jost [4] (following previous ideas in Desai and Rao [15]). For convenience, we recall the proof here briefly.

Proof. Denote $P(f) := \{v \in V : f(v) > 0\}$, $N(f) := \{v \in V : f(v) < 0\}$. By the assumption, $P(f) \cup N(f) \neq \emptyset$. Now construct a new graph $G' = (V', E', w')$ from the original graph G in the following way. Duplicate all the vertices in $P(f)$ and $N(f)$. Denote by u' the new vertices duplicated from u. For any edge $\{u, v\}$ such that $u, v \in P(f)$ or $u, v \in N(f)$, replace it by two new edges $\{u, v'\}, \{v, u'\}$ with the same weight $w'_{uv'} = w'_{vu'} = w_{uv}$. All the other vertices, edges and weights are unchanged.

Consider the function $g : V' \to \mathbb{R}$,

$$g(u) = \begin{cases} |f(u)|, & \text{if } u \in P(f) \cup N(f); \\ 0, & \text{otherwise.} \end{cases}$$

Then the above construction converts the inside edges of $P(f)$, $N(f)$ into the boundary edges of $\mathrm{supp}(g)$. Furthermore, one can check that

$$\overline{\mathcal{R}}(f) \geq \mathcal{R}'(g).$$

Now by Lemma 2.1, we know there exists a subset $\emptyset \neq S \subseteq \mathrm{supp}(g) = P(f) \cup N(f)$, such that $\mathcal{R}'(g) \geq 1 - \sqrt{1 - \phi'(S)^2}$. Denote $S_P = S \cap P(f)$, $S_N = S \cap N(f)$. Then we have $S_P \cap S_N = \emptyset$, $S_P \cup S_N \neq \emptyset$ and

$$\phi'(S) = \frac{|E'(S, \overline{S})|}{\mathrm{vol}'(S)} = \frac{|E(S_P, S_P)| + |E(S_N, S_N)| + |E(S_P \cup S_N, \overline{S_P \cup S_N})|}{\mathrm{vol}(S_P \cup S_N)}$$

$$= 1 - \overline{\phi}(S_P, S_N),$$

where for the last equality we used

$$\text{vol}(S_P \cup S_N) = 2\big|E(S_P, S_N)\big| + \big|E(S_P, S_P)\big| + \big|E(S_N, S_N)\big|$$
$$+ \big|E(S_P \cup S_N, \overline{S_P \cup S_N})\big|. \tag{20}$$

This completes the proof of the lemma. $\quad\square$

2.2. Padded random partitions of doubling metric space

Random partition theory of metric spaces was firstly developed in theoretical computer science. It has found many important applications in pure mathematics, see e.g. [25,22,26]. We discuss a result of that in this section which is needed in our arguments later.

We first introduce the concept of doubling metric spaces. There are two kinds of doubling properties: metric doubling and measure doubling.

The metric doubling constant ρ_X of a metric space (X, d) is defined as

$$\rho_X := \inf\bigg\{ c \in \mathbb{N} : \forall x \in X,\ r > 0,\ \exists x_1, \ldots, x_c \in X,$$

$$\text{such that } B(x, r) \subseteq \bigcup_{i=1}^{c} B\left(x_i, \frac{r}{2}\right)\bigg\},$$

where $B(x, r)$ is the closed ball in X with center x and radius r. (X, d) is called a metric doubling space if $\rho_X < +\infty$. The metric doubling dimension of (X, d) is then defined as $\dim_d(X) := \log_2 \rho_X$.

A Borel measure μ on (X, d) is called a doubling measure if there exists a number C_μ such that for any $x \in X$, $r > 0$,

$$0 < \mu\big(B(x, r)\big) \leq C_\mu \mu\left(B\left(x, \frac{r}{2}\right)\right) < +\infty.$$

Similarly we call $\dim_\mu(X) := \log_2 C_\mu$ the measure doubling dimension. Note that the measure doubling dimension of $\mathbb{R}^k$ with the standard Euclidean volume measure is exactly k.

The two doubling dimensions are related by the following result (see e.g. the remark on p. 67 of [13]).

Lemma 2.3. *If a metric space (X, d) has a doubling measure μ, then*

$$\dim_d(X) \leq 4 \dim_\mu(X). \tag{21}$$

A partition of a metric space (X, d) is a series of subsets $P = \{S_i\}_{i=1}^m$ for some number m, where $S_i \cap S_j = \emptyset$, for any $i \neq j$ and $X = \bigcup_{i=1}^m S_i$. A partition can also be considered as a map $P : X \to 2^X$, such that $P(x)$ is the unique set in $\{S_i\}_{i=1}^m$ that contains x. A random partition $\mathcal{P}$ of X is a distribution ν over the space of partitions of X. The following padded random partition theorem is a slight modification of Theorem 3.2 in Gupta, Krauthgamer and Lee [20] (see also Lemma 3.11 in [25]).

Theorem 2.4. *Let (X, d) be a finite metric subspace of (Y, d). Then for every $r > 0$, $\delta \in (0, 1)$ there exists a random partition $\mathcal{P}$, i.e. a distribution ν over all possible partitions of X, such that*

- *$\mathrm{diam}(S) \leq r$, for any S in every partition P in the support of ν;*
- *$\mathbb{P}_\nu[B_d(x, \frac{r}{\alpha}) \subseteq \mathcal{P}(x)] \geq 1 - \delta$ for all x, where $\alpha = \frac{32 \dim_d(Y)}{\delta}$.*

The random partition obtained in the above theorem is called an $(r, \alpha, 1 - \delta)$-padded random partition in [26].

Proof. We refer the readers to [20] for the proof of this theorem. But we comment here that one can replace the $\dim_d(X)$ in Theorem 3.2 of [20] by $\dim_d(Y)$ as we do here in the conclusion. The reason is that the only point where the metric doubling dimension plays a role in the proof is the following fact (this is more clear in [25]). Let $Z \subseteq X$ be a subset in which each pair of distinct points has a distance at least ϵ. Then the cardinality of $B_d(x, t) \cap Z$ is less or equal to $2^{\dim_d(X)\lceil \log_2 \frac{2t}{\epsilon} \rceil}$, for any $x \in X$ and radius $t \geq \epsilon$. Surely one can estimate the same cardinality in the bigger space Y. $\quad\square$

3. Relations between $h(k)$ and its dual $\bar{h}(k)$

In this section, we explore some interesting relations between the multi-way Cheeger and dual Cheeger constants. The following two propositions can be considered as strong extensions of Theorem 3.1 and Proposition 3.1 in Bauer and Jost [4].

Proposition 3.1. *Let G be any graph, then for each $1 \leq k \leq N$ we have*

$$\bar{h}(k) \leq 1 - h(k). \tag{22}$$

Moreover, we have

- (i) *$\bar{h}(k) = 1$ if and only if G has at least k connected components, each of which is bipartite.*
- (ii) *If G is bipartite, then $h(k) + \bar{h}(k) = 1$, for each $1 \leq k \leq N$.*
- (iii) *If $h(k) + \bar{h}(k) = 1$, then for the k-sub-bipartition $\{(V_{2i-1}, V_{2i})\}_{i=1}^k$ which assumes*

$\bar{h}(k)$, *the pair* (V_{2i_0-1}, V_{2i_0}) *with the maximal expansion, i.e.* $\phi(V_{2i_0-1} \cup V_{2i_0}) = \max_{1 \leq i \leq k} \phi(V_{2i-1} \cup V_{2i})$, *is bipartite in the sense of*

$$\left|E(V_{2i_0-1}, V_{2i_0-1})\right| = \left|E(V_{2i_0}, V_{2i_0})\right| = 0.$$

Proof. By definition of $h(k)$, $\bar{h}(k)$ and the formula (20), we have

$$1 - \bar{h}(k)$$

$$= \min_{\substack{\{(V_{2i-1}, V_{2i})\}_{i=1}^k \in \mathrm{Pair}(k)}} \max_{1 \leq i \leq k} \frac{\mathrm{vol}(V_{2i-1} \cup V_{2i}) - 2|E(V_{2i-1}, V_{2i})|}{\mathrm{vol}(V_{2i-1} \cup V_{2i})}$$

$$= \min_{\substack{\{S_i\}_{i=1}^k \\ S_i \neq \emptyset, \forall i \\ S_i \cap S_j = \emptyset, \forall i \neq j}} \min_{\substack{\{(V_{2i-1}, V_{2i})\}_{i=1}^k \\ V_{2i-1} \cup V_{2i} = S_i \\ V_{2i-1} \cap V_{2i} = \emptyset, \forall i}} \max_{1 \leq i \leq k} \frac{|E(S_i, \overline{S_i})| + |E(V_{2i}, V_{2i})| + |E(V_{2i-1}, V_{2i-1})|}{\mathrm{vol}(S_i)}$$

$$\geq \min_{\substack{\{S_i\}_{i=1}^k \\ S_i \neq \emptyset, \forall i \\ S_i \cap S_j = \emptyset, \forall i \neq j}} \max_{1 \leq i \leq k} \frac{|E(S_i, \overline{S_i})|}{\mathrm{vol}(S_i)} = h(k). \tag{23}$$

Observe further in the above calculation, that the equality in (23) can be achieved when the graph G is bipartite. Since then for each S_i, we can always find a bipartition $S_i = V_{2i-1} \cup V_{2i}$ such that $|E(V_{2i-1}, V_{2i-1})| = |E(V_{2i}, V_{2i})| = 0$. This actually proves (ii).

For (i), if $\bar{h}(k) = 1$, then there exists $\{(V_{2i-1}, V_{2i})\}_{i=1}^k \in \mathrm{Pair}(k)$ such that

$$\min_{1 \leq i \leq k} \bar{\phi}(V_{2i-1}, V_{2i}) = 1.$$

Hence for each i,

$$1 = \frac{2|E(V_{2i-1}, V_{2i})|}{\mathrm{vol}(V_{2i-1} \cup V_{2i})}.$$

Then recalling (20), we obtain

$$0 = \left|E(V_{2i}, V_{2i})\right| = \left|E(V_{2i-1}, V_{2i-1})\right| = \left|E(V_{2i-1} \cup V_{2i-1}, \overline{V_{2i-1} \cup V_{2i-1}})\right|,$$

which implies $\{V_{2i-1} \cup V_{2i}\}_{i=1}^k$ are k connected components, each of which is bipartite. Conversely, if we know that G has k connected components, each of which is bipartite, we can choose the k-sub-bipartition to be the bipartitions of those k components. Then by definition, we have

$$\bar{h}(k) \geq \min_{1 \leq i \leq k} \bar{\phi}(V_{2i-1}, V_{2i}) = 1.$$

Together with (22), we know $\bar{h}(k) = 1$.

For (iii), suppose $\{(V_{2i-1}, V_{2i})\}_{i=1}^{k} \in \mathrm{Pair}(k)$ assumes $\bar{h}(k)$, then by definition,

$$1 = \bar{h}(k) + h(k) \leq \min_{1 \leq i \leq k} \bar{\phi}(V_{2i-1}, V_{2i}) + \max_{i=1}^{k} \phi(V_{2i-1} \cup V_{2i}).$$

Let $V_{2i_0-1} \cup V_{2i_0}$ attain the maximum in the above inequality. Then we have

$$
\begin{aligned}
1 = \bar{h}(k) + h(k) &\leq \bar{\phi}(V_{2i_0-1}, V_{2i_0}) + \phi(V_{2i_0-1} \cup V_{2i_0}) \\
&= \frac{2|E(V_{2i_0-1}, V_{2i_0})| + |E(V_{2i-1} \cup V_{2i-1}, \overline{V_{2i-1} \cup V_{2i-1}})|}{\mathrm{vol}(V_{2i_0-1} \cup V_{2i_0})} \leq 1.
\end{aligned}
$$

This implies $0 = |E(V_{2i_0-1}, V_{2i_0-1})| = |E(V_{2i_0}, V_{2i_0})|$. $\quad\square$

Remark 3.2. The property (ii) above shows the duality between $h(k)$ and $\bar{h}(k)$. Recalling the fact that bipartiteness is equivalent to whenever λ is an eigenvalue, so is $2 - \lambda$ (see e.g. Lemma 1.8 in [9]), and employing property (i) above, we conclude

$$G \text{ is bipartite} \quad \Leftrightarrow \quad \lambda_k + \lambda_{N-k+1} = 2, \quad \forall k \quad \Leftrightarrow \quad h(k) + \bar{h}(k) = 1, \quad \forall k. \quad (24)$$

We note that the equality $h(k) + \bar{h}(k) = 1$ only for certain k does not imply that G is bipartite. For example, we have trivially for any graph with N vertices, when $k > \frac{N}{2}$,

$$h(k) = 1, \qquad \bar{h}(k) = 0,$$

i.e. $h(k) + \bar{h}(k) = 1$. We also have the following example.

Example 3.3. Consider the unweighted (i.e. every edge has a weight 1) complete graph K_{2n} with $2n$ vertices. By choosing n disjoint edges, it is not hard to check that

$$h(n) = \frac{2n-2}{2n-1}, \qquad \bar{h}(n) = \frac{1}{2n-1}.$$

Therefore we have $h(n) + \bar{h}(n) = 1$.

In fact even when $h(k) + \bar{h}(k) = 1$ for $2 \leq k \leq N$, the graph still can be non-bipartite. An example is an odd cycle (see Proposition 7.2). However, this graph is already very close to a bipartite graph.

By (24), it is immediately to see that for bipartite graphs the classical Cheeger inequality (5) and Theorem 1.1 are equivalent to the following dual estimates.

Corollary 3.4. *Let G be any bipartite graph. Then*

(i) $2 - \lambda_{N-1} \geq \frac{(1-\bar{h}(2))^2}{2}$.
(ii) $2 - \lambda_{N-k+1} \geq \frac{1}{C^2 k^4}(1 - \bar{h}(k))^2$ *for each natural number $1 \leq k \leq N$, where C is a universal number.*

It is interesting to note that the fact dual to Bauer–Jost's dual Cheeger inequality (9) is not the inequality (5) but the identity $\lambda_1 = 0$ for bipartite graphs.

Proposition 3.5. *For any graph G, we have for each $1 \le k \le N$*

$$\bar{h}(k) \ge \frac{1}{2}\big(1 - h(k)\big). \tag{25}$$

To prove this proposition, we need the following lemma of Bauer and Jost [4] (see also Theorem 4.2 in [3]).

Lemma 3.6. *For any subset $S \subseteq V$, there exists a partition $S = V_1 \cup V_2$ such that*

$$\big|E(V_1, V_2)\big| \ge \max\big\{\big|E(V_1, V_1)\big|, \big|E(V_2, V_2)\big|\big\}. \tag{26}$$

Proof of Proposition 3.5. For any k-subpartition $S_1, S_2, \ldots, S_k$ of V, by Lemma 3.6, for each $1 \le i \le k$ we have a partition $S_i = V_{2i-1} \cup V_{2i}$, such that

$$\big|E(V_{2i-1}, V_{2i})\big| \ge \max\big\{\big|E(V_{2i-1}, V_{2i-1})\big|, \big|E(V_{2i}, V_{2i})\big|\big\}. \tag{27}$$

By definition, we know that

$$\bar{h}(k) \ge \min_{1 \le i \le k} \bar{\phi}(V_{2i-1}, V_{2i})$$

$$= \min_{1 \le i \le k}\left(\frac{2|E(V_{2i-1}, V_{2i})| + \frac{1}{2}|E(V_{2i-1} \cup V_{2i}, \overline{V_{2i-1} \cup V_{2i}})|}{\mathrm{vol}(V_{2i-1} \cup V_{2i})} - \frac{1}{2}\phi(V_{2i-1} \cup V_{2i})\right).$$

Combining (20) and (27), we arrive at

$$\mathrm{vol}(V_{2i-1} \cup V_{2i}) \le 4\big|E(V_{2i-1}, V_{2i})\big| + \big|E(V_{2i-1} \cup V_{2i}, \overline{V_{2i-1} \cup V_{2i}})\big|.$$

Therefore, we obtain

$$\bar{h}(k) \ge \min_{1 \le i \le k}\left(\frac{1}{2} - \frac{1}{2}\phi(S_i)\right) = \frac{1}{2}\left(1 - \max_{1 \le i \le k}\phi(S_i)\right).$$

Since $S_1, S_2, \ldots, S_k$ are chosen arbitrarily, we obtain

$$\bar{h}(k) \ge \frac{1}{2}\big(1 - h(k)\big). \qquad \square$$

4. Lower bound estimate of λ_{N-k+1}

In this section, we prove the lower bound estimate of λ_{N-k+1}. For any k-sub-bipartition $\{(V_{2i-1}, V_{2i})\}_{i=1}^k$, let us denote $V^* := V \setminus \bigcup_{i=1}^k (V_{2i-1} \cup V_{2i})$. Then for every $k \in \mathbb{N}$, we define a new constant which is greater or equal to $\bar{h}(k)$, namely

$$\bar{h}^*(k) = \max_{\{(V_{2i-1}, V_{2i})\}_{i=1}^k \in \mathrm{Pair}(k)} \min_{1 \le i \le k} \frac{2|E(V_{2i-1}, V_{2i})| + \frac{1}{2}|E(V_{2i-1} \cup V_{2i}, V^*)|}{\mathrm{vol}(V_{2i-1} \cup V_{2i})}.$$

We prove the following result.

Theorem 4.1. *For any graph G and each $1 \le k \le N$, we have*

$$\lambda_{N-k+1} \ge 2\bar{h}^*(k). \tag{28}$$

Observe that the right hand side of (11) is an immediate corollary of this result.

Proof. Given any $\{(V_{2i-1}, V_{2i})\}_{i=1}^k \in \mathrm{Pair}(k)$, we choose $l^2(V, \mu)$-orthogonal functions as follows.

$$f_i(v) = \begin{cases} 1, & \text{if } v \in V_{2i-1}; \\ -1, & \text{if } v \in V_{2i}; \\ 0, & \text{otherwise}, \end{cases} \qquad i = 1, 2, \dots, k.$$

By construction, we know that every f_i is not identically 0. Then (17) tells us that

$$\lambda_{N-k+1} \ge \min_{a_1, a_2, \dots, a_k} \left\{ \frac{\sum_{e=(u,v)\in E}(f(u) - f(v))^2 w_{uv}}{\sum_{u \in V} f(u)^2 \mu(u)} \, \middle| \, f = \sum_{i=1}^k a_i f_i \right\}, \tag{29}$$

where the maximum is taken over all collections of k constants, at least one of which is non-zero. It is straightforward to see

$$\sum_{u \in V} f(u)^2 \mu(u) = \sum_{u \in V} \sum_{i=1}^k a_i^2 f_i(u)^2 \mu(u) = \sum_{i=1}^k a_i^2 \, \mathrm{vol}(V_{2i-1} \cup V_{2i}). \tag{30}$$

For the numerator of the quotient in (29), we have

$$\sum_{e=(u,v)\in E} w_{uv}\big(f(u) - f(v)\big)^2$$

$$= \frac{1}{2} \sum_{u \in V} \sum_{v \in V} w_{uv}\big(f(u) - f(v)\big)^2$$

$$= \frac{1}{2} \sum_{i=1}^k \left(\sum_{u \in V_{2i-1}} \sum_{v \in V_{2i}} + \sum_{u \in V_{2i}} \sum_{v \in V_{2i-1}} \right) w_{uv} 4a_i^2$$

$$+ \frac{1}{2} \sum_{i=1}^k \left(\sum_{u \in V_{2i-1} \cup V_{2i}} \sum_{v \in V^*} + \sum_{u \in V^*} \sum_{v \in V_{2i-1} \cup V_{2i}} \right) w_{uv} a_i^2$$

$$+ \frac{1}{2} \sum_{i=1}^{k} \sum_{u \in V_{2i-1} \cup V_{2i}} \sum_{l=1, l \neq i}^{k} \sum_{v \in V_{2l-1} \cup V_{2l}} w_{uv} \big(f(u) - f(v)\big)^2$$

$$\geq \sum_{i=1}^{k} a_i^2 \big(4 \big|E(V_{2i-1}, V_{2i})\big| + \big|E(V_{2i} \cup V_{2i-1}, V^*)\big|\big).$$

Then we obtain

$$\lambda_{N-k+1} \geq \min_{a_1, a_2, \ldots, a_k} \frac{\sum_{i=1}^{k} a_i^2 (4|E(V_{2i-1}, V_{2i})| + |E(V_{2i-1} \cup V_{2i}, V^*)|)}{\sum_{i=1}^{k} a_i^2 \operatorname{vol}(V_{2i-1} \cup V_{2i})}$$

$$\geq 2 \min_{1 \leq i \leq k} \frac{2|E(V_{2i-1}, V_{2i})| + \frac{1}{2}|E(V_{2i-1} \cup V_{2i}, V^*)|}{\operatorname{vol}(V_{2i-1} \cup V_{2i})}.$$

Since $\{(V_{2i-1}, V_{2i})\}_{i=1}^{k} \in \operatorname{Pair}(k)$ are chosen arbitrarily, we obtain

$$\lambda_{N-k+1} \geq 2 \bar{h}^*(k). \qquad \square$$

5. The metric for clustering via top k eigenfunctions

In this section, we start to prove the lower bound estimate of $2 - \lambda_{N-k+1}$ in (11). Recall that the max-min problem in (17) is solved by the corresponding eigenfunctions. Hence for the top k eigenfunctions $f_{N+k-1}, \ldots, f_N$, we have

$$\lambda_{N-k+1} = \mathcal{R}(f_{N-k+1}) = \min_{N-k+1 \leq j \leq N} \mathcal{R}(f_j).$$

Therefore

$$\lambda_{N-k+1} \leq \frac{\sum_{j=N-k+1}^{N} \sum_{e=\{u,v\} \in E} (f_j(u) - f_j(v))^2 w_{uv}}{\sum_{j=N-k+1}^{N} \sum_{u \in V} f_j(u)^2 \mu(u)}.$$

Then it is straightforward to calculate

$$2 - \lambda_{N-k+1} \geq \frac{\sum_{j=N-k+1}^{N} \sum_{e=\{u,v\} \in E} (f_j(u) + f_j(v))^2 w_{uv}}{\sum_{j=N-k+1}^{N} \sum_{u \in V} f_j(u)^2 \mu(u)}$$

$$= \frac{\sum_{e=\{u,v\} \in E} \|F(u) + F(v)\|^2 w_{uv}}{\sum_{u \in V} \|F(u)\|^2 \mu(u)} = \overline{\mathcal{R}}(F), \tag{31}$$

where F is the map from V to $\mathbb{R}^k$ defined by $F(v) = (f_{N-k+1}(v), \ldots, f_N(v))$, i.e. in the way of (12). One can also obtain the fact (31) by applying the min-max principle directly to the operator $I + P$.

Following the route in [26] for dealing with Rayleigh quotient of the bottom k eigenfunctions, we will localize F to be k disjointly supported maps $\{\Psi_i\}_{i=1}^{k}$ for which $\overline{\mathcal{R}}(\Psi_i)$

can be controlled from above by $\overline{\mathcal{R}}(F)$. Afterwards, we will apply Lemma 2.2 to handle each $\overline{\mathcal{R}}(\Psi_i)$ further. More explicitly, our requirements for the localization are for each i

- $\sum_{u \in V} \|\Psi_i(u)\|^2 \mu(u)$ can be bounded from below by a certain fraction of $\sum_{u \in V} \|F(u)\|^2 \mu(u)$;
- $\|\Psi_i(u) + \Psi_i(v)\|$ can be controlled from above by $\|F(u) + F(v)\|$.

The first requirement will be realized by the theory of random partitions on doubling metric spaces combined with the crucial Lemma 5.2 below. The second requirement is solved by Lemma 5.3 below. Before all those arguments, we first need to introduce our new metric.

5.1. Real projective space with a rough metric

We can use the standard Riemannian metric on real projective spaces inherited from the spheres via the canonical antipodal projection

$$Pr : \mathbb{S}^{k-1} \to P^{k-1}\mathbb{R}, \quad \{x, -x\} \mapsto [x].$$

But for ease of calculations, we adopt a rough metric. That is, for any $[x], [y] \in P^{k-1}\mathbb{R}$, we define

$$\bar{d}\big([x], [y]\big) := \min\{\|x + y\|, \|x - y\|\}, \tag{32}$$

where $\|\cdot\|$ is the Euclidean norm of vectors in $\mathbb{S}^{k-1} \subset \mathbb{R}^k$. It is easy to check that $\bar{d}$ is a metric on $P^{k-1}\mathbb{R}$.

Proposition 5.1. *For the metric space $(P^{k-1}\mathbb{R}, \bar{d})$, we have*

(i) $\mathrm{diam}(P^{k-1}\mathbb{R}, \bar{d}) = \sqrt{2}$;
(ii) $\dim_{\bar{d}}(P^{k-1}\mathbb{R}) \le 4(\log_2 \pi - \frac{1}{2})(k - 1)$.

Proof. Let us denote the distance function deduced from the standard Riemannian metric on $P^{k-1}\mathbb{R}$ by d_{Rie}, and the Riemannian volume measure by μ_{Rie}. (i) is easy. (Compare the fact that $\mathrm{diam}(P^{k-1}\mathbb{R}, d_{\mathrm{Rie}}) = \frac{\pi}{2}$.) One can further observe that

$$\frac{2\sqrt{2}}{\pi} d_{\mathrm{Rie}} \le \bar{d} \le d_{\mathrm{Rie}}. \tag{33}$$

Since $(P^{k-1}\mathbb{R}, d_{\mathrm{Rie}})$ has constant sectional curvature 1 (find more geometric properties of projective spaces in [19]), by the Bishop–Gromov comparison theorem,

$$\frac{\mu_{\mathrm{Rie}}(B_{\bar{d}}(x, r))}{\mu_{\mathrm{Rie}}(B_{\bar{d}}(x, \frac{r}{2}))} \le \frac{\mu_{\mathrm{Rie}}(B_{d_{\mathrm{Rie}}}(x, \frac{\pi}{2\sqrt{2}}r))}{\mu_{\mathrm{Rie}}(B_{d_{\mathrm{Rie}}}(x, \frac{r}{2}))} \le \left(\frac{\pi}{\sqrt{2}}\right)^{k-1}.$$

Furthermore, recalling Lemma 2.3, this implies that

$$\dim_{\bar{d}}\big(P^{k-1}\mathbb{R}\big) \leq 4\log_2\left(\frac{\pi}{\sqrt{2}}\right)^{k-1} = 4\left(\log_2\pi - \frac{1}{2}\right)(k-1). \qquad \square$$

Consider the vertex set V of a graph G, and a nontrivial map $F : V \to \mathbb{R}^k$. We write $\widetilde{V}_F := \operatorname{supp} F$ for convenience. Then we define a map to the real projective space,

$$Pr \circ \widetilde{F} : \widetilde{V}_F \to P^{k-1}\mathbb{R}, \quad v \mapsto Pr\left(\frac{F(v)}{\|F(v)\|}\right). \tag{34}$$

Via the metric $\bar{d}$ defined above, we obtain a non-negative symmetric function on $\widetilde{V}_F \times \widetilde{V}_F$

$$\begin{aligned}
\bar{d}_F(u,v) &:= \bar{d}\big(Pr \circ \widetilde{F}(u), Pr \circ \widetilde{F}(v)\big) \\
&= \min\left\{\left\|\frac{F(u)}{\|F(u)\|} - \frac{F(v)}{\|F(v)\|}\right\|, \left\|\frac{F(u)}{\|F(u)\|} + \frac{F(v)}{\|F(v)\|}\right\|\right\},
\end{aligned}$$

which satisfies the triangle inequality on $\widetilde{V}_F$. That is, we obtain a pseudo metric space $(\widetilde{V}_F, \bar{d}_F)$.

5.2. Spreading lemma

We prove the following spreading lemma for the new metric extending Lemma 3.2 in Lee, Oveis Gharan and Trevisan [26]. For any map $F : S \subseteq V \to \mathbb{R}^k$, let us call the quantity

$$\sum_{u \in S} \mu(u)\big\|F(u)\big\|^2$$

the l^2 mass of F on S, denoted by $\mathcal{E}_S$ for short. By spreading, we mean that the l^2 mass of F distributes evenly on $\widetilde{V}_F$.

Lemma 5.2. *Let F be the map constructed from $l^2(V,\mu)$-orthonormal functions $f_1, f_2, \ldots, f_k$ in (12). If $S \subseteq V$ satisfies $\operatorname{diam}(S \cap \widetilde{V}_F, \bar{d}_F) \leq r$, for some $0 < r < 1$, then we have*

$$\mathcal{E}_S \leq \frac{1}{k(1-r^2)}\mathcal{E}_V.$$

In the above, we used the fact that $\mathcal{E}_S = \mathcal{E}_{S\cap\widetilde{V}_F}$, $\mathcal{E}_V = \mathcal{E}_{\widetilde{V}_F}$. The map F is said to be $(r, \frac{1}{k(1-r^2)})$-spreading if it satisfies the conclusion of the lemma. This property tells us that when the subset is of small size, the l^2 mass of F of it cannot be too large. The l^2 mass of F cannot concentrate in a particular small region.

Proof. Since $\mathcal{E}_S = \mathcal{E}_{S \cap \widetilde{V}_F}$, we can suppose w.l.o.g. that $S \subseteq \widetilde{V}_F$. As in [26], for a unit vector $x \in \mathbb{R}^k$, i.e., $\|x\|^2 = \langle x, x \rangle = \sum_{i=1}^k x_i^2 = 1$, we have

$$\sum_{v \in V} \mu(v) \langle x, F(v) \rangle^2 = \sum_{v \in V} \mu(v) \left(\sum_{i=1}^k x_i f_i(v) \right)^2 = \sum_{v \in V} \sum_{i,j=1}^k x_i x_j f_i(v) f_j(v) \mu(v)$$

$$= \sum_{i,j=1}^k x_i x_j \sum_{v \in V} f_i(v) f_j(v) \mu(v) = \sum_{i=1}^k x_i^2 = 1.$$

We also have

$$\mathcal{E}_V = \sum_{v \in V} \sum_{i=1}^k \mu(v) f_i^2(v) = \sum_{i=1}^k \sum_{v \in V} \mu(v) f_i^2(v) = k.$$

Now, for any $u \in S$, we obtain

$$\frac{\mathcal{E}_V}{k} = 1 = \sum_{v \in V} \mu(v) \left\langle F(v), \frac{F(u)}{\|F(u)\|} \right\rangle^2 = \sum_{v \in V} \mu(v) \|F(v)\|^2 \left\langle \frac{F(v)}{\|F(v)\|}, \frac{F(u)}{\|F(u)\|} \right\rangle^2. \quad (35)$$

Note that

$$\left\langle \frac{F(v)}{\|F(v)\|}, \frac{F(u)}{\|F(u)\|} \right\rangle^2 = \left\langle \frac{F(v)}{\|F(v)\|}, \frac{-F(u)}{\|F(u)\|} \right\rangle^2$$

$$= \left[\frac{1}{2} \left(2 - \left\| \frac{F(v)}{\|F(v)\|} - \frac{F(u)}{\|F(u)\|} \right\|^2 \right) \right]^2$$

$$= \left[\frac{1}{2} \left(2 - \left\| \frac{F(v)}{\|F(v)\|} + \frac{F(u)}{\|F(u)\|} \right\|^2 \right) \right]^2.$$

Therefore we have

$$\left\langle \frac{F(v)}{\|F(v)\|}, \frac{F(u)}{\|F(u)\|} \right\rangle^2 = \left[1 - \frac{1}{2} \bar{d}_F(u,v)^2 \right]^2.$$

Inserting this back into (35), we arrive at

$$\frac{\mathcal{E}_V}{k} \geq \sum_{v \in S} \mu(v) \|F(v)\|^2 \left[1 - \frac{1}{2} \bar{d}_F(u,v)^2 \right]^2 \geq \left(1 - r^2 + \frac{r^4}{4} \right) \mathcal{E}_S \geq (1 - r^2) \mathcal{E}_S.$$

This proves the lemma. □

5.3. Localization lemma

Let F be a map from V to $\mathbb{R}^k$. Given $\epsilon > 0$, we define the ϵ-neighborhood of a subset $\widetilde{S} \subseteq \widetilde{V}_F$ with respect to the metric $\bar{d}_F$ as

$$N_\epsilon(\widetilde{S}, \bar{d}_F) := \left\{ v \in \widetilde{V}_F : \bar{d}_F(v, \widetilde{S}) < \epsilon \right\}. \tag{36}$$

Now for any given subset $S \subseteq V$, we define a cut-off function,

$$\theta(v) := \begin{cases} 0, & \text{if } F(v) = 0; \\ \max\left\{0, 1 - \frac{\bar{d}_F(v, S \cap \widetilde{V}_F)}{\epsilon}\right\}, & \text{otherwise.} \end{cases} \tag{37}$$

Then we can localize F to be

$$\Psi := \theta \cdot F : V \to \mathbb{R}^k.$$

It is obvious that $\Psi|_S = F|_S$ and $\mathrm{supp}(\Psi) \subseteq N_\epsilon(S \cap \widetilde{V}_F, \bar{d}_F)$. We extend Lemma 3.3 in [26] to our new metric in the following localization lemma.

Lemma 5.3. *Given $\epsilon < 2$, let Ψ be the localization of F via θ as above. Then for any $e = \{u, v\} \in E$, we have*

$$\left\| \Psi(u) + \Psi(v) \right\| \leq \left(1 + \frac{2}{\epsilon}\right) \left\| F(u) + F(v) \right\|. \tag{38}$$

Proof. First observe that if $F(u) = F(v) = 0$, (38) is trivial. If only one of $F(u)$, $F(v)$ vanishes, (38) is implied by the fact that $|\theta| \leq 1$. Therefore we only need to consider the case that both $u, v \in \widetilde{V}_F$.

It is direct to calculate for $\{u, v\} \in E$

$$\begin{aligned} \left\| \Psi(u) + \Psi(v) \right\| &= \left\| \theta(u)F(u) + \theta(v)F(v) \right\| \\ &\leq \left| \theta(u) \right| \left\| F(u) + F(v) \right\| + \left| \theta(u) - \theta(v) \right| \left\| F(v) \right\|. \end{aligned} \tag{39}$$

If $\{u, v\}$ satisfies $\bar{d}_F(u, v) = \left\| \frac{F(u)}{\|F(u)\|} - \frac{F(v)}{\|F(v)\|} \right\|$, then by Proposition 5.1(i) we know

$$\left\| \frac{F(u)}{\|F(u)\|} - \frac{F(v)}{\|F(v)\|} \right\| \leq \sqrt{2},$$

which implies $\langle F(u), F(v) \rangle \geq 0$. Therefore

$$\left| \theta(u) - \theta(v) \right| \left\| F(v) \right\| \leq \left\| F(v) \right\| \leq \left\| F(u) + F(v) \right\|.$$

If $\{u, v\}$ satisfies $\bar{d}_F(u, v) = \left\| \frac{F(u)}{\|F(u)\|} + \frac{F(v)}{\|F(v)\|} \right\|$, then $\langle F(u), F(v) \rangle \leq 0$, and we have

$$\big|\theta(u) - \theta(v)\big|\big\|F(v)\big\| \le \frac{1}{\epsilon}\bar{d}_F(u,v)\big\|F(v)\big\| = \frac{1}{\epsilon}\left\|\frac{\|F(v)\|}{\|F(u)\|}F(u) + F(v)\right\|$$

$$\le \frac{1}{\epsilon}\left(\big\|F(v) + F(u)\big\| + \left\|\frac{\|F(v)\|}{\|F(u)\|}F(u) - F(u)\right\|\right)$$

$$\le \frac{2}{\epsilon}\big\|F(u) + F(v)\big\|.$$

In conclusion, we have

$$\big|\theta(u) - \theta(v)\big|\big\|F(v)\big\| \le \frac{2}{\epsilon}\big\|F(u) + F(v)\big\|. \tag{40}$$

Recalling (39) and the fact $|\theta| \le 1$, we obtain inequality (38), finishing the proof of the lemma. $\square$

6. Finding k-sub-bipartition with small $1 - \overline{\phi}$

In this section, we will prove

Theorem 6.1. *For a map $F : V \to \mathbb{R}^k$ constructed from $l^2(V,\mu)$-orthonormal functions $f_1, f_2, \ldots, f_k$ as in (12), there exists a k-sub-bipartition $\{(V_{2i-1}, V_{2i})\}_{i=1}^k$, such that for each i*

$$\frac{(1 - \overline{\phi}(V_{2i-1}, V_{2i}))^2}{2} \le 1 - \sqrt{1 - \big(1 - \overline{\phi}(V_{2i-1}, V_{2i})\big)^2} \le Ck^6\overline{\mathcal{R}}(F), \tag{41}$$

where C is a universal constant.

It was shown in (31) that once we take the k functions above to be the top k orthonormal eigenfunctions, we will have $2 - \lambda_{N-k+1} \ge \overline{\mathcal{R}}(F)$. Then by the definition of $\overline{h}(k)$, (11) follows immediately from this theorem.

We need the following lemma (modified from Lemma 3.5 of [26]) to get k-disjoint subsets of $(\widetilde{V}_F, \bar{d}_F)$.

Lemma 6.2. *Suppose F is $(r, \frac{1}{k}(1 + \frac{1}{8k}))$-spreading, and $(\widetilde{V}_F, \bar{d}_F)$ has an $(r, \alpha, 1 - \frac{1}{4k})$-padded random partition; then there exists k non-empty, mutually disjoint subsets $T_1, T_2, \ldots, T_k \subseteq \widetilde{V}_F$ such that*

- *for any $1 \le i \ne j \le k$, $\bar{d}_F(T_i, T_j) \ge 2\frac{r}{\alpha}$;*
- *for any $1 \le i \le k$, $\mathcal{E}_{T_i} \ge \frac{1}{2k}\mathcal{E}_V$.*

Proof. Let $\mathcal{P}$ be the $(r, \alpha, 1 - \frac{1}{4k})$-padded random partition in the assumption. Denote by $I_{B_{\bar{d}_F}(v, \frac{r}{\alpha}) \subseteq \mathcal{P}(v)}$ the indicator function for the event that $B_{\bar{d}_F}(v, \frac{r}{\alpha}) \subseteq \mathcal{P}(v)$ happens.

Then we calculate the expectation

$$\mathbb{E}_{\mathcal{P}}\left[\sum_{v\in V}\mu(v)\big\|F(v)\big\|^2 I_{B_{\bar{d}_F}(v,\frac{r}{\alpha})\subseteq\mathcal{P}(v)}\right] \geq \sum_{v\in V}\mu(v)\big\|F(v)\big\|^2\left(1-\frac{1}{4k}\right).$$

If we denote $\widehat{S} := \{v \in S : B_{\bar{d}_F}(v,\frac{r}{\alpha}) \subseteq S\}$, this is equivalent to

$$\sum_{P\in\mathcal{P}}\left(\sum_{S\in P}\sum_{v\in\widehat{S}}\mu(v)\big\|F(v)\big\|^2\right)\mathbb{P}(P) \geq \left(1-\frac{1}{4k}\right)\mathcal{E}_V.$$

Therefore there exists at least one partition $P = \{S_i\}_{i=1}^m$ of $\widetilde{V}_F$ for some natural number m such that

$$\sum_{i=1}^m \mathcal{E}_{\widehat{S}_i} \geq \left(1-\frac{1}{4k}\right)\mathcal{E}_V. \tag{42}$$

By the spreading property in the assumption, we know for every $1 \leq i \leq m$

$$\mathcal{E}_{\widehat{S}_i} \leq \mathcal{E}_{S_i} \leq \frac{1}{k}\left(1+\frac{1}{8k}\right)\mathcal{E}_V. \tag{43}$$

We can construct the desired k disjoint subsets from $\widehat{S}_1, \widehat{S}_2, \ldots, \widehat{S}_m$ by the following procedure. If we can find two of these sets, say $\widehat{S}_i$, $\widehat{S}_j$, such that

$$\mathcal{E}_{\widehat{S}_i} < \frac{1}{2k}\mathcal{E}_V, \qquad \mathcal{E}_{\widehat{S}_j} < \frac{1}{2k}\mathcal{E}_V, \tag{44}$$

then we replace them by the set $\widehat{S}_i \cup \widehat{S}_j$. Note that in this process we did not violate the fact (43) since

$$\mathcal{E}_{\widehat{S}_i\cup\widehat{S}_j} < \frac{1}{k}\mathcal{E}_V < \frac{1}{k}\left(1+\frac{1}{8k}\right)\mathcal{E}_V. \tag{45}$$

We repeat the above operation until we cannot find two sets anymore such that (44) holds. Therefore, when we stop, we get a series of subsets $T_1, T_2, \ldots, T_r$ for some number r such that

$$\mathcal{E}_{T_i} \leq \frac{1}{k}\left(1+\frac{1}{8k}\right)\mathcal{E}_V, \quad \text{for } i = 1, 2, \ldots, r, \tag{46}$$

and

$$\mathcal{E}_{T_i} \geq \frac{1}{2k}\mathcal{E}_V, \quad \text{for } i = 1, 2, \ldots, r-1. \tag{47}$$

We are not sure about the lower bound of at most one of those $\{\mathcal{E}_{T_i}\}_{i=1}^r$, here we suppose w.l.o.g. that it is $\mathcal{E}_{T_r}$.

Observe that

$$(k-1) \cdot \frac{1}{k}\left(1 + \frac{1}{8k}\right) \leq 1 - \frac{1}{4k} - \frac{1}{2k}. \tag{48}$$

Recalling (42) and (46), we know $r \geq k$, and if we take

$$T_k := \bigcup_{j=k}^{r} T_j,$$

we will have

$$\mathcal{E}_{T_k} \geq \frac{1}{2k}\mathcal{E}_V.$$

This proves the lemma. $\quad\square$

Proof of Theorem 6.1. Choosing $r = \frac{1}{3\sqrt{k}}$, by Lemma 5.2 we have that F is $(r, \frac{1}{k}(1 + \frac{1}{8k}))$-spreading. If we further take $\delta = \frac{1}{4k}$, by Theorem 2.4, V has an $(r, \alpha, 1 - \frac{1}{4k})$-padded random partition with

$$\alpha = 128k \dim_{\bar{d}}\big(P^{k-1}\mathbb{R}\big).$$

Recalling Proposition 5.1(ii), we know that there exists an absolute constant $C = 4(\log_2 \pi - \frac{1}{2})$ such that $\alpha \leq 128Ck(k-1)$. Then we can apply Lemma 6.2 to find k disjoint subsets $T_1, T_2, \ldots, T_k$, such that

- for any $1 \leq i \neq j \leq k$, $\bar{d}_F(T_i, T_j) \geq 2\frac{r}{\alpha} \geq \frac{2}{3\sqrt{k}}\frac{1}{128Ck(k-1)}$;
- for any $1 \leq i \leq k$, $\mathcal{E}_{T_i} \geq \frac{1}{2k}\mathcal{E}_V$.

Now let $\{\theta_i\}_{i=1}^k$, $k \geq 2$ be the k cut-off functions defined in (37) (replacing S there by T_i) with $\epsilon = \frac{1}{3\sqrt{k}}\frac{1}{128Ck(k-1)}$. Then we get k localizations $\Psi_i = \theta_i F$ of F with disjoint supports. Note that $\Psi_i|_{T_i} = F|_{T_i}$. Applying Lemma 5.3, we arrive at

$$\overline{\mathcal{R}}(\Psi_i) \leq 2k\big(1 + 768C\sqrt{k}k(k-1)\big)^2\overline{\mathcal{R}}(F) \leq 2 \times (768C)^2k^6\overline{\mathcal{R}}(F). \tag{49}$$

Let us write $\Psi_i(u) = (\Psi_i^1(u), \Psi_i^2(u), \ldots, \Psi_i^k(u))$. Then we can conclude that there exists a $j_0 \in \{1, 2, \ldots, k\}$ such that $\Psi_i^{j_0}$ is not identically zero and

$$\overline{\mathcal{R}}\big(\Psi_i^{j_0}\big) \leq \overline{\mathcal{R}}(\Psi_i). \tag{50}$$

Then by Lemma 2.2, for each T_i, we can find two subsets $V_{2i-1}, V_{2i} \subseteq T_i$ which satisfy $V_{2i-1} \cap V_{2i} = \emptyset$, $V_{2i-1} \cup V_{2i} \neq \emptyset$ such that

$$1 - \sqrt{1 - \left(1 - \overline{\phi}(V_{2i-1}, V_{2i})\right)^2} \leq \overline{\mathcal{R}}\left(\Psi_i^{j_0}\right). \tag{51}$$

Combining (49), (50) and (51), we prove the theorem. $\square$

Applying (49) and (50) to the top k orthonormal eigenfunctions, we arrive at the following lemma.

Lemma 6.3. *For every graph G, and each natural number $1 \leq k \leq N$, there exist k disjointly supported functions $\psi_1, \ldots, \psi_k : V \to \mathbb{R}$ such that for each $1 \leq i \leq k$,*

$$\overline{\mathcal{R}}(\psi_i) \leq Ck^6(2 - \lambda_{N-k+1}), \tag{52}$$

where C is a universal constant.

Note in the above lemma that the case $k = 1$ is trivial. Combining this lemma with Theorem 4.10 in [23], we prove the following *improved* version of higher order dual Cheeger inequalities.

Theorem 6.4. *For every graph G and $1 \leq k \leq l \leq N$, we have*

$$1 - \overline{h}(k) \leq Clk^6 \frac{2 - \lambda_{N-k+1}}{\sqrt{2 - \lambda_{N-l+1}}}, \tag{53}$$

where C is a universal constant.

(53) can be seen as a dual result to Corollary 1.3(i) in [23].

7. Trees and cycles

In this section, we explore the relations between spectra and k-way Cheeger and dual Cheeger constants on trees (i.e. graphs without cycles) and cycles. In particular, we discuss the following kind of inequalities,

$$\lambda_k \geq C_1(k)h(k)^2 \quad \text{and} \quad 2 - \lambda_{N-k+1} \geq C_2(k)\left(1 - \overline{h}(k)\right)^2. \tag{54}$$

It is proved by Miclo [29] and Daneshgar, Javadi and Miclo [14] that for trees, $C_1(k)$ can be $\frac{1}{2}$, and for cycles it can be

$$C_1(k) = \begin{cases} \frac{1}{2}, & \text{if } k = 1 \text{ or } k \text{ is even;} \\ \frac{1}{48}, & \text{if } k \geq 3 \text{ is odd.} \end{cases} \tag{55}$$

That is, for those special classes of graphs, $C_1(k)$ is even independent of k. (Of course the case $k = 1$ is trivial and we list it here just for completeness.) Recalling (24) in

Remark 3.2, we can take $C_2(k) = C_1(k)$ to be independent of k for even cycles and trees since they are all bipartite.

If we replace $\bar{h}(k)$ in (54) by $\bar{h}(k-1)$, we can prove the following result for cycles.

Theorem 7.1. *For any cycle $\mathcal{C}_N$, we have for every $1 \leq k \leq N$*

$$2 - \lambda_{N-k+1} \geq C_3 \left(1 - \bar{h}(k-1)\right)^2, \tag{56}$$

with the notation $\bar{h}(0) := \bar{h}(1)$ and

$$C_3 = \begin{cases} \frac{1}{2}, & \text{if } N-k+1 \text{ is odd or } N-k+1 = N-1, N; \\ \frac{1}{48}, & \text{if } N-k+1 \leq N-2 \text{ is even.} \end{cases}$$

As commented above, we only need to prove the theorem for odd cycles, for which we need the following observation.

Proposition 7.2. *For odd cycles, we have*

$$h(k) + \bar{h}(k) = 1, \quad \text{for } 2 \leq k \leq N. \tag{57}$$

Proof. Observe the fact that any proper subset S of V possesses a bipartition $V_1 \cup V_2 = S$ such that $|E(V_1, V_1)| = |E(V_2, V_2)| = 0$. Therefore, the inequality (23) in the proof of Proposition 3.1(ii) is in fact an equality, proving $h(k) + \bar{h}(k) = 1$ for any $2 \leq k \leq N$. $\quad \square$

Proof of Theorem 7.1. Let $\mathcal{C}_{N-1}$ be a cycle obtained from $\mathcal{C}_N$ by contracting one edge $\{u_0, v_0\}$, i.e., by removing $\{u_0, v_0\}$ from the edge set of $\mathcal{C}_N$ and identifying vertices u_0, v_0 to be one vertex renamed as η in $\mathcal{C}_{N-1}$. Let us use a prime to indicate quantities of the new graph $\mathcal{C}_{N-1}$. Then we have $d'_\eta = d_{u_0} + d_{v_0} - 2w_{u_0 v_0}$ and for $u_1 \sim u_0 \sim v_0 \sim v_1$ in $\mathcal{C}_N$, $w'_{\eta u_1} = w_{u_0 u_1}$, $w_{\eta v_1} = w_{v_0 v_1}$.

Now by an interlacing idea in Butler [6], we claim

$$\lambda'_k \geq \lambda_k, \quad \text{for } 1 \leq k \leq N-1. \tag{58}$$

(Note that the interlacing results for the so-called weak coverings in [6] do not apply to our case since different weights and vertex degrees of the new graph $\mathcal{C}_{N-1}$ are chosen there. See also the general result Theorem 4.1 for contracting operations in Horak and Jost [21] which works for unweighted graphs.) Indeed, by (17) we have

$$\lambda'_k = \max_{\mathcal{F}^{N-k} \subseteq \mathbb{R}^{N-1}} \min_{f' \in \mathcal{F}^{N-k}, f' \not\equiv 0} \mathcal{R}'\left(f'\right),$$

where the maximum is taken over all possible $N - k$ dimensional subspaces $\mathcal{F}^{N-k}$ of $\mathbb{R}^{N-1}$. Given f', we define a new function f on $\mathcal{C}_N$ as

$$f(u) = \begin{cases} f'(\eta), & \text{if } u = u_0 \text{ or } v_0; \\ f'(u), & \text{otherwise.} \end{cases}$$

This satisfies

$$\mathcal{R}'\big(f'\big) = \frac{\sum_{u \sim v} w_{uv}(f(u) - f(v))^2}{\sum_u d_u f^2(u) - 2w_{u_0 v_0}(f'(\eta))^2} \geq \mathcal{R}(f).$$

Hence we have

$$\lambda'_k = \max_{\mathcal{F}^{N-k+1} \subseteq \mathbb{R}^N} \min_{\substack{f \in \mathcal{F}^{N-k+1}, f \not\equiv 0 \\ f(u_0) = f(v_0)}} \frac{\sum_{u \sim v} w_{uv}(f(u) - f(v))^2}{\sum_u d_u f^2(u) - 2w_{u_0 v_0}(f(u_0))^2}$$

$$\geq \max_{\mathcal{F}^{N-k+1} \subseteq \mathbb{R}^N} \min_{f \in \mathcal{F}^{N-k+1}, f \not\equiv 0} \mathcal{R}(f) = \lambda_k.$$

Using this and the bipartiteness of $\mathcal{C}_{N-1}$ for odd N, we obtain for $k \geq 2$,

$$2 - \lambda_{N-k+1} \geq 2 - \lambda'_{N-k+1} = 2 - \lambda'_{(N-1)-(k-1)+1} = \lambda'_{k-1} \geq \lambda_{k-1}. \tag{59}$$

By the result (55) of [14] and Proposition 7.2, we prove the theorem for odd N and $k \geq 3$. For $k = 1, 2$, by the dual Cheeger inequality (9), $2 - \lambda_{N-1} \geq 2 - \lambda_N \geq \frac{1}{2}(1 - \bar{h}(1))^2$. $\quad\square$

In the following, we consider the special class of unweighted cycles as an example. We will see that the constants in (55) can be better for unweighted cycles and $C_2(k)$ can also be independent of k for unweighted odd cycles.

Proposition 7.3. *For an unweighted cycle, we have $h(1) = 0$ and*

$$\bar{h}(1) = \begin{cases} \frac{N-1}{N}, & \text{if } N \text{ is odd;} \\ 1, & \text{if } N \text{ is even,} \end{cases}$$

$$h(k) = 1 - \bar{h}(k) = \frac{1}{\lfloor \frac{N}{k} \rfloor}, \quad \text{for } 2 \leq k \leq N.$$

Proof. For $2 \leq k \leq N$, we only need to calculate $h(k)$ since we always have $h(k) = 1 - \bar{h}(k)$.

Let $\{S_i\}_{i=1}^k$ be the k-subpartition of a cycle achieving $h(k)$. Then we can always suppose that $V^* = V \setminus \bigcup_{i=1}^k S_i = \emptyset$ and every S_i is connected, since otherwise we can construct k connected partitions from them without increasing their expansions as follows. For every i, if S_i has connected components $\{S_i^j\}_{j=1}^t$ for some natural number $t \geq 2$, we replace S_i by $S_i^{j_1} \in \{S_i^j\}_{j=1}^t$ which has the minimal expansion. Note that

$$\phi\big(S_i^{j_1}\big) = \min_{1 \leq j \leq t} \phi\big(S_i^j\big) \leq \frac{\sum_{j=1}^t |E(S_i^j, \overline{S_i^j})|}{\sum_{j=1}^t \mathrm{vol}(S_i^j)} = \phi(S_i).$$

For the new k-subpartition $\{S_i\}_{i=1}^k$, if $V^* \neq \emptyset$, we can combine each connected component of V^* with one of its adjacent subsets in $\{S_i\}_{i=1}^k$. Since the boundary measure is

unchanged while the volume increases, we again do not increase their expansion in this process.

In an unweighted cycle for any connected non-empty proper subsets $S_i \subseteq V$, we have $\phi(S_i) = \frac{2}{2\sharp S_i}$, where $\sharp S_i$ represents the number of vertices in S_i. Then it is straightforward to obtain

$$h(k) = \min_{S_1,\ldots,S_k} \max_{1 \leq i \leq k} \frac{1}{\sharp S_i} = \frac{1}{\lfloor \frac{N}{k} \rfloor}.$$

By similar arguments for odd N, we can prove $\bar{h}(1) = \frac{N-1}{N}$. □

It is known that the eigenvalues of an unweighted cycle (see e.g. Example 1.5 in [9]), listed in an increasing order, are

$$\lambda_k = 1 - \cos\left(\frac{2\pi}{N} \left\lfloor \frac{k}{2} \right\rfloor\right), \quad k = 1, 2, \ldots, N.$$

It is then straightforward to check that

$$2 - \lambda_{N-k+1} = \begin{cases} 1 - \cos\frac{(k-1)\pi}{N}, & \text{if } N - k + 1 \text{ is even}; \\ 1 - \cos\frac{k\pi}{N}, & \text{if } N - k + 1 \text{ is odd}. \end{cases}$$

Proposition 7.4. *For an unweighted cycle, we have for every $1 \leq k \leq N$,*

$$C_1 h(k)^2 \leq \lambda_k \leq \frac{\pi^2}{2} h(k)^2, \tag{60}$$

$$C_2\left(1 - \bar{h}(k)\right)^2 \leq 2 - \lambda_{N-k+1} \leq \frac{\pi^2}{2}\left(1 - \bar{h}(k)\right)^2, \tag{61}$$

where

$$C_1 = \begin{cases} 1, & \text{if } k \text{ is even}; \\ \frac{\pi}{9}, & \text{if } k \text{ is odd}, \end{cases} \quad \text{and} \quad C_2 = \begin{cases} \frac{\pi}{9}, & \text{if } N - k + 1 \text{ is even}; \\ 1, & \text{if } N - k + 1 \text{ is odd}. \end{cases}$$

Proof. Recall the following basic inequalities

$$1 - \cos x \leq \frac{x^2}{2}, \quad \forall x \geq 0, \qquad 1 - \cos x \geq \frac{x^2}{\pi}, \quad \forall 0 \leq x \leq \frac{\pi}{2}.$$

We only need to consider $k \geq 2$. When $N - k + 1$ is odd, we have

$$2 - \lambda_{N-k+1} = 1 - \cos\frac{k\pi}{N} \leq \frac{\pi^2 k^2}{2N^2} \leq \frac{\pi^2}{2}\left(\frac{1}{\lfloor \frac{N}{k} \rfloor}\right)^2 = \frac{\pi^2}{2}\left(1 - \bar{h}(k)\right)^2.$$

If $k > \frac{N}{2}$, we have $\bar{h}(k) = 0$, and therefore $2 - \lambda_{N-k+1} > 1 = (1 - \bar{h}(k))^2$. In the case $k \leq \frac{N}{2}$, we have $\frac{N}{k} \leq \lfloor \frac{N}{k} \rfloor + 1 \leq \frac{3}{2}\lfloor \frac{k}{N} \rfloor$, and

$$2 - \lambda_{N-k+1} \geq \frac{\pi k^2}{N^2} \geq \frac{4\pi}{9} \left(\frac{1}{\lfloor \frac{N}{k} \rfloor} \right)^2 = \frac{4\pi}{9} \left(1 - \bar{h}(k) \right)^2.$$

This verifies (61) for odd $N - k + 1$ and in fact also (60) for even k.

When $N - k + 1$ is even, we have similarly

$$2 - \lambda_{N-k+1} = 1 - \cos \frac{(k-1)\pi}{N} \leq \frac{\pi^2}{2} \left(1 - \bar{h}(k) \right)^2.$$

If $k > \frac{N}{2} + 1$, we have $2 - \lambda_{N-k+1} > 1 = (1 - \bar{h}(k))^2$. If $k \leq \frac{N}{2}$, then we have $\frac{N}{k-1} \leq \frac{2N}{k} \leq 3 \lfloor \frac{N}{k} \rfloor$ and

$$2 - \lambda_{N-k+1} \geq \pi \left(\frac{k-1}{N} \right)^2 \geq \frac{\pi}{9} \left(\frac{1}{\lfloor \frac{N}{k} \rfloor} \right)^2 = \frac{\pi}{9} \left(1 - \bar{h}(k) \right)^2.$$

In the case $\frac{N}{2} < k \leq \frac{N}{2} + 1$, we have $\bar{h}(k) = 0$. Observing that $2 - \lambda_{N-k+1}$ is equal to 1 if N is even, and $1 - \cos \frac{(N-1)\pi}{2N} \geq 1 - \cos \frac{\pi}{3} > \frac{\pi}{9}$ if N is odd, we have verified (61) for even $N - k + 1$. This also shows (60) for odd k. $\quad\square$

Remark 7.5. (i) This example shows that unweighted cycles are a class of graphs for which λ_k and $h(k)^2$, as well as $2 - \lambda_{N-k+1}$ and $(1 - \bar{h}(k))^2$, are equivalent, respectively up to an absolute constant. In fact, one can easily extend this conclusion to weighted cycles with uniformly bounded weights, paying the price that the constant will then depend on the uniform weight bounds. In general, such kind of results for λ_k and $h(k)^2$ are called Buser type inequalities. Higher order Buser inequalities for nonnegatively curved graphs are proved in [27].

(ii) We emphasize that this example also shows that it is possible to expect for certain classes of graphs that C_1, C_2 can be improved to be 1 for even k and odd $N - k + 1$, respectively. For related discussions about the Cheeger inequality, see Chapter 5 of [32].

(iii) In [16] (see Examples 3.1–3.4 there), Diaconis and Stroock obtained explicit formulas for $h(2)$ of several interesting non-trivial example graphs, including odd cycles.

8. Essential spectrum of reversible Markov operators

In this last section, we discuss an application of the higher-order dual Cheeger inequality (11) to characterize the essential spectrum of a general reversible Markov operator, in the spirit of Miclo [30] and F.-Y. Wang [41].

Let us start from extending our notations to that abstract setting. Assume that $(X, \mathcal{F}, \mu)$ is a probability space. We define

$$L^2(X, \mu) := \left\{ f : X \to \mathbb{R} \text{ measurable}, \int_X f^2 d\mu < +\infty \right\}.$$

Then let $P : L^2(X, \mu) \to L^2(X, \mu)$ be a linear operator such that

$$P1 = 1 \text{ and, for any } f \in L^2(X, \mu), \ f \geq 0 \text{ implies } Pf \geq 0, \tag{62}$$

where 1 stands for the constant function taking the value one. We will also use 1_S for a measurable subset $S \subseteq X$ to represent the characteristic function of S. The operator P is then called a Markov operator. We will consider a reversible (alternatively called symmetric) Markov operator P with an invariant measure μ. Explicitly, we require for any $f, g \in L^2(X, \mu)$,

$$\int_X g(x) P f(x) d\mu(x) = \int_X f(x) P g(x) d\mu(x). \tag{63}$$

Actually, there exists a symmetric measure J on $X \times X$ (see e.g. [41]), such that

$$d\mu(x) = \int_{y \in X} J(dx, dy). \tag{64}$$

Then we have for any two measurable subsets A, B of X,

$$J(A, B) = (1_A, P 1_B)_\mu, \tag{65}$$

where the inner product notion is extended from the previous finite graph setting.

Remark 8.1. The previous weighted finite graph setting can be fitted into this general framework, see [30] for a dictionary. However, in this section we only discuss the case that $L^2(X, \mu)$ is infinite dimensional.

It is known that the spectrum $\sigma(P)$ of the operator P lies in $[-1, 1]$. In the following we will denote the top and bottom of the essential spectrum $\sigma_{\text{ess}}(P)$ of P by

$$\lambda_{\text{ess}}(P) := \sup \sigma_{\text{ess}}(P) \quad \text{and} \quad \bar{\lambda}_{\text{ess}}(P) := \inf \sigma_{\text{ess}}(P).$$

Miclo [30] extended the notion of multi-way Cheeger constants and the higher-order Cheeger inequalities to reversible Markov operators among many other different settings. He defines the k-way Cheeger constant by

$$h_P(k) := \inf_{S_1, S_2, \ldots, S_k} \max_{1 \leq i \leq k} \frac{(1_{S_i}, P 1_{\overline{S}_i})_\mu}{\mu(S_i)},$$

where the infimum is taken over all possible k-subpartitions, precisely, all collections of k disjoint subsets $S_1, S_2, \ldots, S_k$ such that for each $1 \leq i \leq k$, $\mu(S_i) > 0$. One then has the monotonicity $h_P(k) \leq h_P(k+1)$.

Adapting Miclo's approximation procedure [30], F.-Y. Wang proved the following characterization of $\lambda_{\mathrm{ess}}(P)$.

Theorem 8.2. *(See F.-Y. Wang [41].) Let P be a reversible Markov operator on $L^2(X, \mu)$. Then*

$$\lambda_{\mathrm{ess}}(P) < 1 \quad \Leftrightarrow \quad \sup_{k \geq 1} h_P(k) > 0. \tag{66}$$

In their spirit, we extend the k-way dual Cheeger constant to the present setting as follows.

$$\bar{h}_P(k) := \sup_{(A_1, A_2), \ldots, (A_{2k-1}, A_{2k})} \min_{1 \leq i \leq k} \frac{2(1_{A_{2i-1}}, P1_{A_{2i}})_\mu}{\mu(A_{2i-1} \cup A_{2i})},$$

where the supremum is taken over all possible k-sub-bipartitions of X, precisely, all collections of k pairs of subsets, $(A_1, A_2), \ldots, (A_{2k-1}, A_{2k})$, where for any $1 \leq p \neq q \leq 2k$, A_p and A_q are disjoint and for each $1 \leq i \leq k$, $\mu(A_{2i-1} \cup A_{2i}) > 0$. Accordingly, we have $\bar{h}_P(k) \geq \bar{h}_P(k+1)$.

Then we have the following relations between $h_P(k)$ and $\bar{h}_P(k)$, extending the previous Proposition 3.1.

Proposition 8.3. *For any $k \geq 1$,*

$$h_P(k) + \bar{h}_P(k) \leq 1.$$

Proof. This proposition can be proved in the same way as (23) proves (22), bearing in mind the following fact for a partition $A_1 \cup A_2$ of $S \subseteq X$:

$$\mu(S) = J(S, \overline{S}) + 2J(A_1, A_2) + J(A_1, A_1) + J(A_2, A_2),$$

and recalling (65). $\quad \square$

We prove the following characterization of $\bar{\lambda}_{\mathrm{ess}}(P)$ in terms of the multi-way dual Cheeger constants.

Theorem 8.4. *Let P be a reversible Markov operator on $L^2(X, \mu)$. Then*

$$\bar{\lambda}_{\mathrm{ess}}(P) > -1 \quad \Leftrightarrow \quad \inf_{k \geq 1} \bar{h}_P(k) < 1. \tag{67}$$

An immediate corollary is the following.

Corollary 8.5. *Let P be a reversible Markov operator on $L^2(X, \mu)$. Then*

$$\sup_{k \geq 1} h_P(k) > 0 \quad \Leftrightarrow \quad -1 < \bar{\lambda}_{\mathrm{ess}}(P) \leq \lambda_{\mathrm{ess}}(P) < 1.$$

This is true, because by Proposition 8.3, the condition $\inf_{k\geq 1} \bar{h}_P(k) < 1$ is weaker than $\sup_{k\geq 1} h_P(k) > 0$.

To prove Theorem 8.4, we need to extend the higher-order dual Cheeger inequalities to the present setting. Recalling the comments after (31), the proper operator we should use here is $\bar{L} = I + P$, which is bounded and self-adjoint. Then we can follow [41] to study

$$\bar{\lambda}_k := \sup_{f_1,\ldots,f_{k-1}\in L^2(\mu)} \inf_{\substack{(f,f_i)_\mu=0 \\ \forall 1\leq i\leq k-1}} \frac{(f,\bar{L}f)_\mu}{(f,f)_\mu}.$$

Define $\bar{\lambda}_{\mathrm{ess}}(\bar{L}) := \inf \sigma_{\mathrm{ess}}(\bar{L})$. Then $\bar{\lambda}_k$ is the k-th eigenvalue of $\bar{L}$ if $\bar{\lambda}_k < \bar{\lambda}_{\mathrm{ess}}(\bar{L})$ and $\bar{\lambda}_k = \bar{\lambda}_{\mathrm{ess}}(\bar{L})$ otherwise (see e.g. [41]).

We can now state the following inequalities.

Theorem 8.6. *Let C be the same constant as in (11). Then, for $k \geq 1$,*

$$\frac{1}{C^2 k^6}\big(1 - \bar{h}_P(k)\big)^2 \leq \bar{\lambda}_k \leq 2\big(1 - \bar{h}_P(k)\big). \tag{68}$$

Proof. The upper bound can be proved by the same technique used in the proof of Theorem 4.1. One only needs to keep in mind that

$$\big(f, (I + P)f\big)_\mu = 2(f,f)_\mu - \frac{1}{2}\int_{X\times X} \big(f(x) - f(y)\big)^2 J(dx, dy), \tag{69}$$

for any $f \in L^2(X, \mu)$.

For the lower bound, we refer to the proof of Lemma 2.2 in [41], because we only need to replace the operator $L = I - P$ there by $\bar{L}$ and use the higher-order dual Cheeger inequalities for the finite discrete structure there. Basically, the approximation procedure only involves the operator P. We also recall here the fact that, on a graph with N vertices, $\bar{\lambda}_k$ of the operator $I + P$ equals $2 - \lambda_{N-k+1}$, where λ_{N-k+1} is the $(N-k+1)$-th eigenvalue of $\Delta = I - P$. $\square$

Proof of Theorem 8.4. The proof can be done in the same way as [41]. For the reader's convenience, we recall it here. First observe

$$\bar{\lambda}_{\mathrm{ess}}(P) > -1 \quad \Leftrightarrow \quad \bar{\lambda}_{\mathrm{ess}}(\bar{L}) > 0.$$

If $\bar{\lambda}_{\mathrm{ess}}(\bar{L}) > 0$, then $\sigma(\bar{L}) \cap [0, \bar{\lambda}_{\mathrm{ess}}(\bar{L}))$ is discrete and every eigenvalue in it has finite multiplicity. Therefore, $\bar{\lambda}_k > 0$ for large enough k. Hence by (68), $1 - \bar{h}_P(k) > 0$ for large k.

Otherwise, if $\bar{\lambda}_{\mathrm{ess}}(\bar{L}) = 0$, then $0 \in \sigma_{\mathrm{ess}}(\bar{L})$, and therefore $\bar{\lambda}_k = 0$ for all k. Now using (68) again, we arrive at $\bar{h}_P(k) = 1$ for all k. $\square$

Remark 8.7. Observe that for this application, the order of k in (68) is not important. But we do need the constant in (11) to be universal for any weighted finite graph to derive (68) via the approximation procedure in [29,41].

References

[1] N. Alon, Eigenvalues and expanders, in: Theory of Computing, Singer Island, FL, 1984, Combinatorica 6 (2) (1986) 83–96.

[2] N. Alon, V. Milman, λ_1, isoperimetric inequalities for graphs, and superconcentrators, J. Combin. Theory Ser. B 38 (1) (1985) 73–88.

[3] F. Bauer, B. Hua, J. Jost, The dual Cheeger constant and spectra of infinite graphs, Adv. Math. 251 (2014) 147–194.

[4] F. Bauer, J. Jost, Bipartite and neighborhood graphs and the spectrum of the normalized graph Laplacian, Comm. Anal. Geom. 21 (4) (2013) 787–845.

[5] F. Bauer, J. Jost, S. Liu, Ollivier–Ricci curvature and the spectrum of the normalized graph Laplace operator, Math. Res. Lett. 19 (6) (2012) 1185–1205.

[6] S. Butler, Interlacing for weighted graphs using the normalized Laplacian, Electron. J. Linear Algebra 16 (2007) 90–98.

[7] J. Cheeger, A lower bound for the smallest eigenvalue of the Laplacian, in: Problems in Analysis, Papers Dedicated to Salomon Bochner, 1969, Princeton Univ. Press, Princeton, NJ, 1970, pp. 195–199.

[8] F.R.K. Chung, Laplacians of graphs and Cheeger's inequalities, in: Combinatorics, Paul Erdős is Eighty, vol. 2, Keszthely, 1993, in: Bolyai Soc. Math. Stud., vol. 2, János Bolyai Math. Soc., Budapest, 1996, pp. 157–172.

[9] F.R.K. Chung, Spectral Graph Theory, CBMS Reg. Conf. Ser. Math., vol. 92, American Mathematical Society, Providence, RI, 1997.

[10] F.R.K. Chung, A. Grigor'yan, S.-T. Yau, Upper bounds for eigenvalues of the discrete and continuous Laplace operators, Adv. Math. 117 (2) (1996) 165–178.

[11] F.R.K. Chung, A. Grigor'yan, S.-T. Yau, Higher eigenvalues and isoperimetric inequalities on Riemannian manifolds and graphs, Comm. Anal. Geom. 8 (5) (2000) 969–1026.

[12] F.R.K. Chung, S.-T. Yau, Eigenvalues of graphs and Sobolev inequalities, Combin. Probab. Comput. 4 (1) (1995) 11–25.

[13] R. Coifman, G. Weiss, Analyse harmonique non-commutative sur certains espaces homogènes, Lecture Notes in Math., vol. 242, Springer, Berlin, New York, 1971.

[14] A. Daneshgar, R. Javadi, L. Miclo, On nodal domains and higher-order Cheeger inequalities of finite reversible Markov processes, Stochastic Process. Appl. 122 (4) (2012) 1748–1776.

[15] M. Desai, V. Rao, A characterization of the smallest eigenvalue of a graph, J. Graph Theory 18 (2) (1994) 181–194.

[16] P. Diaconis, D. Strook, Geometric bounds for eigenvalues of Markov chains, Ann. Appl. Probab. 1 (1) (1991) 36–61.

[17] J. Dodziuk, Difference equations, isoperimetric inequality and transience of certain random walks, Trans. Amer. Math. Soc. 284 (2) (1984) 787–794.

[18] K. Funano, Eigenvalues of Laplacian and multi-way isoperimetric constants on weighted Riemannian manifolds, arXiv:1307.3919v1, July 2013.

[19] S. Gallot, D. Hulin, J. Lafontaine, Riemannian Geometry, third edition, Universitext, Springer-Verlag, Berlin, 2004, xvi+322 pp.

[20] A. Gupta, R. Krauthgamer, J.R. Lee, Bounded geometries, fractals, and low-distortion embeddings, in: 44th Symposium on Foundations of Computer Sciences, 2003, pp. 534–543.

[21] D. Horak, J. Jost, Interlacing inequalities for eigenvalues of discrete Laplace operators, Ann. Global Anal. Geom. 43 (2) (2013) 177–207.

[22] J.A. Kelner, J.R. Lee, G.N. Price, S.-H. Teng, Metric uniformization and spectral bounds for graphs, Geom. Funct. Anal. 21 (5) (2011) 1117–1143.

[23] T.-C. Kwok, L.-C. Lau, Y.-T. Lee, S. Oveis Gharan, L. Trevisan, Improved Cheeger's inequality: analysis of spectral partitioning algorithms through higher order spectral gap, in: STOC'13 – Proceedings of the 2013 ACM Symposium on Theory of Computing, ACM, New York, 2013, pp. 11–20.

[24] G.F. Lawler, A.D. Sokal, Bounds on the L^2 spectrum for Markov chains and Markov processes: a generalization of Cheeger's inequality, Trans. Amer. Math. Soc. 309 (2) (1988) 557–580.

[25] J.R. Lee, A. Naor, Extending Lipschitz functions via random metric partitions, Invent. Math. 160 (1) (2005) 59–95.

[26] J.R. Lee, S. Oveis Gharan, L. Trevisan, Multi-way spectral partitioning and higher-order Cheeger inequalities, in: STOC'12 – Proceedings of the 2012 ACM Symposium on Theory of Computing, ACM, New York, 2012, pp. 1117–1130.

[27] S. Liu, N. Peyerimhoff, Eigenvalue ratios of nonnegatively curved graphs, arXiv:1406.6617, June 2014.

[28] A. Lubotzky, Discrete Groups, Expanding Graphs and Invariant Measures, with an appendix by Jonathan D. Rogawski, Progr. Math., vol. 125, Birkhäuser Verlag, Basel, 1994, xii+195 pp.

[29] L. Miclo, On eigenfunctions of Markov processes on trees, Probab. Theory Related Fields 142 (3–4) (2008) 561–594.

[30] L. Miclo, On hyperboundedness and spectrum of Markov operators, Invent. Math. (July 2014), http://dx.doi.org/10.1007/s00222-014-0538-8, online first, in press.

[31] B. Mohar, Isoperimetric numbers of graphs, J. Combin. Theory Ser. B 47 (3) (1989) 274–291.

[32] R. Montenegro, P. Tetali, Mathematical aspects of mixing times in Markov chains, Found. Trends Theor. Comput. Sci. 1 (3) (2006), x+121 pp.

[33] A. Ng, M. Jordan, Y. Weiss, On spectral clustering: analysis and an algorithm, in: T. Dietterich, S. Becker, Z. Ghahramani (Eds.), Adv. Neural Inf. Process. Syst., vol. 14, MIT Press, 2001, pp. 849–856.

[34] N. Peyerimhoff, A. Vdovina, Cayley graph expanders and groups of finite width, J. Pure Appl. Algebra 215 (11) (2011) 2780–2788.

[35] R. Sánchez-García, M. Fennelly, S. Norris, N. Wright, G. Niblo, J. Brodzki, J. Bialek, Hierarchical spectral clustering of power grids, IEEE Trans. Power Syst. 29 (5) (2014) 2229–2237.

[36] B. Simon, R. Høegh-Krohn, Hypercontractive semigroups and two dimensional self-coupled Bose fields, J. Funct. Anal. 9 (1972) 121–180.

[37] A. Sinclair, M. Jerrum, Approximate counting, uniform generation and rapidly mixing Markov chains, Inform. and Comput. 82 (1) (1989) 93–133.

[38] R.M. Tanner, Explicit concentrators from generalized N-gons, SIAM J. Algebr. Discrete Methods 5 (3) (1984) 287–293.

[39] L. Trevisan, Max cut and the smallest eigenvalue, in: STOC'09 – Proceedings of the 2009 ACM International Symposium on Theory of Computing, ACM, New York, 2009, pp. 263–271; SIAM J. Comput. 41 (6) (2012) 1769–1786.

[40] U. von Luxburg, A tutorial on spectral clustering, Stat. Comput. 17 (4) (2007) 395–416.

[41] F.-Y. Wang, Criteria of spectral gap for Markov operators, J. Funct. Anal. 266 (2014) 2137–2152.