Recent Advances in Approximation Algorithms **Spring 2015**

Lecture 17: Cheeger's Inequality and the Sparsest Cut Problem

Lecturer: Shayan Oveis Gharan *June 1st*

Disclaimer: *These notes have not been subjected to the usual scrutiny reserved for formal publications.*

In the rest of this course we use the proof of Marcus, Spielman and Srivastava to prove an upper bound of polyloglog(n) on the integrality gap of the Held-Karp relaxation for ATSP. The materials will be based on the work of Oveis Gharan and Anari [AO14]. We start by introducing the sparsest cut problem and Cheeger's inequalities. The ideas that we develop here will be crucially used later. In particular, we will see that the ideas involved in the semidefinite programming relaxation of the sparsest cut can be used to design an algorithm for finding thin trees.

For a graph $G = (V, E)$, the *conductance* of a set S is the ratio of the fraction of edges in the cut $(S, \overline{S})$ to the volume of S,

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

where $\mathrm{vol}(S) = \sum_{v \in S} d(v)$ is the sum of the degree of vertices in S. Observe that for any set $S \subseteq V$, $0 \leq \phi(S) \leq 1$. If $\phi(S) \approx 0$, S may represent a cluster in G. Conductance is a very well studied measure for graph clustering in the literature (see e.g. [SM00; KVV04; TM06]). The conductance of G, $\phi(G)$ is the smallest conductance among all sets with at most half of the total volume,

$$\phi(G) = \min_{S:\mathrm{vol}(S) \leq \mathrm{vol}(V)/2} \phi(S).$$

For example, the conductance of a complete graph K_n is $\phi(K_n) \approx 1/2$, and the worst set is any set with half of the vertices. The conductance of a cycle (of length n), C_n is about $2/n$, $\phi(C_n) \approx 2/n$ and the worst set is a path of length $n/2$. We say a graph G is an expander if $\phi(G) \geq \Omega(1)$. In this lecture we will see several properties of expander graphs.

Throughout this lecture we assume that G is a d-regular unweighted graph but the statements naturally extend to weighted nonregular graphs.

17.1 Cheeger's inequality

Cheeger's inequality is perhaps one of the most fundamental inequalities in Discrete optimization, spectral graph theory and the analysis of Markov Chains. It relates the eigenvalue of the *normalized Laplacian* matrix to $\phi(G)$. It has many applications in graph clustering [ST96; KVV04], explicit construction of expander graphs [JM85; HLW06; Lee12], analysis of Markov chains [SJ89; JSV04], and image segmentation [SM00].

The normalized Laplacian matrix of G, $\mathcal{L}_G$ is defined as follows

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

In the special case where G is a d-regular graph, we have $\mathcal{L}_G = I - A/d$. The matrix $P = A/d$ is also known as the transition probability matrix of the simple random walk on G. In Assignment 1 we saw that the eigenvalues of the Laplacian matrix of a d regular graph are between $[0, 2d]$. Therefore, the eigenvalues of the normalized Laplacian matrix (of a d regular graph) are between $[0, 2]$. It is not hard to see that this also holds for any nonregular graph.

If G is regular, then an eigenvalue λ of $\mathcal{L}_G$ corresponds to an eigenvalue $1-\lambda$ of $P = A/d$. In part (a) of problem 4 of Assignment 4 we see that the same relation holds even if G is nonregular.

Recall that the first eigenvalue of $\mathcal{L}_G$ is $\lambda_1 = 0$. By variational principle the second eigenvalues λ_2 can be characterized as follows:

$$\lambda_2 = \min_{x \perp \mathbf{1}} \frac{x^\intercal \mathcal{L}_G x}{x^\intercal x} = \min_{x \perp \mathbf{1}} \frac{\sum_{u \sim v}(x(u)-x(v))^2}{d \cdot \sum_v x(v)^2}.$$

The RHS ratio is also known as the *Rayleigh quotient* of x,

$$\mathcal{R}(x) := \frac{\sum_{u \sim v}(x(u)-x(v))^2}{d \cdot \sum_v x(v)^2}.$$

The Rayleigh quotient is closely related to sparsity of cuts. For a set $S \subseteq V$,

$$\mathcal{R}(\mathbf{1}_S) = \frac{|E(S,\overline{S})|}{d \cdot |S|} = \phi(S).$$

In other words, the Rayleigh can be seen as a continuous relaxation of conductance. Note that $\mathbf{1}_S \not\perp \mathbf{1}$. But, if we instead work with $x = \frac{\mathbf{1}_S}{|S|} - \frac{\mathbf{1}_{\overline{S}}}{|\overline{S}|}$, we get $\mathcal{R}(x) \geq \phi(S)/2$. Therefore,

$$\phi(G) = \min_{S:|S| \leq |V|/2} \phi(S) \geq \frac{1}{2} \min_{x \perp \mathbf{1}} \mathcal{R}(x) = \lambda_2/2.$$

The Cheeger's inequality shows that the converse of the above also holds with a quadratic loss.

Theorem 17.1 (Discrete Cheeger's inequality)**.** *For any graph G,*

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

In other words, $\phi(G)$ is close to zero and G has a natural 2-clustering if and only if λ_2 is close to zero. The left side of the above inequality, that we just proved, is usually referred to as the easy direction and the right side is known as the hard direction which we prove in the following section.

Both sides of the above inequality are tight. The left side is tight for the hypercube $\{0,1\}^k$, where $\lambda_2 \approx \phi(G) \approx k$ and the right side is tight for a cycle, where $\lambda_2 \approx 1/n^2$ while $\phi(G) \approx 1/n$. Note that it is very easy to see that in C_n, $\lambda_2 \leq O(1/n^2)$. it is sufficient to plugin the vector

$$x(i) = \begin{cases} -n/4 + i & \text{if} i \leq n/2 \\ 3n/4 - i & \text{if} i \geq n/2. \end{cases}$$

Although the right side is tight it is possible to strengthen it using higher eigenvalues of the normalized Laplacian. It is proved in [Kwo+13] that for any graph G,

$$\phi(G) \leq O(k)\lambda_2/\sqrt{\lambda_k},$$

i.e., if λ_k is large for a small k the square loss is not necessary in the Cheeger's inequality.

The proof of the hard direction is algorithmic, i.e., it gives a set S of conductance $O(\sqrt{\phi(G)})$. This gives an $O(\sqrt{\phi(G)})$ approximation algorithm for the sparsest cut problem. The algorithm is quite simple and is known as the spectral partitioning algorithm.

Algorithm 1 Spectral Partitioning Algorithm

Find a vector x such that $\mathcal{R}(x) \leq O(\lambda_2)$. One can use power iteration algorithm to find x in linear time but we do not discuss the details.
Sort the vertices of V according to x, $x(v_1) \leq x(v_2) \leq \cdots \leq x(v_n)$.
Return the threshold cut $(\{v_1, \ldots, v_i\}, \{v_{i+1}, \ldots, v_n\})$ with the smallest conductance.

17.1.1 Proof of the Hard Direction

Given a vector x, we construct two nonnegative vectors y, z with disjoint support such that $\mathcal{R}(y), \mathcal{R}(z) \leq O(\mathcal{R}(x))$. Say the size of the support of y is smaller than z, we use Lemma 17.3 below to show that there is a set S in the support of y of conductance $\phi(S) \leq \sqrt{2\mathcal{R}(y)} \leq O(\sqrt{\mathcal{R}(x)})$. Since the size of the support of y is at most $n/2$, $|S| \leq n/2$ and this completes the proof.

It remains to construct y, z.

Lemma 17.2. *For any vector $x : V \to \mathbb{R}$, there are disjointly supported nonnegative vectors y, z such that*

$$\mathcal{R}(y), \mathcal{R}(z) \leq 4\mathcal{R}(x).$$

Proof. Let $x \perp \mathbf{1}$ and $\|x\| = 1$. Ideally we would like to let $y = \max\{0, x\}$ and $z = \min\{0, x\}$. Then,

$$\sum_{u \sim v} (y(u) - y(v))^2 + (z(u) - z(v))^2 \leq \sum_{u \sim v} (x(u) - x(v))^2.$$

In addition, $\|y\|^2 + \|x\|^2 = \|x\|^2 = 1$. If $\|y\|^2, \|z\|^2 \geq \Omega(1)$, then we are done. So, suppose $\|y\|^2 < 1/4$ (and $\|z\|^2 \geq 3/4$.

Then, all we need to do is to shift x a little bit. Let $x' = x + 1/2\sqrt{n}$, and let as before $y' = \max\{0, x'\}$ and $z' = \min\{0, x'\}$. All we need to show is that $\|y'\|^2, \|z'\|^2 \geq \Omega(1)$. First of all, since $x \perp 1$, $\|x'\|^2 = \|x\|^2 + 1/4 = 5/4$. Also, since $\|z'\|^2 \leq \|x\|^2 \leq 1$ and $\|z'\|^2 + \|y'\|^2 = 5/4$, we have $\|y'\|^2 \geq 1/4$. Finally,

$$3/4 \leq \|z\|^2 = \sum_v z(v)^2 \leq \sum_v \left(z'(v) - \frac{1}{2\sqrt{n}}\right)^2 \leq \sum_v 2z'(v)^2 + \frac{1}{2} = 2\|z'\|^2 + 1/2,$$

so $\|z'\|^2 \geq 1/8$. $\square$

From now on, we only need to round a nonnegative vector to a set of small conductance.

Lemma 17.3. *For any nonnegative function $x : V \to \mathbb{R}$, there is a set S in the support of x such that*

$$\phi(S) \leq \sqrt{2\mathcal{R}(x)}.$$

Before proving the above lemma, we prove an easier statement, we show that if we drop all squares in the Rayleigh quotient we get an exact characterization of the conductance.

Lemma 17.4. *For every non-negative $x : V \to \mathbb{R}$, there is a set S in the support of x such that*

$$\phi(S) \leq \frac{\sum_{u \sim v} |x(v) - x(u)|}{d \cdot \sum_v x(v)}. \tag{17.1}$$

Proof. Since the right hand side is homogeneous in x, we may assume that $\max_{v\in V} x(v) \leq 1$. Let $t \sim [0,1]$ be chosen uniformly at random, and let $S_t := \{v : x(v) > t\}$. Then, by linearity of expectation,

$$\mathbb{E}\left[|S_t|\right] = \sum_v \mathbb{P}\left[v \in S_t\right] = \sum_v x(v).$$

Also, for any edge (u,v) the probability that this edge is cut is exactly $|x(u) - x(v)|$. Therefore,

$$\mathbb{E}\left[|E(S_t, \overline{S_t})|\right] = \sum_{u\sim v} \mathbb{P}\left[(u,v) \text{ is cut}\right] = \sum_{u\sim v} |x(u) - x(v)|.$$

Therefore,

$$\frac{\mathbb{E}\left[|E(S_t, \overline{S_t})|\right]}{d \cdot \mathbb{E}\left[|S_t|\right]} = \frac{\sum_{u\sim v} |x(u) - x(v)|}{d \cdot \sum_v x(v)}.$$

Note that in general if $\frac{\mathbb{E}A}{\mathbb{E}B} \leq c$ it may be that $A/B > c$ with probability 1. For example, if $A = B = -1$ with probability $1/2$ and $A = 1, B = 2$ with probability $1/2$. Then, $\mathbb{E}A = \frac{\mathbb{E}A}{\mathbb{E}B} = 0$, but $A/B \geq 1/2$ with probability 1. But, if A, B are nonnegative (with probability 1), then $\frac{\mathbb{E}A}{\mathbb{E}B} \leq c$ implies that $A/B \leq c$ with a positive probability. This follows from the following simple inequality. For any set of nonnegative numbers $a_1, a_2, \ldots, a_m, b_1, b_2 \ldots, b_m \geq 0$, we have

$$\min_{1\leq i\leq m} \frac{a_i}{b_i} \leq \frac{a_1 + \ldots + a_m}{b_1 + \ldots + b_m} \leq \max_{1\leq i\leq m} \frac{a_i}{b_i}. \tag{17.2}$$

Therefore, there is a set S_t that satisfies the conclusion of the lemma. □

By the above lemma, to prove Lemma 17.3 all we need to do is to construct a vector y such that $\frac{\sum_{u\sim v} |y(u)-y(v)|}{\sum_v y(v)} \leq O(\sqrt{\mathcal{R}(x)})$, it is easy to see that $y = x^2$ satisfies this claim.

Proof of Lemma 17.3. Let $y(v) := y(v)^2$ for all $v \in V$. We use Lemma 17.4. So, all we need to show is that

$$\frac{\sum_{u\sim v} |y(u) - y(v)|}{d \cdot \sum_v y(v)} \leq \sqrt{2\mathcal{R}(x)}.$$

The proof simply follows from one application of Cauchy-Schwarz inequality.

$$\begin{aligned}
\frac{\sum_{u\sim v} |y(u) - y(v)|}{d \cdot \sum_v y(v)} &= \frac{\sum_{u\sim v} \left|x(u)^2 - x(v)^2\right|}{d \cdot \sum_v x(v)^2} \\
&= \frac{\sum_{u\sim v} |x(u) - x(v)| \cdot |x(u) + x(v)|}{d \cdot \sum_v x(v)^2} \\
&\leq \frac{\sqrt{\sum_{u\sim v} |x(u) - x(v)|^2} \cdot \sqrt{\sum_{u\sim v} |x(u) + x(v)|^2}}{d \cdot \sum_v x(v)^2} \\
&= \sqrt{\mathcal{R}(f)} \cdot \sqrt{\frac{\sum_{u\sim v} |x(u) + x(v)|^2}{d \cdot \sum_v x(v)^2}}
\end{aligned}$$

It remains to upper bound $\sum_{u\sim v} |x(u) + x(v)|^2$. We use the simple inequality $(a+b)^2 \leq 2a^2 + 2b^2$ that holds for any $a, b \in \mathbb{R}$. Therefore,

$$\sum_{u\sim v} |x(u) + x(v)|^2 \leq \sum_{u\sim v} 2x(u)^2 + 2x(v)^2 = \sum_v 2d \cdot x(v)^2.$$

□

17.2 Expander Mixing Lemma

Recall that a graph G is an expander if $\phi(G) \geq \Omega(1)$. By Theorem 17.1, G is an expander if and only if $\lambda_2 \geq \Omega(1)$. In addition, 1 is a simple linear time algorithm to test if a given graph is an expander. Expander graphs have many nice properties, for example, their diameter is $O(\log n)$, a simple *lazy* random walk on them mixes in time $O(\log n)$, and that they are approximately the same as a complete graph. The latter statement is known as the expander mixing lemma that we want to prove in this section.

Let $\tilde{\lambda}_1 \geq \tilde{\lambda}_2 \geq \ldots \tilde{\lambda}_n$ be the eigenvalues of the adjacency matrix of G. It is easy to see that $\tilde{\lambda}_i = d(1-\lambda_i)$. Therefore, G is an expander if $\tilde{\lambda}_2 \leq c \cdot d$ for some $c < 1$. Also, we write

$$\tilde{\lambda} = \max\{\tilde{\lambda}_2, |\tilde{\lambda}_n|\}.$$

We say G is a stronger expander if $\tilde{\lambda} = o(d)$.

Definition 17.5. *A d regular graph is a Ramanujan graph if $\tilde{\lambda} \leq 2\sqrt{d-1}$.*

It turns out that Ramanujan graphs are the strongest form of expanders. This is due to a theorem of Alon and Boppana,

Theorem 17.6 (Alon-Boppana). *For any d regular graph*

$$\tilde{\lambda} \geq 2\sqrt{d-1} - \epsilon$$

where $\epsilon \to 0$ as $n \to \infty$.

On the other hand, a random d-regular graph is almost Ramanujan

Theorem 17.7 ([Fri04]). *In a random d-regular, with high probability, $\tilde{\lambda} \leq 2\sqrt{d-1} + \epsilon$.*

For a long time it was an open problem to give explicit construction of Ramanujan graphs. Lubotzky, Philips and Sarnak [LPP88] and independently Margulis [] gave explicit construction Ramanujan graphs when d is a prime number. It remains an open problem to construct (or prove the existence of) Ramanujan graphs of arbitrary degree. Marcus, Spielman and Srivastava used the method of interlacing polynomials to prove the existence of bipartite Ramanujan graphs of arbitrary degree [MSS13]. Note that, the bipartiteness is because of a limitation of the method of interlacing polynomials, that is one can only bound one eigenvalue.

The Alon's expander mixing lemma shows that (strong) expanders are very similar to random d-regular graphs.

Theorem 17.8 (Expander Mixing Lemma). *For any d regular graph G, and disjoint subsets of vertices S, T*

$$\left| |E(S,T)| - \frac{d}{n} \cdot |S| \cdot |T| \right| \leq \tilde{\lambda} \cdot \sqrt{|S| \cdot |T|}.$$

Note that $\frac{d}{n}|S| \cdot |T|$ is the expected number of edges between S, T in a random d-regular graph. For example, if G is a Ramanujan graph and $|S|, |T| \geq \Omega(n)$, then $|E(S,T)|$ is the same as the expected number of edges between S, T in a random graph up to a $O(\sqrt{d \cdot n})$ additive error. Next, we prove the above theorem, the proof is based on the lecture notes of Trevisan.

Proof. Let J be the all ones matrix. We can write

$$\begin{aligned} |E(S,T)| &= \mathbf{1}_S^\intercal A \mathbf{1}_T \\ \frac{d}{n} \cdot |S| \cdot |T| &= \frac{d}{n} \mathbf{1}_S^\intercal J \mathbf{1}_T, \end{aligned}$$

where as usual A is the adjacency matrix. Therefore,

$$\begin{aligned} \left| |E(S,T)| - \frac{d}{n} \cdot |S| \cdot |T| \right| &= \left| \mathbf{1}_S^{\intercal}(A - \frac{d}{n}J)\mathbf{1}_T \right| \\ &\le \|\mathbf{1}_S\| \left\| A - \frac{d}{n}J \right\| \|\mathbf{1}_T\| \end{aligned}$$

where for any matrix B, $\|B\| = \max_x \frac{\|Bx\|}{\|x\|}$, of $\|B\|$ is the largest eigenvalue of B. In particular, in the last inequality we use that $\|(A - dJ/n)\mathbf{1}_T\| \le \|A - dJ/n\| \cdot \|\mathbf{1}_T\|$, and that for any two vectors x, y, $x^{\intercal} y \le \|x\| \cdot \|y\|$.

Now, we have $\|\mathbf{1}_S\| = \sqrt{|S|}$, $\|\mathbf{1}_T\| = \sqrt{|T|}$, and the largest eigenvalue of $A - \frac{d}{n}J$ is $\tilde{\lambda}$ which proves the claim. □

17.3 The Sparsest Cut Problem

In the above we showed that $\min_{x \perp \mathbf{1}} \mathcal{R}(x)$ is a relaxation of the sparsest problem (up to a factor of 2). We can eliminate the pesky condition $x \perp 1$ by looking at $\sum_{u,v}(x(u) - x(v))^2$ in the denominator.

$$\min_{x \perp \mathbf{1}} \mathcal{R}(x) = \min_x \frac{\sum_{u \sim v}(x(u) - x(v))^2}{\frac{d}{n}\sum_{u,v}(x(u) - x(v))^2}$$

To see the identity note that $\sum_{u,v}(x(u) - x(v))^2 = n\sum_v x(v)^2 - (\sum_v x(v))^2$. So, to maximize the denominator of the RHS we need x to be orthogonal to $\mathbf{1}$. Since the numerator is invariant under shifting x, in the optimum we have $x \perp \mathbf{1}$.

We can generalize the above relaxation by looking at the minimum over any set of vectors of $\mathbb{R}^m$ assigned to the vertices.

$$\min_x \frac{\sum_{u \sim v}(x(u) - x(v))^2}{\frac{d}{n}\sum_{u,v}(x(u) - x(v))^2} = \min_{F: V \to \mathbb{R}^m} \frac{\sum_{u \sim v} \|F(u) - F(v)\|^2}{\frac{d}{n}\sum_{u,v} \|F(u) - F(v)\|^2}$$

It is easy to see that the optimum vector F is one dimensional and equal to the second eigenvector of $\mathcal{L}_G$.

The vector formulation allows us to write a SDP relaxation for the sparsest cut problem. Let X be the matrix of inner products, i.e., let $X(u,v) = \langle F(u), F(v) \rangle$ for all $u, v \in V$. Then, $X \preceq 0$. In addition for any pair of vertices u, v,

$$X \bullet L_{u,v} = X(u,u) + X(v,v) - X(u,v) - X(v,u) = \|F(u) - F(v)\|^2.$$

Therefore, $X \bullet L_G = \sum_{u \sim v} \|F(u) - F(v)\|^2$. So, we can write the following relaxation for the sparsest cut problem.

$$\begin{aligned} \min \quad & X \bullet L_G \\ \text{subject to} \quad & \frac{d}{n} X \bullet L_K = 1, \\ & X \succeq 0. \end{aligned} \tag{17.3}$$

Note that L_k is the Laplacian matrix of a complete graph on n vertices. Similar to the Cheeger's inequality the value of the above relaxation can be as small as $\phi(G)^2$.

It is instructive to understand the dual of the above relaxation. It is

$$\begin{aligned} \max \quad & \gamma, \\ \text{subject to} \quad & \gamma \cdot \frac{d}{n} L_K \preceq L_G. \end{aligned} \tag{17.4}$$

The optimum is simply the second smallest eigenvalue of L_G.

Next, we use the metric property of the optimum to strengthen (17.3). First, observe that the optimum is a 0/1 vector. Therefore, if x is the indicator vector of the optimum set, $x(u) - x(v)$ is either zero or one. So, for any triple of vertices u, v, w,

$$(x(u) - x(v))^2 \leq (x(u) - x(w))^2 + (x(w) - x(v))^2.$$

Leighton and Rao wrote an LP relaxation using the metric property and showed that its integrality gap is $O(\log n)$.

This is also known as the squared triangle inequality. Since the optimum satisfies this we can strengthen (17.3) by enforcing that the distance function must be a metric.

$$\begin{aligned} \min \quad & X \bullet L_G \\ \text{subject to} \quad & \frac{d}{n} X \bullet L_k = 1, \\ & X \bullet L_{u,v} \leq X \bullet L_{u,w} + X \bullet L_{w,v}, \quad \forall u, v, w, \\ & X \succeq 0. \end{aligned} \tag{17.5}$$

The above relaxation is also known as the Goemans, Linial's relaxation, or the Arora, Rao and Vazirani's relaxation.

Next, we want to write the dual of the above relaxation. First, note that we can rewrite the triangle inequality constraint as follows. For any pair of vertices u, v and any path P in G from u to v,

$$X \bullet L_{u,v} \leq X \bullet L_P,$$

where as usual $L_P = \sum_{e \in P} L_e$. In the dual, we will have one variable y_p for each path from u, v.

$$\begin{aligned} \max \quad & \gamma, \\ \text{subject to} \quad & \gamma \cdot \frac{d}{n} L_k \preceq L_G + \sum_{u,v} \sum_{P \in P_{u,v}} y_P (L_{u,v} - L_P), \\ & y_P \geq 0 \qquad \qquad \qquad \qquad \qquad \forall P. \end{aligned} \tag{17.6}$$

where $P_{u,v}$ is the set of all paths from u to v in G.

First, observe that if $y_P = 0$ for all P we get the same SDP relaxation (17.3). Roughly speaking, y_P shortcuts the long paths of G to increase the second eigenvalue of (normalized Laplacian of) G and to improve the performance of Cheeger's inequality. Let us give a clarifying example. Consider the cycle C_n. Recall that $\lambda_2 \approx 1/n^2$ and $\phi(G) \approx 1/n$. Suppose we add the following edges $(1, \sqrt{n}), (\sqrt{n}, 2\sqrt{n}), \ldots, (n - \sqrt{n}, n)$ all with weight $1/2$ and we decrease the weight of every original edge to $1/2$ (see Figure 17.1) and let H be the new graph. It is easy to see that the size of every cut in H is at most what it was in C_n, so $\phi(H) \leq \phi(C_n)$. On the other hand, it is not hard to see that the second smallest eigenvalue of $\mathcal{L}_H$ is $\Omega(1/n)$. Therefore, by Cheeger's inequality

$$\phi(C_n) \geq \phi(H) \geq \lambda_2(\mathcal{L}_H)/2 \geq \Omega(1/n).$$

This shows that SDP (17.3), the optimum value of (17.6) is $\Omega(1/n)$ for C_n.

In general, the matrix

$$L_G + \sum_{u,v} \sum_{P \in P_{u,v}} y_p (L_{u,v} - L_P)$$

is less than L_G across all cuts. We write $A \preceq_C B$ if for any set $S \subset V$,

$$\mathbf{1}_S^\intercal A \mathbf{1}_S \leq \mathbf{1}_S^\intercal B \mathbf{1}_S.$$

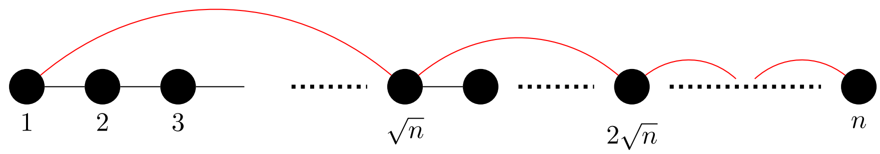

Figure 17.1: The weight of every (black and red) edge in the above graph is 1/2. The size of every cut is at most the size of the corresponding cut of C_n. So, $\phi(G) \leq \phi(C_n)$ is smaller. But the second eigenvalue of the $\mathcal{L}_G$ is $\Omega(1/n)$.

We can write,

$$L_G + \sum_{u,v} \sum_{P \in P_{u,v}} y_P (L_{u,v} - L_P) \preceq_C L_G.$$

This is because for any path P from u to v,

$$(L_{u,v} - L_P) \preceq_C 0,$$

or equivalently,

$$L_{u,v} \preceq_C L_P.$$

In words, in (17.6) we would like to route a multi-commodity flow along the paths P such that the second eigenvalue of L_G plus the Laplacian of the demand graph of the flow and minus of the Laplacian of the actual flow is as large as possible. Since this routing does not increase the value of cuts, the second eigenvalue of $L_G + \sum_{u,v} \sum_{P \in P_{u,v}} (L_{u,v} - L_P)$ lower bounds $\phi(G)$.

Arora, Rao and Vazirani called this flow an *expander flow*. This is because the performance of Cheeger's inequality is the best in an expander. So, ideally, if the demand graph of the flow is an expander, the ARV relaxation gives a constant factor approximation for sparsest cut. Arora, Rao and Vazirani showed that the integrality gap of (17.5) is $O(\sqrt{\log n})$ for any graph G [ARV09]. Note that loosing a factor 2 we can simply SDP (17.6) as follows.

$$\begin{aligned} \max \quad & \gamma, \\ \text{subject to} \quad & \gamma \cdot \frac{d}{n} L_k \preceq \sum_{u,v} \sum_{P \in P_{u,v}} L_P, \\ & \sum_{u,v} \sum_{P \in P_{u,v}} L_P \preceq L_G, \qquad y_P \geq 0 \quad \forall P. \end{aligned} \tag{17.7}$$

In fact the analysis of Arora, Rao, Vazirani also bounds the integrality gap of a simpler variant of the above SDP. In particular, instead of requiring the multi-commodity flow to be smaller than L_G in the PSD sense, we can simply ask that the congestion of every edge to be at most 1 in the multicommodity flow.

$$\begin{aligned} \max \quad & \gamma, \\ \text{subject to} \quad & \gamma \cdot \frac{d}{n} L_k \preceq \sum_{u,v} \sum_{P \in P_{u,v}} L_P, \\ & \sum_{P : e \in P} L_P \leq 1 \qquad \forall e, \\ & y_P \geq 0 \qquad \forall P. \end{aligned} \tag{17.8}$$

The above characterization of the semidefinite programming relaxation of the sparsest cut has been used in designing fast primal dual approximation algorithms for this problem [AHK10; She09]

In the next lecture we will see this idea can be extended to reduce the effective resistance of the edges of a graph which makes the graph ready for Marcus, Spielman, Srivastava's theorem.

References

[AHK10] S. Arora, E. Hazan, and S. Kale. "O(sqrt(log(n)) Approximation to SPARSEST CUT in $\tilde{\mathrm{O}}(\mathrm{n}^2)$ Time". In: *SIAM J. Comput.* 39.5 (2010), pp. 1748–1771 (cit. on p. 17-8).

[AO14] N. Anari and S. Oveis Gharan. "Effective-Resistance-Reducing Flows and Asymmetric TSP". 2014. URL: http://arxiv.org/abs/1411.4613 (cit. on p. 17-1).

[ARV09] S. Arora, S. Rao, and U. Vazirani. "Expander flows, Geometric Embeddings and Graph Partitioning". In: *J. ACM* 56 (2 Apr. 2009), 5:1–5:37 (cit. on p. 17-8).

[Fri04] J. Friedman. "A proof of Alon's second eigenvalue conjecture and related problems". 2004. URL: http://arxiv.org/abs/cs/0405020 (cit. on p. 17-5).

[HLW06] S. Hoory, N. Linial, and A. Wigderson. "Expander graphs and their applications". In: *Bulletin of the American Mathematical Society* 43.4 (2006), pp. 439–561 (cit. on p. 17-1).

[JM85] S. Jimbo and A. Maruoka. "Expanders obtained from affine transformations". In: *STOC*. Providence, Rhode Island, United States: ACM, 1985, pp. 88–97 (cit. on p. 17-1).

[JSV04] M. Jerrum, A. Sinclair, and E. Vigoda. "A polynomial-time approximation algorithm for the permanent of a matrix with nonnegative entries". In: *J. ACM* 51.4 (July 2004), pp. 671–697 (cit. on p. 17-1).

[KVV04] R. Kannan, S. Vempala, and A. Vetta. "On clusterings: Good, bad and spectral". In: *J. ACM* 51.3 (May 2004), pp. 497–515 (cit. on p. 17-1).

[Kwo+13] T. C. Kwok, L. C. Lau, Y. T. Lee, S. O. Gharan, and L. Trevisan. "Improved Cheeger's inequality: analysis of spectral partitioning algorithms through higher order spectral gap". In: *STOC*. 2013, pp. 11–20 (cit. on p. 17-2).

[Lee12] J. R. Lee. http://tcsmath.wordpress.com/2012/04/18/gabber-galil-analysis-of-margulis-expanders/. 2012 (cit. on p. 17-1).

[LPP88] A. Lubotzky, R. Phillips, and S. P. "Ramanujan graphs". In: *Combinatorica* 9 (1988), pp. 261–277 (cit. on p. 17-5).

[MSS13] A. Marcus, D. A. Spielman, and N. Srivastava. "Interlacing Families I: Bipartite Ramanujan Graphs of All Degrees". In: *FOCS*. 2013, pp. 529–537 (cit. on p. 17-5).

[She09] J. Sherman. "Breaking the Multicommodity Flow Barrier for $\mathrm{O}(\sqrt{\log n})$-Approximations to Sparsest Cut". In: *FOCS*. 2009, pp. 363–372 (cit. on p. 17-8).

[SJ89] A. Sinclair and M. Jerrum. "Approximate counting, uniform generation and rapidly mixing Markov chains". In: *Inf. Comput.* 82.1 (July 1989), pp. 93–133 (cit. on p. 17-1).

[SM00] J. Shi and J. Malik. "Normalized Cuts and Image Segmentation". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22.8 (2000), pp. 888–905 (cit. on p. 17-1).

[ST96] D. A. Spielman and S.-H. Teng. "Spectral Partitioning Works: Planar Graphs and Finite Element Meshes". In: *FOCS*. 1996, pp. 96–105 (cit. on p. 17-1).

[TM06] D. A. Tolliver and G. L. Miller. "Graph Partitioning by Spectral Rounding: Applications in Image Segmentation and Clustering". In: *CVPR*. IEEE Computer Society, 2006, pp. 1053–1060 (cit. on p. 17-1).