University of Michigan–Ann Arbor

Department of Electrical Engineering and Computer Science

EECS 498 004 **Advanced Graph Algorithms**, Fall 2021

Lecture 9: Spectral Equivalences, Expander Mixing Lemma, and Ramanujan Graphs

September 30, 2021

Instructor: Thatchaphol Saranurak

Scribe: Nikhil Shagrithaya

1 Introduction

Last week, we saw that the eigenvalues of the normalized Laplacian of a graph can contain useful information about the 'connectedness' of a graph. In particular, we saw that the second smallest eigenvalue of ($\mathbf{N}_G$) is closely related to the conductance of the graph. The precise relationship is given by Cheeger's inequality:

$$\lambda_2(\mathbf{N}_G)/2 \leq \Phi(G) \leq \sqrt{2\lambda_2(\mathbf{N}_G)}$$

Moreover, we also saw an algorithm which gave us a $O(1)$-approximation of the sparest cut, when the conductance of the graph was $\Omega(1)$. That is, using the power method, we were able to obtain a method cut for which the following was true:

$$\Phi_G(S) \approx \sqrt{2\lambda_2(\mathbf{N}_G)} \leq 2\sqrt{\Phi(G)}$$

2 Spectral equivalency of graph

We can also define a new notion of connectivity for graphs based on the eigenvalues of the normalized Laplacian itself. In order to do this, we need some definitions.

Definition 2.1. We define a partial order on graphs based on their eigenvalues. We say that:

$$H \preccurlyeq^{\text{spec}} G$$

if $\boldsymbol{x}^T\mathbf{L}_H\boldsymbol{x} \leq \boldsymbol{x}^T\mathbf{L}_G\boldsymbol{x}$ for all $\boldsymbol{x} \in \mathbb{R}^V$. Or equivalently, $\mathbf{L}_H \preceq \mathbf{L}_G$ or $\mathbf{L}_G - \mathbf{L}_H$ is psd (positive semi-definite).

Definition 2.2. We say that $\mathbf{H}$ and $\mathbf{G}$ are α**-spectral-equivalent** if $\mathbf{H}' \preccurlyeq^{\text{spec}} \mathbf{G} \preccurlyeq^{\text{spec}} \alpha\mathbf{H}'$ for some $\mathbf{H}' = c \cdot \mathbf{H}$. We write $\mathbf{H} \approx_{\alpha}^{\text{spec}} \mathbf{G}$.

Note that $\preccurlyeq^{\text{spec}}$ is stronger than $\preccurlyeq^{\text{cut}}$, because we can define the partial order based on cut in another, equivalent way: $H \preccurlyeq^{\text{cut}} G$ iff $\boldsymbol{x}^T \mathbf{L}_H \boldsymbol{x} \leq \boldsymbol{x}^T \mathbf{L}_G \boldsymbol{x}$ for all $\boldsymbol{x} \in \{0,1\}^V$. Because the partial order based on eignenvalues requires $\boldsymbol{x}^T \mathbf{L}_H \boldsymbol{x} \leq \boldsymbol{x}^T \mathbf{L}_G \boldsymbol{x}$ for all $\boldsymbol{x} \in \mathbb{R}^V$, and not just $\boldsymbol{x} \in \{0,1\}^V$, $\preccurlyeq^{\text{spec}}$ is a stronger requirement than $\preccurlyeq^{\text{cut}}$, namely $H \preccurlyeq^{\text{spec}} G$ implies $H \preccurlyeq^{\text{cut}} G$.

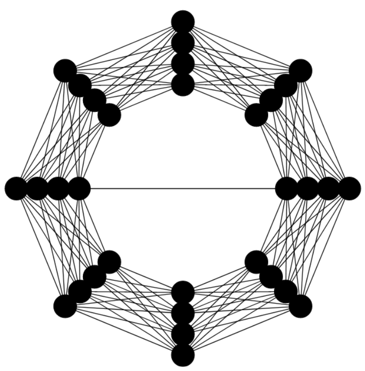
A natural question to then ask, is whether the converse is true, i.e. $H \preccurlyeq^{\text{cut}} G$ implies $H \preccurlyeq^{\text{spec}} G$? And if not, does it hold approximately?

We'll see that the answer is no to both of these questions, in the following lemma:

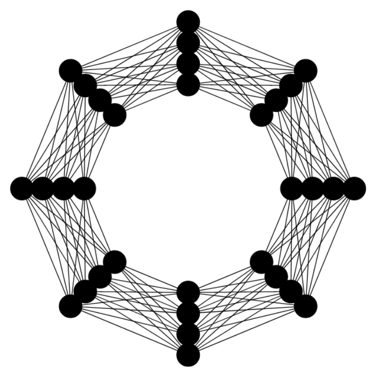
Lemma 2.3 ($\preccurlyeq^{\text{spec}}$ is strictly stronger than $\preccurlyeq^{\text{cut}}$)**.** *There exist graphs G and $\tilde{G}$ such that:*

- $\tilde{G} \preccurlyeq^{\text{cut}} G \preccurlyeq^{\text{cut}} (1+\epsilon)\tilde{G}$, *but*
- $\tilde{G} \preccurlyeq^{\text{spec}} G$ *but* $G \not\preccurlyeq^{\text{spec}} \frac{\epsilon^2 n}{10000}\tilde{G}$

Proof. We will construct unweighted graphs G and $\tilde{G}$ with two parameters: n and $k = \frac{50}{\epsilon}$. Both graphs will have $n \times k$ vertices, having n clusters with k vertices each. The vertex set of $\tilde{G}$ is $\{0, 1, \ldots, n-1\} \times \{1, \ldots, k\}$, where n is even. The graph $\tilde{G}$ will consist of n complete bipartite graphs, connecting all pairs of vertices $\{u, i\}$ and $\{v, j\}$ where $v = u + 1 mod n$. G is identical $\tilde{G}$, except that it has one extra edge going from vertex $\{0, 1\}$ to vertex $\{n/2, 1\}$. The graphs for $n = 8$ and $k = 4$ look like this:



G: $n = 8$ sets of $k = 4$ vertices arranged in a ring and connected by complete bipartite graphs, plus one edge across.



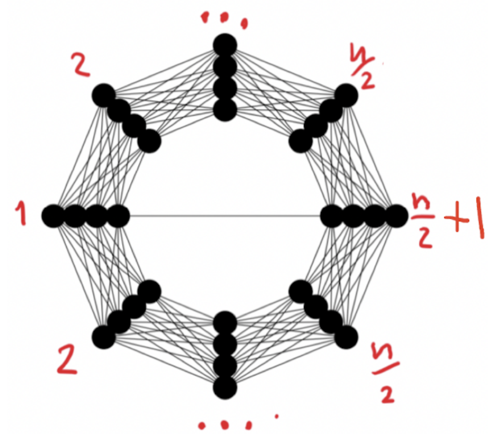
$\tilde{G}$: A good cut sparsifier of G, but a poor spectral sparsifier

[1]

It is easy to see that $\tilde{G} \preccurlyeq^{\text{cut}} G \preccurlyeq^{\text{cut}} (1+\epsilon)\tilde{G}$, because the only cuts whose value will change are those where the extra edge is crossing between them. Because any cut must have size at least k^2, adding one extra edge does not change the cut value by much.

Moreover, $\tilde{G} \preccurlyeq^{\text{spec}} G$ holds because $\tilde{G}$ is a subgraph of G. All that remains is to show $G \not\preccurlyeq^{\text{spec}} \frac{\epsilon^2 n}{10000}\tilde{G}$. Consider the following $\boldsymbol{x}$:

[1]Image from https://arxiv.org/abs/0808.4134



where $\forall u \in \{0, 1, \cdots, n/2\}$, $\forall i \in [k]$, $x_{\{u,i\}} = u + 1$; $\forall u \in \{n/2+1, \cdots, n-1\}$ $\forall i \in [k]$, $x_{\{u,i\}} = n + 1 - u$. Then, we have:

$$\boldsymbol{x}^T \boldsymbol{L}_G \boldsymbol{x} = nk^2 + \left(\frac{n}{2}\right)^2$$

but

$$\boldsymbol{x}^T \boldsymbol{L}_{\tilde{G}} \boldsymbol{x} = nk^2$$

So $\boldsymbol{x}^T \boldsymbol{L}_G \boldsymbol{x} - \frac{\epsilon^2 n}{10000} \boldsymbol{x}^T \boldsymbol{L}_{\tilde{G}} \boldsymbol{x} = nk^2 > 0 \implies G \not\preccurlyeq^{\text{spec}} \frac{\epsilon^2 n}{10000} \tilde{G}$. □

This lemma shows that spectral-equivalence between graphs are *strictly stronger* than cut-equivalence, and moreover, cut-equivalence cannot even imply a slightly weaker spectral-equivalence, as was the case for flow, because the lemma shows a gap of $\Omega(n)$, which is the largest possible gap.

As an aside, it is easy to check if $H \preccurlyeq^{\text{spec}} G$, because all we need to do is check whether $\boldsymbol{L}_G - \boldsymbol{L}_H$ is positive semi-definite, and this can be done in polynomial time.

Question 2.4 (Open)**.** Is there a **fast** algorithm for checking if $H \preccurlyeq^{\text{spec}} G$ or even $H \approx^{\text{spec}}_{(1+\epsilon)} G$?

Question 2.5 (Open)**.** Is there a **polynomial-time** algorithm for checking if $H \preccurlyeq^{\text{cut}} G$ or even $H \approx^{\text{cut}}_{(1+\epsilon)} G$?

From the previous lecture, checking if $H \preccurlyeq^{\text{flow}} G$ is checking the existence of feasible H-concurrent flow in G, which can be done by solving LP in polynomial time. Since $H \preccurlyeq^{\text{flow}} G \implies H \preccurlyeq^{\text{cut}} G$, Q2.5 can also be done in polytime. Remains to check the existence of faster algorithms for checking $H \preccurlyeq^{\text{flow}} G$ or $H \preccurlyeq^{\text{cut}} G$.

3 Spectral Equivalence between Expanders

In Lecture 03_2, we saw that expanders with the same degree profile are approximately equivalent, in both the cut and the flow sense. But are they also spectral-equivalent?

They indeed are. We begin with the following lemma:

Lemma 3.1. *Suppose G is a ϕ-expander. If $H \preccurlyeq^{\text{deg}} G$, then $H \preccurlyeq^{\text{spec}} O(\frac{1}{\phi^2})G$.*

Proof. We can increase the degree of any vertex in H to be equal to the degree of the corresponding vertex in G by adding self-loops, this will only make our task easier, and importantly, $\mathbf{L}_H$ remains the same. Now, both H and G have the same degree profile $\boldsymbol{d}$ (the degree profile is a vector whose entries are the degrees of the vertices of the graph). For any $\boldsymbol{x} \perp \boldsymbol{d}$, we have:

$$\boldsymbol{x}^\top \boldsymbol{L}_G \boldsymbol{x} \geq \frac{\phi^2}{4} \boldsymbol{x}^\top \boldsymbol{L}_H \boldsymbol{x}.$$

The above statement is true because, by Cheeger's, $\boldsymbol{x}^\top \boldsymbol{L}_G \boldsymbol{x} \geq \frac{\phi^2}{2} \boldsymbol{x}^\top \boldsymbol{D} \boldsymbol{x}$ for any $\boldsymbol{x} \perp \boldsymbol{d}$ (recall that $\lambda_2(\boldsymbol{N}_G) = \min_{\boldsymbol{x} \perp \boldsymbol{d}} \frac{\boldsymbol{x}^\top \boldsymbol{L}_G \boldsymbol{x}}{\boldsymbol{x}^\top \boldsymbol{D} \boldsymbol{x}}$). Also, for any $\boldsymbol{x}$, we claim $2\boldsymbol{x}^\top \boldsymbol{D} \boldsymbol{x} \geq \boldsymbol{x}^\top \boldsymbol{L}_H \boldsymbol{x}$ because:

$$\begin{aligned}
2\boldsymbol{y}^\top \boldsymbol{D} \boldsymbol{y} - \boldsymbol{y}^\top \boldsymbol{L}_H \boldsymbol{y} &= \sum_u 2 \deg_G(u) y_u^2 - \sum_{uv \in E} w_{uv}(y_u - y_v)^2 = \sum_u 2 \deg_H(u) y_u^2 - \sum_{uv \in E} w_{uv}(y_u - y_v)^2 \\
&= \sum_{uv \in E} w_{uv}(y_u + y_v)^2 \geq 0.
\end{aligned}$$

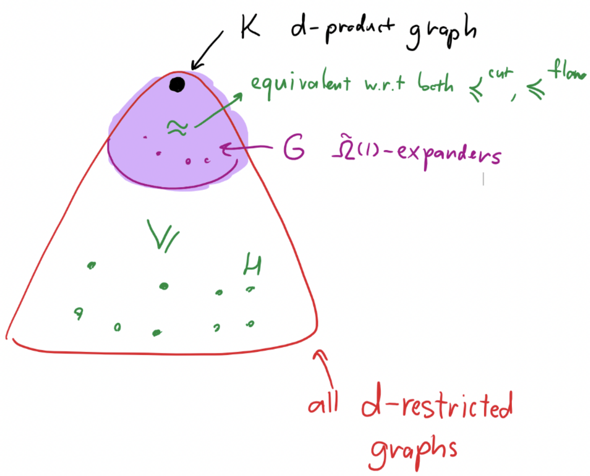
Now we have to show the same inequality also holds for all vectors $\boldsymbol{x}$ such that $\boldsymbol{x} \parallel \boldsymbol{d}$. Let $\boldsymbol{y} = \boldsymbol{x} - \sigma \mathbf{1}$ where $\boldsymbol{y} \perp \boldsymbol{d}$. Then:

$$\boldsymbol{y}^\top \boldsymbol{L}_G \boldsymbol{y} = (\boldsymbol{x} - \sigma\mathbf{1})^\top \boldsymbol{L}_G (\boldsymbol{x} - \sigma\mathbf{1}) = \boldsymbol{x}^T \boldsymbol{L}_G \boldsymbol{x} - \sigma \mathbf{1}^T \mathbf{L}_G \boldsymbol{x} - \boldsymbol{x}^T \mathbf{L}_G \sigma \mathbf{1} + \sigma \mathbf{1}^T \mathbf{L}_G \sigma \mathbf{1}$$

The last two terms are zero because $\mathbf{1}$ is in the kernel of L_G. The second term is also zero, because $\sigma \mathbf{1}^T \mathbf{L}_G \boldsymbol{x} = (\mathbf{L}_G^T \sigma \mathbf{1})^T \boldsymbol{x} = (\mathbf{L}_G^\sigma \mathbf{1})^T \boldsymbol{x} = 0$, where the second equality is true because $\mathbf{L}_G$ is symmetric. A similar argument holds for $\mathbf{L}_H$. Combining everything, we have that $\boldsymbol{x}^\top \boldsymbol{L}_G \boldsymbol{x} = \boldsymbol{y}^\top \boldsymbol{L}_G \boldsymbol{y}$ and $\boldsymbol{x}^\top \boldsymbol{L}_H \boldsymbol{x} = \boldsymbol{y}^\top \boldsymbol{L}_H \boldsymbol{y}$, and because we've already shown that $\boldsymbol{x}^\top \boldsymbol{L}_G \boldsymbol{x} \geq \frac{\phi^2}{4} \boldsymbol{x}^\top \boldsymbol{L}_H \boldsymbol{x}$ when $\boldsymbol{x} \perp \boldsymbol{d}$, we are done. ∎

Corollary 3.2. *Let G and G' be ϕ-expanders with degree profile $\boldsymbol{d}$. We have:*

- *$G \preccurlyeq^{\text{spec}} O(\frac{1}{\phi^2}) G'$ and $G' \preccurlyeq^{\text{spec}} O(\frac{1}{\phi^2}) G$. So G and G' are $O(\frac{1}{\phi^4})$-spectral-equivalent.*
- *For all $\boldsymbol{d}$-restricted graphs H, we have $H \preccurlyeq^{\text{spec}} O(\frac{1}{\phi^2}) G$.*
- We have same picture we saw. It holds for for $\preccurlyeq^{\text{spec}}$



4 Expander Mixing Lemma

4.1 Motivation

Let G be a graph with degree profile $\boldsymbol{d}$. If G has high conductance, then we have:

$$|E_G(S, V \setminus S)| \approx \boldsymbol{d}(S)$$

for all sets S such that $vol(S) \leq vol(V \backslash S)$. But the good conductance only guarantees us that the number of edges going out of a set is high. But what can we say about the number of edges **between two sets ?**

For the $\boldsymbol{d}$-product graph, it can be easily seen that:

$$|E(S, T)| \approx \frac{\boldsymbol{d}(S)\boldsymbol{d}(T)}{\boldsymbol{d}(V)}$$

for all $S, T \subseteq V$. Intuitively, one should think about $\frac{\boldsymbol{d}(S)\boldsymbol{d}(T)}{\boldsymbol{d}(V)}$ as the expectation of $|E(S, T)|$ in the random graph with degree profile $\boldsymbol{d}$. This is a stronger property than conductance, as conductance guarantees this only when $T = V \setminus S$.

A natual question to then ask, is: do expanders satisfies even this stronger property? We will see that this stronger property has applications.

4.2 The Normalized Adjacency Matrix

We begin by defining a variant of the Laplacian matrix, whose eigenvalues will give us a good approximation of the property we defined in the previous subsection. Recall that we previously defined:

$$\boldsymbol{N}_G = \boldsymbol{I} - \boldsymbol{D}^{-1/2}\boldsymbol{A}_G\boldsymbol{D}^{-1/2}$$

, which had eigenvalues satisfying:

$$0 = \lambda_1(\boldsymbol{N}_G) \leq \lambda_2(\boldsymbol{N}_G) \leq \cdots \leq \lambda_n(\boldsymbol{N}_G) \leq 2.$$

We will call

$$\boldsymbol{A}'_G = \boldsymbol{D}^{-1/2}\boldsymbol{A}_G\boldsymbol{D}^{-1/2}$$

as the normalized adjacency matrix. Note that $\boldsymbol{A}'_G = \boldsymbol{I} - \boldsymbol{N}_G$ has eigenvalues satisfying:

$$1 = \lambda_1(\boldsymbol{A}'_G) \geq \lambda_2(\boldsymbol{A}'_G) \geq \cdots \geq \lambda_n(\boldsymbol{A}'_G) \geq -1.$$

The following properties follow from the corresponding ones we proved earlier, about $\mathbf{N_G}$:

- $\lambda_2(\boldsymbol{A}'_G) = 1$ iff G is not connected.
- $\lambda_2(\boldsymbol{A}'_G) < 1 - \Omega(1)$ iff G is a $\Omega(1)$-expander.
- $\lambda_n(\boldsymbol{A}'_G) = -1$ iff G is bipartite.

We further define σ_2 as follows:

$$\sigma_2 := \max\{|\lambda_2(\boldsymbol{A}'_G)|, |\lambda_3(\boldsymbol{A}'_G)|, \ldots, |\lambda_n(\boldsymbol{A}'_G)|\} = \max\{|\lambda_2(\boldsymbol{A}'_G)|, |\lambda_n(\boldsymbol{A}'_G)|\}.$$

It is the second largest eigenvalue in absolute value of $\boldsymbol{A}'_G$. So from what we learned before, if $\sigma_2 \ll 1$ is small, then G is both "far" from being disconnected and bipartite. Moreover, in the following section, we will see that σ_2 describes when two arbitrary disjoint sets are chosen in the graph, how much the edges between them behave like these of $\boldsymbol{d}$-product graph. The smaller σ_2 is, the more similar they are. We will be formalizing this statement, and make it mathematically precise, in the following sections.

4.3 The Statement and the Proof

Lemma 4.1 (Expander Mixing Lemma)**.** *For a graph G with degree profile $\mathbf{d}$, and for every $S, T \subseteq V$, we have*

$$\left| |E(S,T)| - \frac{\boldsymbol{d}(S)\boldsymbol{d}(T)}{\boldsymbol{d}(V)} \right| \leq \sigma_2 \sqrt{\boldsymbol{d}(S)\boldsymbol{d}(T)}.$$

Therefore, when σ_2 is small, $|E(S,T)| \approx \frac{\boldsymbol{d}(S)\boldsymbol{d}(T)}{\boldsymbol{d}(V)}$ holds for all $S, T \subseteq V$.

Proof. We will show a slightly more general statement, which is:

$$\sigma_2 = \max_{x,y} \frac{|\boldsymbol{x}^\top(\boldsymbol{A}_G - \boldsymbol{R})\boldsymbol{y}|}{\sqrt{\sum_v \boldsymbol{d}(v)x_v^2} \cdot \sqrt{\sum_v \boldsymbol{d}(v)y_v^2}}$$

where $\boldsymbol{R}$ is equal to $\frac{1}{\boldsymbol{d}(V)}\boldsymbol{d} \cdot \boldsymbol{d}^\top$. Looking at $\boldsymbol{R}$ entry-wise, we get $\boldsymbol{R}_{u,v} = \frac{\boldsymbol{d}(u)\boldsymbol{d}(v)}{\boldsymbol{d}(V)}$. So in other words, $\boldsymbol{R}$ is the adjacency matrix of the $\boldsymbol{d}$-product graph. Now, observe that if $\boldsymbol{x} = \mathbf{1}_S$ and $\boldsymbol{y} = \mathbf{1}_T$, we get $\boldsymbol{x}^\top\boldsymbol{A}_G\boldsymbol{y} = |E(S,T)|$, and $\boldsymbol{x}^\top\boldsymbol{R}\boldsymbol{y}$ equals $\boldsymbol{d}(S)\boldsymbol{d}(T)/\boldsymbol{d}(V)$. Therefore, we have:

$$\sigma_2 \geq \frac{\left| |E(S,T)| - \frac{\boldsymbol{d}(S)\boldsymbol{d}(T)}{\boldsymbol{d}(V)} \right|}{\sqrt{\boldsymbol{d}(S)} \cdot \sqrt{\boldsymbol{d}(T)}}$$

which gives the lemma.

We can rewrite $\boldsymbol{A}'_G$ by making use of eigenvalue decomposition:

$$\boldsymbol{A}'_G = \boldsymbol{v}_1\boldsymbol{v}_1^\top + \lambda_2(\boldsymbol{A}'_G)\boldsymbol{v}_2\boldsymbol{v}_2^\top + \cdots + \lambda_n(\boldsymbol{A}'_G)\boldsymbol{v}_n\boldsymbol{v}_n^\top.$$

By how we defined it, $\lambda_1(\boldsymbol{A}'_G) = 1$. Moreover, $\boldsymbol{v}_1 = \frac{1}{\sqrt{\boldsymbol{d}(V)}}\boldsymbol{d}^{1/2}$. This we can see by combining the fact that $\boldsymbol{A}'_G = \boldsymbol{I} - \boldsymbol{N}_G$ has the same eigenvectors as $\boldsymbol{N}_G$, and that the first eigenvector of $\boldsymbol{N}_G$ is $\frac{1}{\sqrt{\boldsymbol{d}(V)}}\boldsymbol{d}^{1/2}$.

So we have:

$$\begin{aligned}
\sigma_2 &= \max_{z,w} \frac{|z^\top (A'_G - v_1 v_1^\top) w|}{\|z\| \cdot \|w\|} \\
&= \max_{x,y} \frac{|x^\top D^{1/2} (A'_G - v_1 v_1^\top) D^{1/2} y|}{\sqrt{\sum_v d(v) x_v^2} \cdot \sqrt{\sum_v d(v) y_v^2}} && z \mapsto D^{1/2} x, w \mapsto D^{1/2} y \\
&= \max_{x,y} \frac{|x^\top (A_G - R) y|}{\sqrt{\sum_v d(v) x_v^2} \cdot \sqrt{\sum_v d(v) y_v^2}}
\end{aligned}$$

The first equality is by the *variational characterization of singular values* (we prove this below). The last equality is because $A'_G = D^{-1/2} A_G D^{-1/2}$, and we can verify that

$$\begin{aligned}
D^{1/2} v_1 v_1^\top D^{1/2} &= D^{1/2} \left(\frac{1}{\sqrt{d(V)}} d^{1/2}\right) \left(\frac{1}{\sqrt{d(V)}} d^{1/2}\right)^\top D^{1/2} \\
&= \frac{1}{d(V)} d \cdot d^\top \\
&= R
\end{aligned}$$

This completes the proof.

□

Corollary 4.2. *Let G be a d-regular graph. For every $S, T \subseteq V$, we have*

$$\left| |E(S,T)| - \frac{d \cdot |S| \cdot |T|}{dn} \right| \leq \sigma_2 d \sqrt{|S| \cdot |T|}.$$

It turns out that the converse of the Expander Mixing Lemma also holds.

Lemma 4.3. [2]*Let G be a d-regular graph. If*

$$\left| |E(S,T)| - \frac{d \cdot |S| \cdot |T|}{dn} \right| \leq \theta d \sqrt{|S| \cdot |T|}$$

for all two disjoint sets S, T. Then $\sigma_2 = O(\theta(1 + \log(\frac{d}{\theta}))$.

5 Ramanujan Graphs: The Best Expander w.r.t. Eigenvalues

To simplify the discussion, we will only talk about d-regular in this section. A natural question to ask at this point is: How small σ_2 can be? A smaller σ_2 means that $E(S,T)$ is closer to $\frac{d|S||T|}{dn}$, for all $S, T \subseteq V$.

It turns out that σ_2 can be as small as $\frac{2\sqrt{d-1}}{d} \leq \frac{2}{\sqrt{d}}$. Graphs that achieve this bound are called **Ramanujan graphs**. Furthermore, this bound is tight! Roughly, we have that:

[2]https://link.springer.com/article/10.1007/s00493-006-0029-7

Theorem 5.1. *For every n_0 and d_0, there exist $n \in [n_0, O(n_0)]$ and $d \in [d_0, O(d_0)]$ such that, there is a d-regular graph G with n vertices satisfying:*

$$\sigma_2 \leq \frac{2\sqrt{d-1}}{d} = O(\frac{1}{\sqrt{d}}).$$

That is, its second eigenvalue (in absolute values) of the normalized adjacency matrix is at most $\frac{2\sqrt{d-1}}{d}$.

Question 5.2 (Open)**.** Can we prove that this holds for all n and d?

Adam Marcus, Daniel Spielman and Nikhil Srivatsava [3] proved that it holds for bipartite graphs. That is, $\lambda_n(A'_G) = -1$, but $\max_{i \neq 1,n} \lambda_i(A'_G) \leq \frac{2\sqrt{d-1}}{d}$. This led to the resolution of the *Kardison Singer problem*[4]

5.1 Tightness of Ramanujan Graphs

One can ask whether Ramanujan expanders are the best expanders with respect to eigenvalues. That is: does there exist another family of graphs for which σ_2 is strictly smaller than $\frac{2\sqrt{d-1}}{d}$?

It is known that this is not the case, it has been shown that $\sigma_2 \geq \frac{2\sqrt{d-1}}{d}(1 - o(1))$.[5] That is, Ramanujan graphs are the best expanders w.r.t. eigenvalues. In this lecture, we will show a slightly less general version of this bound.

Lemma 5.3. *For any d-regular graph where $d \leq 0.99 \cdot n$, we have $\sigma_2 \geq \Omega(\frac{1}{\sqrt{d}})$.*

Proof. We have

$$\begin{aligned}\mathrm{Tr}[(A'_G)^2] &= \sum_{i=1}^{n} \lambda_i(A'_G)^2 \\ &\leq 1 + (n-1)\sigma_2^2\end{aligned}$$

upon rearranging the terms, we get:

$$\sigma_2 \geq \sqrt{\frac{\mathrm{Tr}[(A'_G)^2] - 1}{(n-1)}}$$

Note $(A_G^2)_{u,v}$ is just the number of 2-step walks from u to v. So $\mathrm{Tr}[A_G^2] = nd$. Also, when G is d-regular, we have $A'_G = A_G/d$, and therefore, $\mathrm{Tr}[(A'_G)^2] = (nd)/d^2 = n/d$ holds. Substituting these values into the last inequality above, we get:

$$\sigma_2 \geq \sqrt{\frac{n/d - 1}{(n-1)}} = \sqrt{\frac{(n-d)}{d(n-1)}} = \Omega(1/\sqrt{d})$$

when $d \leq 0.99 \cdot n$. □

[3]https://arxiv.org/abs/1304.4132
[4]*https://arxiv.org/abs/1306.3969*
[5]https://www.sciencedirect.com/science/article/pii/0012365X9190112F?via%3Dihub

5.2 Some Properties of Ramanujan Graphs

Let G be a d-regular Ramanujan graph. Then:

Corollary 5.4. *For any $S, T \subseteq V$, if $|S| \cdot |T| \cdot d > 4n^2$, then $E(S,T) \neq \emptyset$.*

That is, if we take two vertex sets which are large enough, then in Ramanujan graphs, there is guaranteed to be an edge between those two sets.

Proof. We prove the contra-positive. Suppose $E(S,T) = \emptyset$. The Expander Mixing Lemma for d-regular graphs gives us:

$$\left| |E(S,T)| - \frac{d \cdot |S| \cdot |T|}{dn} \right| \leq \sigma_2 d \sqrt{|S| \cdot |T|}$$

which implies that

$$\frac{d|S| \cdot |T|}{n} \leq 2\sqrt{d} \cdot \sqrt{|S| \cdot |T|} \iff |S| \cdot |T| \cdot d \leq 4n^2.$$

□

Example 5.5. Say $d = \sqrt{n}$. Consider S and T. If $|S|, |T| \geq 2n^{3/4}$, then $E(S,T) \neq \emptyset$. But in fact, if $|S|, |T| \geq 4n^{3/4}$ which are a bit bigger than what we required before, then there will be considerable amount of edges between then, i.e. $|E(S,T)| = \Omega(d|S||T|/n) = \Omega(n)$.

Corollary 5.6. *Any independent set S has size at most $2n/\sqrt{d}$.*

Proof. As $E(S,S) = \emptyset$ by the definition of an independent set, then $|S|^2 \cdot d \leq 4n^2$. So $|S| \leq 2n/\sqrt{d}$. □

Exercise 5.7. Is there a d-regular graph that is better than Ramanujan graph? Say, for all S, T where $|S|, |T| \geq 2\sqrt{n}$, we have $E(S,T) \neq \emptyset$?

6 Expander: Conductance vs Eigenvalue

Recall Cheeger's inequality:

$$\frac{\lambda_2(\boldsymbol{N}_G)}{2} \leq \Phi(G) \leq \sqrt{2\lambda_2(\boldsymbol{N}_G)}.$$

Roughly, it claims the following fact: $\Phi(G)$ is big $\iff$ $\lambda_2(\boldsymbol{N}_G)$ is big. But in the *very well-connected* regime, the implication may not necessarily go both ways.

6.1 Eigenvalue can be stronger than Conductance

When the conductance $\Phi(G) = 1 - o(1)$ is almost maximum, we cannot always conclude that $|E(S,T)| \approx \frac{d(S)d(T)}{d(V)}$ for all $S, T \subseteq V$. Take the case of the star graph: it has conductance 1. But for S, T such that they are equal sized partitions of the outer vertices, we get $|E(S,T)| = 0$, whereas $\frac{d(S)d(T)}{d(V)}$ is roughly $O(n)$.

In general, by Cheeger's inequality, $\lambda_2(\boldsymbol{N}_G) \geq \Phi(G)^2/2 = \frac{1}{2} - o(1)$. Therefore, we can only bound $\sigma_2 \leq \frac{1}{2} + o(1)$. In this particular case, even when $\mathrm{vol}(S), \mathrm{vol}(T) = \mathrm{vol}(V)/2$, the Expander Mixing Lemma does not guarantee an edge between S and T. That is, $E(S,T) = \emptyset$ possibly.

6.2 Conductance can be stronger than Eigenvalue

Going the other way, there are cases when the eigenvalue can imply strong expansion, but the conductance of that graph might not be as good. To see one such case, suppose G is d-regular and $\sigma_2 \leq O(\frac{1}{\sqrt{d}})$ is almost minimum. Then, we have:

$$\lambda_2(\boldsymbol{N}_G) \geq 1 - \frac{1}{O(\sqrt{d})}.$$

By Cheeger's inequality, we only get:

$$\Phi(G) \geq \lambda_2(\boldsymbol{N}_G)/2 \geq \frac{1}{2} - o(\frac{1}{O(\sqrt{d})}).$$

Therefore, Cheeger's does not allow us to get strong lower bounds on the conductance, in this case. In particular, it does not imply $\Phi(G) \approx (1-\epsilon)$.

A d-regular graph G where $\Phi(G) \approx (1-\epsilon)$ is called a **lossless expander**, and it has lots of applications in:

- Complexity theory (pseudo-randomness)
- Streaming algorithms (deterministic sparse recovery).
- Dynamic algorithms (dynamic matching)[6]

Note that d should be small, and G should be regular. (Otherwise easy: consider clique and star). There is a bipartite version which is important as well.

6.3 Proof of variational characterization of singular values

Theorem 6.1. *Let* $\mathbf{A}'_\mathbf{G} = \mathbf{I} - \mathbf{N}_\mathbf{G}$. *Let* σ_2 *denote the second largest eigenvalue in terms of absolute value. Then,*

$$\sigma_2 = \max_{z,w} \frac{|\boldsymbol{z}^\top(\boldsymbol{A}'_G - \boldsymbol{v}_1\boldsymbol{v}_1^\top)\boldsymbol{w}|}{\|\boldsymbol{z}\| \cdot \|\boldsymbol{w}\|}$$

Proof. For any matrix $\boldsymbol{M}$, we write the singular value decomposition of $\boldsymbol{M}$ as

$$\boldsymbol{M} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^\top = \sum_i \sigma_i(\boldsymbol{M}) \cdot \boldsymbol{u}_i\boldsymbol{v}_i^\top$$

where $\boldsymbol{u}_i, \boldsymbol{v}_i$ are unit vectors and $\sigma_i(\boldsymbol{M})$ are the singular values of $\boldsymbol{M}$. Note that this generalizes eigenvalue decomposition, and holds for any $m \times n$ matrix $\mathbf{M}$. Chapter 3 of this book gives a gentle introduction to singular value decomposition and mentions a lot of applications:

https://home.ttic.edu/~avrim/book.pdf

We will need the following fact:

[6] https://arxiv.org/abs/2108.10461

Fact 6.2. *For each i, we have :*

$$\boldsymbol{M}\boldsymbol{v}_i = \sigma_i(\boldsymbol{M})\boldsymbol{u}_i.$$

Furthermore, the first (largest) singular value $\sigma_1(\boldsymbol{M})$ of $\boldsymbol{M}$ is such that

$$\sigma_1(\boldsymbol{M}) = \|\boldsymbol{M}\|_2 = \max_{\boldsymbol{x}} \frac{\|\boldsymbol{M}\boldsymbol{x}\|}{\|\boldsymbol{x}\|}.$$

Lemma 6.3 (Variational Characterization of the First Singular Values)**.** *We have*

$$\sigma_1(\boldsymbol{M}) = \max_{\boldsymbol{z},\boldsymbol{w}} \frac{\boldsymbol{z}^\top \boldsymbol{M}\boldsymbol{w}}{\|\boldsymbol{z}\| \cdot \|\boldsymbol{w}\|}.$$

Proof. For any unit vector $\boldsymbol{z}$ and $\boldsymbol{w}$, by Cauchy-Schwarz we have

$$\boldsymbol{z}^\top \boldsymbol{M}\boldsymbol{w} \leq \|z\| \cdot \|\boldsymbol{M}\boldsymbol{w}\| \leq \|z\| \cdot \|\boldsymbol{M}\|_2 \cdot \|\boldsymbol{w}\| = \sigma_1(\boldsymbol{M})\|z\| \cdot \|\boldsymbol{w}\|.$$

That is,

$$\sigma_1(\boldsymbol{M}) \geq \max_{\boldsymbol{z},\boldsymbol{w}} \frac{\boldsymbol{z}^\top \boldsymbol{M}\boldsymbol{w}}{\|\boldsymbol{z}\| \cdot \|\boldsymbol{w}\|}.$$

But we know that this equality is attained for the unit vectors $\boldsymbol{u}_1$ and $\boldsymbol{v}_1$ because

$$\boldsymbol{u}_1^\top \boldsymbol{M}\boldsymbol{v}_1 = \boldsymbol{u}_1^\top(\sigma_1(\boldsymbol{M})\boldsymbol{u}_1) = \sigma_1(\boldsymbol{M}).$$

This prove the lemma. □

Claim 6.4. *For symmetric matrix $\boldsymbol{M}$, $\sigma_1(\boldsymbol{M}) = \max\{|\lambda_1(\boldsymbol{M})|, |\lambda_n(\boldsymbol{M})|\}$.*

Proof. We have that

$$\begin{aligned}
(\sigma_1(\boldsymbol{M}))^2 = \|\boldsymbol{M}\|_2^2 &= \max_{\boldsymbol{x}} \frac{\|\boldsymbol{M}\boldsymbol{x}\|^2}{\|\boldsymbol{x}\|^2} \\
&= \max_{\boldsymbol{x}} \frac{\boldsymbol{x}^\top \boldsymbol{M}^\top \boldsymbol{M}\boldsymbol{x}}{\boldsymbol{x}^\top \boldsymbol{x}} \\
&= \lambda_1(\boldsymbol{M}^2).
\end{aligned}$$

But we know that $\lambda_1(\boldsymbol{M}^2) = (\max\{|\lambda_1(\boldsymbol{M})|, |\lambda_n(\boldsymbol{M})|\})^2$. This proves the claim. □

When we let $\boldsymbol{M} = \boldsymbol{A}'_G - \boldsymbol{v}_1\boldsymbol{v}_1^\top$, we have

$$\begin{aligned}
\sigma_1(\boldsymbol{M}) &= \max\{|\lambda_1(\boldsymbol{A}'_G - \boldsymbol{v}_1\boldsymbol{v}_1^\top)|, |\lambda_n(\boldsymbol{A}'_G - \boldsymbol{v}_1\boldsymbol{v}_1^\top)|\} \\
&= \max\{|\lambda_2(\boldsymbol{A}'_G)|, |\lambda_n(\boldsymbol{A}'_G)|\} \\
&= \sigma_2
\end{aligned}$$

by the definition of σ_2 from the lecture. (Basically, we defined $\sigma_2 = \sigma_2(\boldsymbol{A}'_G)$.) We can conclude now that

$$\sigma_2(\boldsymbol{A}'_G) = \sigma_1(\boldsymbol{M}) = \max_{\boldsymbol{z},\boldsymbol{w}} \frac{|\boldsymbol{z}^\top(\boldsymbol{A}'_G - \boldsymbol{v}_1\boldsymbol{v}_1^\top)\boldsymbol{w}|}{\|\boldsymbol{z}\| \cdot \|\boldsymbol{w}\|}$$

by the variational characterization of $\boldsymbol{M}$.

□