CS 270 Algorithms **Spring 2021**

Lecture 15: Applications of Cheeger's Inequality

Lecturer: Prasad Raghavendra Scribe: Ben Hoberman, Catherine Huang, Frank Wang

15.1 Miscellaneous Discussion on Cheeger's Inequality

First, we'll remind ourselves of some definitions and results from prior lectures:

Definition 15.1. *For a d-regular graph, the* normalized Laplacian of G L_G *(or L when G is clear) is given by* $L_G = \mathbb{I} - \frac{1}{d}A_G$.

Definition 15.2. $\phi(G)$, the conductance of G, *measures the connectivity of a graph and is defined by*

$$\begin{aligned}\phi(G) &\triangleq \min_{S\subset V, |S|\leq n/2} \phi(S) \\ &= \min_{S\subset V, |S|\leq n/2} \frac{E(S, V\setminus S)}{\mathrm{vol}(S)} \\ &= \min_{S\subset V, |S|\leq n/2} \frac{E(S, V\setminus S)}{d|S|} \text{ when } G \text{ is } d\text{-regular.}\end{aligned}$$

Theorem 15.3 (Cheeger's Inequality)**.** *For a d-regular graph G,*

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

Calculating $\phi(G)$ exactly is an NP-complete problem. Since eigenvalues can be calculated efficiently, Cheeger's Inequality allows us to *bound* $\phi(G)$ efficiently using the spectral properties of the graph's normalized Laplacian.

Since Cheeger's Inequality gives us approximate bounds, it's natural to ask how tight these bounds are. We can see, however, that Cheeger's Inequality can indeed be tight: consider the cycle on n vertices C_n. Doing a bit of algebra, one can see that $\sqrt{2\lambda_2(L_{C_n})} = \phi(C_n)$.

Another natural question we might ask is why just the second eigenvalue? Can we find similar types of bounds on λ_k for $k > 2$? As it turns out, we can, though these "higher" versions of Cheeger's Inequality are less clear and were proved much later than Cheeger's Inequality itself. Specifically, λ_2 measures how connected a graph is in terms of it being split into 2 components, so for higher k, we'd be interested in how connected the graph is when separated into k components.

Theorem 15.4 (Higher Cheeger's Inequality)**.** *In a graph G, there exists a partition $S_1, S_2, \ldots, S_n$ of V such that*

$$\max_{i\in[k]} \phi(S_i) \leq \sqrt{\lambda_k * poly(k)}$$

Cheeger's Inequality isn't our only tool for efficiently estimating a graph's conductance, though. Recall that in previous lectures, we provied how to use Bourgain's embedding and linear programming to efficiently find a set $S \subset V$ s.t. $\phi(S) \leq \phi(G) \log n$ (where $n = |V|$ as usual). Our constructive proof of Cheeger's Inequality, on the other hand, finds an embedding with expansion $\phi(S) \leq \sqrt{\phi(G)}$. These are fundamentally qualitatively different types of bounds, which is worth noting and thinking about.

Definition 15.5. *A* combinatorial expander *is a d-regular graph G such that $\phi(G) \geq \Omega(1)$.*

Definition 15.6. *A* spectral expander *is a d-regular graph G such that $\lambda_2\left(\frac{L_G}{d}\right) \geq \Omega(1)$.*

Intuitively, both combinatorial expanders and spectral expanders are sparse graphs which have high connectivity. A graph being a combinatorial expander is more obviously useful to us (in proofs, applications, etc.), but determining directly whether a graph is a combinatorial expander is difficult, as discussed earlier. Cheeger's Inequality, however, tells us that mathematically speaking, a graph G is a combinatorial expander if and only if it is a spectral expander. And since we can easily determine whether a graph is a spectral expander, we can essentially hack our way into figuring out this hard but useful property of general graphs.

15.2 Second Eigenvalue of a Planar Graph

In this section, we'll upper bound λ_2 of a planar graph. Recall that a planar graph is a graph which can be embedded into $\mathbb{R}^2$ (drawn on the plane) without any of the edges crossing one another. This will be a fun proof that will require some unfamiliar tools from geometry.

Lemma 15.7 (Circle-Packing Theorem)**.** *For every connected planar graph G, there exists a circle packing that is isomorphic to the graph G. Concretely, this means that there exist circles $C_1, C_2, \ldots, C_n$ with disjoint interiors such that $(i, j) \in E$ if and only if C_i touches C_j. See figure 15.1 for an illustration.*

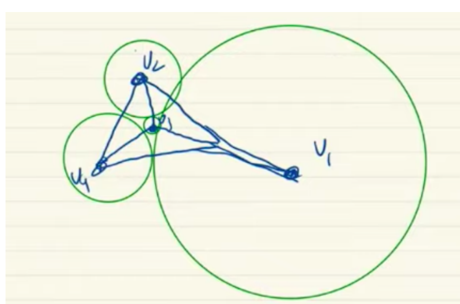

Figure 15.1: Circle-packing representation of a simple planar graph

Definition 15.8. *A* stereographic projection *projects $\mathbb{R}^2$ on to a sphere by mapping each point $v \in \mathbb{R}^2$ to a point v' on the sphere using the line connecting v and x, the sphere's north pole. This is illustrated in figure 15.2.*

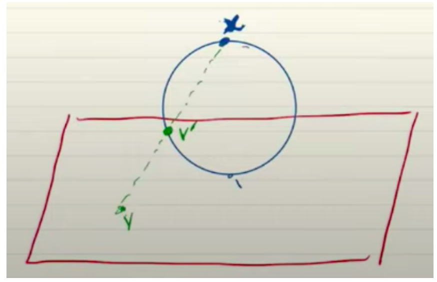

Figure 15.2: An example of stereographic projection: the sphere is tangent to the plane, and the point x is chosen to be opposite the point of tangency. This figure shows the point $v \in \mathbb{R}^2$ being projected to v' on the sphere.

One useful property of stereographic projections is that they map circles to circles, so projecting our circle-packing representation of a planar graph $G = (V, E)$ still maintains the property that $(i, j) \in E$ iff C_i touches C_j.

Theorem 15.9. *For any planar d-regular graph G, $\lambda_2(L_G) \leq 8/n$.*

Proof. To prove this, we use the embedding of G in $\mathbb{R}^3$ obtained by assigning vectors $v_1, v_2, \ldots, v_n \in \mathbb{R}^3$ to the points in $\mathbb{R}^2$ given by taking the stereographic projection of the circle-packing embedding of G into $\mathbb{R}^3$. We'll use the Rayleigh quotient under some constraints to show our result. In particular, we'll show that

$$\frac{\sum_{(i,j)\in E} \|v_i - v_j\|^2}{d \sum_{i \in V} \|v_i\|^2} \leq \frac{8}{n}. \tag{15.1}$$

subject to the constraint that $\sum_{i=1}^{n} v_i = 0$, which enforces that our embedding is orthogonal to the all ones vector (which is the eigenvector of λ_1). Note that Rayleigh quotient argument works because $\min\limits_{x \perp \vec{1}} \frac{x^\top L_G x}{x^\top x} = \lambda_2(L_G)$.

From the circle-packing theorem, we can represent the planar graph as disjoint, tangent circles. Then, we can do a stereographic project of the entire plane onto a unit sphere and assign each of the centers of C_i to be v_i, with each v_i lying on the surface of the sphere. Circles on the sphere form "caps" on the surface of the sphere. R_i is defined as the ℓ_2 distance from the center of the cap formed by C_i to the edge of C_i. This is illustrated in figure 15.3.

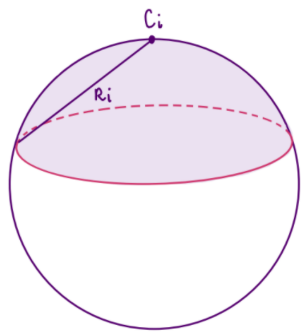

Figure 15.3: Spherical cap

We now bound the Rayleigh quotient by consider its numerator and denominator separately.

Since each vertex gets mapped to a point on the unit sphere centered at the origin (we can stipulate this just fine), $\|v_i\| = 1$ for all v_i, so the denominator of the LHS of 15.1 is $d\sum_{i=1}^{n} \|v_i\|^2 = nd$.

To bound the numerator, by the triangle inequality, $\|v_i - v_j\|_2 \leq R_i + R_j$. Then, squaring both sides, $\|v_i - v_j\|_2^2 \leq (R_i + R_j)^2 \leq 2(R_i^2 + R_j^2)$. Summing over all edges in the graph and using d-regularity, $\sum_{(i,j)\in E} \|v_i - v_j\|_2^2 \leq 2d\left(\sum_{i=1}^{n} R_i^2\right)$.

The surface area of the unit sphere is 4π, and since the area of each cap is πR_i^2, we can bound the sum of the caps by $\sum_{i=1}^{n} \pi R_i^2 \leq 4\pi$ to yield $\sum_{i=1}^{n} R_i^2 \leq 4$. Plugging this in, we can see that $\sum_{(i,j)\in E} \|v_i - v_j\|_2^2 \leq 8d$.

Combining our bounds for the numerator and denominator,

$$\lambda_2 = \frac{\sum_{(i,j)\in E} \|v_i - v_j\|_2^2}{d\sum_{i=1}^{n} |v_i|^2} = \frac{\sum_{(i,j)\in E} \|v_i - v_j\|_2^2}{dn} \leq \frac{8d}{nd} = \frac{8}{n}$$

This shows our desired bound, but we have not yet guaranteed that $\sum_{i=1}^{n} v_i = 0$. There are many ways in which the sphere could have been chosen, and only one configuration satisfying the condition $\sum_{i=1}^{n} v_i = 0$ is needed. A proof can be done using Brouwer's Fixed Point Theorem and ideas from topology, but these ideas are "too fun for this course."

$\square$

15.3 Random Walks & Mixing Via Cheeger's Inequality

Definition 15.10. *A* random walk *on G is a stochastic process over state space V: we pick an arbitrary starting state $u \in V$. For each time step, if we are currently at node v, we take a step by picking one of v's neighbors uniformly at random and moving to it.*

Let $G = (V, E)$ be a d-regular graph (with labels $V = [n]$ as usual).

To analyze random walks, let $P_t(v)$ be the probability that our random walk is at vertex v after t time steps with initial starting vertex u. We can naturally define $P_t \in \mathbb{R}_+^n$ by $P_t[v] = P_t(v)$. P_t is a probability distribution and thus satisfies $\sum_{v=1}^n P_t(v) = 1$. Note that we have a nice closed form for $P_0 = \begin{cases} 1, & \text{if } v = u \\ 0, & \text{else} \end{cases}$.

We can now define the transition function nicely. Because we're working with a d-regular graph, all connected vertices can be reached with probability $1/d$. Thus,

$$\begin{aligned}
\Pr(\text{at } v \text{ at time } t) &= \sum_{(w,v)\in E} \frac{1}{d} \Pr(\text{at } w \text{ at time } t-1) \\
&= \frac{1}{d} A \cdot P_{t-1} \\
&= \tilde{A} \cdot P_{t-1} \\
&= \tilde{A}^t P_0, \text{ taking } \tilde{A} \triangleq \frac{1}{d} A.
\end{aligned}$$

P_t eventually reaches steady-state corresponding to the eigenvector-eigvenvalue pair (of $\tilde{A}$) $\lambda_n = 1, v_n = \vec{1}/n$.

Definition 15.11. ϵ-mixing time *is the number of time steps after which* $\left\| P_t - \frac{1}{n}\vec{1} \right\|_2 \le \epsilon$ *for* $\epsilon > 0$.

This quantity measures the deviation of the probability distribution from the uniform distribution.

Doing some algebra, we can see

$$\begin{aligned}
P_t &= \tilde{\text{A}}\, P_{t-1} \\
\implies P_t - \frac{1}{n}\vec{1} &= \tilde{\text{A}}\, P_{t-1} - \frac{1}{n}\vec{1} = \tilde{\text{A}} \left(P_{t-1} - \frac{1}{n}\vec{1} \right) \text{ since } \vec{1} \text{ is an eigenvector of } \tilde{\text{A}} \\
&= \sum_{i=1}^{n} \lambda_i(\tilde{\text{A}})^t c_i \vec{e}_i \text{ where } c_i \text{ are coordinates in the eigenbasis } \{\vec{e}_i : i \in [n]\} \\
\implies \left\| P_t - \frac{1}{n}\vec{1} \right\|_2 &\le \max_{i\in[n-1]} \left[|\lambda_i|^t \right]
\end{aligned}$$

Setting the expression equal to ϵ, we can see that the ϵ-mixing time for G is $\frac{\log(1/\epsilon)}{\log\left(\max_{i\in[n-1]} \lambda_i(\tilde{\text{A}})\right)}$.

Note, however, that since $L = dI - A$, $\lambda_i(L) = d - \lambda_i(A) = d - d\lambda_i(\tilde{\text{A}})$. We can also see that if the graph isn't connected, we'll end up with other steady states (other eigenvalues equal to one) and the steady state mixing time will go to infinity, which is exactly what we'd expect.

Now, applying Cheeger's inequality to our arbitrary connected d-regular graph G, $\lambda_2(\tilde{L}) \ge \frac{1}{2}\phi(G)^2$. But $\phi(G) \ge \frac{1}{dn}$ since we need to cut at least one edge (recall the definition of G's conductance) for any split, so

$\lambda_2(L) \geq O\left(\frac{1}{(nd)^2}\right)$. Since λ_2 lower-bounds all of the non-steady state eigenvalues, we can upper bound the ϵ-mixing time by a polynomial in n (and d) (roughly $O(n^3)$).

This generates a very simple randomized algorithm to test a graph G for $s-t$ connectivity: set $u = s$ and run a random walk for n^5 steps. If we see t, then we report that G is $s-t$ connected, otherwise, that it's not.

This is obviously not efficient in terms of time, but most deterministic algorithms use $\Theta(n)$ space, and this algorithm only has to keep track of a current position ($\log n$ space) and a step counter ($\log n^5 = O(\log n)$ space), meaning this simple algorithm uses only $O(\log n)$ space. There does exist a deterministic algorithm to test G for $s-t$ connectivity in $O(\log n)$ space, but it was only discovered as recently as 2005.