ORIE 6334 Spectral Graph Theory	November 22, 2016
Lecture 25	
Lecturer: David P. Williamson	*Scribe: Pu Yang*

In the remaining three lectures, we will cover a prominent result by Arora, Rao, and Vazirani for the sparsest cut problem. In this lecture, we will set the scene by giving a prior result by Leighton and Rao, and explain what the Arora-Rao-Vazirani algorithm has to do with the topic of this course.

Recall that the sparsity of a cut $S \subseteq V$ in a graph $G = (V, E)$ is defined as

$$\rho(S) \equiv \frac{|\delta(S)|}{|S||V-S|},$$

where $\delta(S)$ is the set of edges with exactly one end in S. The sparsest cut of a graph is the defined as

$$\rho(G) \equiv \min_{S \subseteq G} \rho(S).$$

Recall also that the sparsest cut is related to two other concepts. The *edge expansion* of a set $S \subseteq V$, $|S| \leq n/2$ is defined as

$$\alpha(S) \equiv \frac{|\delta(S)|}{|S|},$$

and the edge expansion of a graph

$$\alpha(G) \equiv \min_{S \subseteq V, |S| \leq n/2} \alpha(S).$$

The *conductance* of a set $S \subseteq V$ is defined as

$$\phi(S) \equiv \frac{|\delta(S)|}{\min(\mathrm{Vol}(S), \mathrm{Vol}(V-S))},$$

where

$$\mathrm{Vol}(S) \equiv \sum_{i \in S} \deg(i),$$

and

$$\phi(G) \equiv \min_{S \subseteq G} \phi(S).$$

We'll cover the following two results in the next three lectures.

[0]Part of this lecture is taken from Leighton and Rao 1999, and another from an unpublished note of Sudan and W; an analogous result to that of Sudan and W also appears in the Arora, Rao, and Vazirani paper.

Theorem 1 (Leighton and Rao 1988) *There is an $O(\log n)$-approximation algorithm for the sparsest cut problem.*

Theorem 2 (Arora, Rao, and Vazirani 2004) *There is an $O(\sqrt{\log n})$-approximation algorithm for the sparsest cut problem.*

Let us first look at a relaxation of the sparsest cut problem:

$$\begin{aligned} \rho_{LR} \equiv \min \sum_{e \in E} x(e) \\ s.t. \sum_{i,j:i \neq j} d_x(i,j) \geq 1 \\ x(e) \geq 0 \qquad \forall e \in E, \end{aligned}$$

where $d_x(i,j)$ is the shortest path distance from i to j using x as edge lengths.

We claim that this relaxation can be solved in polynomial time. To see that it is a relaxation, let S^* be the sparsest cut. Set

$$x(e) = \begin{cases} \frac{1}{|S^*||V-S^*|} & \text{if } e \in S^* \\ 0 & \text{o.w.} \end{cases}$$

Then

$$\sum_{i,j:i \neq j} d_x(i,j) \geq |S^*||V-S^*| \frac{1}{|S^*||V-S^*|} = 1.$$

This is true because for any path that connects a node in S^* and a node in $V - S^*$, it must use at least one edge in $\delta(S^*)$, and the number of pairs of nodes with one in S^* and the other in $V - S^*$ is $|S^*||V - S^*|$. The objective function for x set in this way is

$$\sum_{e \in E} x(e) = \frac{|\delta(S^*)|}{|S^*||V-S^*|} = \rho(S^*) = \rho(G).$$

So x set in this way is a feasible solution to this problem, and the optimal solution for this problem gives a lower bound on $\rho(G)$.

In Lecture 14 on low-stretch trees, we proved the following lemma:

Lemma 3 *Let D be some parameter. There is a partition of G into clusters s.t.*

- *each cluster has diameter (w.r.t. number of edges) less than or equal to D.*
- *there are at most $\alpha|E|$ intercluster edges with $\alpha \leq \frac{4 \ln n}{D}$.*

We claim that by a similar proof, one can get the following extension of the lemma.

Lemma 4 *Let D be some parameter, and $x(e) \geq 0$ be lengths of edges in E. There is a partition of G into clusters s.t.*

- *each cluster has diameter (w.r.t. lengths x) less than or equal to D.*
- *there are at most $\alpha \sum_{e \in E} x(e)$ intercluster edges with $\alpha \leq \frac{4 \ln n}{D}$.*

With this lemma and a little extra work we will arrive at the algorithm of Leighton and Rao. Let's see why.

First, by choosing appropriate D in Lemma 2, we have the following result.

Lemma 5 *Let $D = 1/(4n^2)$. Divide the graph G into clusters that satisfies Lemma 4. Then either there is a cluster C where $|C| \geq \frac{2n}{3}$, or there exists $S \subseteq V$ such that $\rho(S) = O(\log n) \sum_{e \in E} x(e) \leq O(\log n)\rho(G)$, where $x(e)$ is the solution for the relaxed problem.*

Proof: If there is no cluster with size at least $\frac{2}{3}n$, then order the clusters by nonincreasing size, and add clusters in this order to S until $|S| \geq \frac{1}{3}n$. Then it must be the case that $|V - S| \geq \frac{1}{3}n$ as well, so that

$$\rho(S) = \frac{|\delta(s)|}{|S||V-S|} \leq \frac{4n^2 \cdot 4 \ln n \sum_{e \in E} x(e)}{\frac{1}{9}n^2} = O(\log n) \sum_{e \in E} x(e).$$

The inequality follows $|\delta(S)|$ is at most the number of intercluster edges, which is at most $(4 \ln n / D) \sum_{e \in E} x(e)$ by Lemma 3. □

Therefore, when there is no large cluster with size at least $\frac{2}{3}n$, we have found such a cut. What if there exists a cluster C such that $|C| \geq \frac{2}{3}n$?

Lemma 6 *If there exists a cluster C such that $|C| \geq \frac{2}{3}n$, with diameter no more than $\frac{1}{4n^2}$, we can find S such that $\rho(S) \leq 6 \sum_{e \in E} x(e) \leq 6\rho(G)$.*

Proof: Let $d(i, C) = \min_{j \in C} d(i, j)$. Order the vertices as $i_1, \ldots, i_n$ in non-increasing order of $d(i, C)$. Let $S_k = \{i_1, \ldots, i_k\}$ for $k = 1, \ldots, n-1$.

We claim that $\sum_{i,j \in V} |d(i,C) - d(j,C)| \geq \frac{1}{6}$. We will prove this claim later. The intuition behind this claim is that given we have many nodes in C which are close to each other because the size of C is large and the diameter of C is small, in order for the constraint in the relaxation problem to hold the total of the difference between the distance to C of each pair of nodes must be large.

With this claim, we have that

$$\begin{aligned}
\min_{1 \leq k \leq n-1} \rho(S_k) &= \min_{1 \leq k \leq n-1} \frac{|\delta(S_k)|}{|S_k||V - S_k|} \\
&\leq \frac{\sum_{k=1}^{n-1} |d(i_{k+1}, C) - d(i_k, C)||\delta(S_k)|}{\sum_{k=1}^{n-1} |d(i_{k+1}, C) - d(i_k, C)||S_k||V - S_k|} \\
&= \frac{\sum_{(i,j) \in E} |d(j,C) - d(i,C)|}{\sum_{i,j \in V} |d(j,C) - d(i,C)|} \\
&\leq \frac{\sum_{e \in E} x(e)}{1/6} = 6 \sum_{e \in E} x(e) \leq 6\rho(G).
\end{aligned}$$

The first inequality is easy to check by noticing that the weight $|d(i_{k+1}, C) - d(i_k, C)|$ for $|\delta(S_k)|$ and $|S_k||V - S_k|$ depends only on k. The last inequality is from the claim. Why is the second equality true? We notice that a given edge (i, j) appears in $\delta(S_k)$ for the indices k such that exactly one of i and j is in S_k. Thus if we rewrite

$$\sum_{k=1}^{n-1} |d(i_{k+1}, C) - d(i_k, C)||\delta(S_k)|$$

as a sum over the edges, we get that

$$\sum_{(i,j)\in E} |d(i, C) - d(j, C)|.$$

Similarly, a given pair of vertices i and j appears in the product $|S_k||V - S_k|$ for exactly the indices k such that exactly one of i and j is in S_k, so we can rewrite

$$\sum_{k=1}^{n-1} |d(i_{k+1}, C) - d(i_k, C)||S_k||V - S_k|$$

as

$$\sum_{i,j\in V} |d(j, C) - d(i, C)|.$$

It remains to prove the claim. Pick some $i' \in C$. For any $i \in V$, there is some $j \in C$ such that $d(i, C) = d(i, j)$. So $d(i, i') \le d(i, j) + d(j, i') \le d(i, C) + \frac{1}{4n^2}$. Then

$$\begin{aligned} 1 \le \sum_{i,j:i\ne j} d(i, j) \le \sum_{i,j:i\ne j} (d(i, i') + d(i', j)) = 2n \sum_{i\in V} d(i, i') \\ \le 2n \sum_{i\in V} \left(d(i, C) + \frac{1}{4n^2} \right) = 2n \sum_{i\in v} d(i, C) + \frac{1}{2}. \end{aligned}$$

Therefore

$$\sum_{i\in V} d(i, C) = \sum_{i\notin C} d(i, C) \ge \frac{1}{4n}.$$

Then

$$\sum_{i,j\in V} |d(i, C) - d(j, C)| \ge \sum_{i\notin C, j\in C} d(i, C) = |C| \sum_{i\notin C} d(i, C) \ge \frac{2n}{3} \cdot \frac{1}{4n} = \frac{1}{6}.$$

□ Thus we have proved the existence of an $O(\log n)$-approximate algorithm for sparsest cut.

To improve on this result we look at a related relaxation of the sparsest cut problem proposed independently by Goemans and Linial:

$$\rho_N \equiv \min \sum_{e \in E} x(e)$$
$$s.t. \sum_{i,j: i \neq j} d_x(i,j) \geq 1$$
$$d_x \text{ is a negative type metric,}$$

where d is a negative type metric iff $\sum_{i,j \in V} d(i,j)z(i)z(j) \leq 0$, for all z such that $z^T e = 0$. The following is a useful theorem of determining whether a metric has negative type, and will be used next time to show that we can solve the relaxation in polynomial time.

Theorem 7 *d has negative type iff $\exists f : V \to \Re^n$ s.t. $d(i,j) = \|f(i) - f(j)\|^2$.*

To see why this is still a relaxation, again let

$$x(e) = \begin{cases} \frac{1}{|S^*||V-S^*|} & \text{if } e \in S^* \\ 0 & \text{o.w.} \end{cases}$$

Then for any z $s.t.$ $z^T e = 0$,

$$\begin{aligned} \sum_{i,j \in V} d_x(i,j)z(i)z(j) &= \frac{1}{|S^*||V-S^*|} \sum_{i \in S^*, j \notin S^*} z(i)z(j) \\ &= \frac{1}{|S^*||V-S^*|} \left(\sum_{i \in S^*} z(i) \right) \left(\sum_{j \notin S^*} z(j) \right) \\ &= \frac{1}{|S^*||V-S^*|} \left(\sum_{i \in S^*} z(i) \right) \left(- \sum_{j \in S^*} z(j) \right) \leq 0. \end{aligned}$$

Thus d_x for x so defined is of negative type.[1]

We now show that this relaxation is very related to previous topics of the course. To do this, we introduce the concept of a flow packing. Let P_{ij} be the set of all i-j paths in G. Call $F = (f_{ij})$, symmetric a *flow packing* into G if for all paths $P^k \in P_{ij}$, there exist $f_{ij}^k \geq 0$ such that

$$f_{ij} = f_{ji} = \sum_k f_{ij}^k$$
$$\sum_{i,j,k: (a,b) \in p^k \in P_{ij}} f_{ij}^k \leq 1 \qquad \forall (a,b) \in E.$$

Then the following can be shown.

[1] We also need to show that for the optimal cut the graphs induced by S^* and $V - S^*$ are both connected; it is possible to do so.

Theorem 8 *For flow packing $F = (f_{ij})$, let L_F be weighted Laplacian with weights $w(i,j) = f_{ij}$. Then*

$$\rho_N = \frac{1}{n} \max_{\textit{Flow packings } F} \lambda_2(L_F).$$

The proof of this theorem uses duality, and the details are omitted here.

A natural example of a flow packing is the adjacency matrix A. Then from the theorem, $\rho_N \geq \frac{1}{n}\lambda_2(L_G)$.

What is the difference between the two relaxations? It turns out there are cases for which the negative-type metric relaxation is much stronger than the relaxation without the negative-type restriction. We have the following lemma.

Lemma 9 *There exist graphs G such that $\rho(G) = \Omega(\log n)\rho_{LR}$ but for which $\rho(G) = O(1)\rho_N$.*

We did not have time to prove this lemma in class, but here is the proof.

Proof: There exist 3-regular expander graphs $G = (V, E)$ such that $\alpha(G) \geq \beta$ for some constant $\beta > 0$. For these,

$$\rho(G) = \min_{S \subseteq V} \frac{|\delta(S)|}{|S||V - S|} \geq \min_{S \subseteq V} \frac{\beta}{\max(|S|, |V - S|)} = \frac{\beta}{n - 1}.$$

If we set $x(e) = 4/(n^2(\log n - 2))$ for all $e \in E$, then since G is 3-regular, there are at most $1 + 3 + 3 \cdot 2 + \cdots 3 \cdot 2^{d-1}$ nodes within d hops, or

$$1 + \sum_{k=1}^{d} 3 \cdot 2^{k-1} = 1 + 3(2^d - 1).$$

So within $d = \log_2 n - 3$, there are at most $3 \cdot 2^{\log_2 n - 3} - 2 = \frac{3}{8}n - 2 \leq \frac{1}{2}n$ within this distance. Therefore,

$$\begin{aligned}
\sum_{i,j \in V} d_x(i,j) &= \frac{1}{2} \sum_{i \in V} \sum_{j \neq i} d_x(i,j) \\
&\geq \frac{1}{2} \sum_{i \in V} \frac{n}{2} \cdot \left(\frac{4}{n^2(\log n - 2)} \right) (\log n - 2) \\
&= 1,
\end{aligned}$$

because we need to traverse at least $\log n - 2$ edges to reach $n/2$ vertices from any given vertex i. Thus x is feasible for the linear program.

Then since there are $3n/2$ edges in a 3-regular graph, we have that

$$\rho_{LR} \leq \frac{3n}{2} \cdot \frac{4}{n^2(\log n - 2)} = \frac{6}{n(\log n - 2)},$$

so that $\rho(G) = \Omega(\log n)\rho_{LR}$.

However, recall that

$$\beta \leq \alpha(G) \leq 3\phi(G)$$

for a 3-regular graph, and that by Cheeger's inequality,

$$\phi(G) \leq \sqrt{2\lambda_2(\mathcal{L})} = \sqrt{\frac{2}{3}\lambda_2(L_G)},$$

so it follows that $\lambda_2(L_G) = \Omega(\beta^2)$. In a 3-regular graph $|\delta(S)| \leq 3\min(|S|, |V - S|)$, so that

$$\rho(G) \leq \min_{S \subseteq V} \frac{|\delta(S)|}{|S||V - S|} \leq \min_{S \subseteq V} \frac{3}{\max(|S|, |V - S|)} \leq \frac{6}{n}.$$

Since the adjacency matrix is a flow packing, we have that $\rho_N \geq \lambda_2(L_G)/n$. Thus

$$\rho(G) \leq \frac{6}{n} \leq O(1) \cdot \frac{\lambda_2(L_G)}{n} \leq O(1)\, \rho_N.$$

□