ORIE 6334 Spectral Graph Theory	December 1, 2016
Lecture 27	
Lecturer: David P. Williamson	*Scribe: Qinru Shi*

1 Recap of Previous Lecture

Last time we started to prove the following theorem.

Theorem 1 (Arora, Rao, Vazirani, 2004) *There is an $O(\sqrt{\log n})$-approximation algorithm for sparsest cut.*

The proof of the theorem uses a SDP relaxation in terms of vectors $v_i \in \mathbb{R}^n$ for all $i \in V$. Define distances to be $d(i,j) \equiv \|v_i - v_j\|^2$ and balls to be $B(i,r) \equiv \{j \in V \mid d(i,j) \leq r\}$. We first showed that if there exists a vertex $i \in V$ such that $|B(i,1/4)| \geq n/4$, then we can find a cut of sparsity $\leq O(1) \cdot OPT$. If there does not exist such a vertex in V, then we can find $U \subseteq V$ with $|U| \geq n/2$ such that for any $i \in U$, $1/4 \leq \|v_i\|^2 \leq 4$ and there are at least $n/4$ vertices $j \in U$ such that $d(i,j) > 1/4$.

Then we gave the ARV algorithm.

Algorithm 1: ARV Algorithm

Pick a random vector r such that $r(i) \sim N(0,1)$
Let $L = \{i \in V : v_i \cdot r \leq -1\}$ and $R = \{i \in V : v_i \cdot r \geq 1\}$
Find a maximal matching $M \subseteq \{(i,j) \in L \times R : d(i,j) \leq \Delta\}$
Let L', R' be the vertices in L, R respectively that remain uncovered
Sort $i \in V$ by increasing distance to L' (i.e. $d(i, L')$) to get $i_1, i_2, \ldots, i_n$
Let $S_k = \{i_1, \ldots, i_k\}$ and return $S = \arg\min_{1 \leq k \leq n-1} \rho(S_k)$

Observation 1 *At the end of the ARV algorithm, for any $i \in L'$ and $j \in R'$, $d(i,j) > \Delta$.*

Assume the matching algorithm gives the same matching for r as for $-r$. Then, we can assume that the probability of i being matched if $i \in L$ is the same as the probability of i being matched if $i \in R$.

Next, we stated the following two theorems and proved the first one.

Theorem 2 *There exists some constant c' such that* $\Pr[|L|, |R| \geq c'n] \geq c'$.

Theorem 3 (Structure Theorem) *For $\Delta = \Omega(1/\sqrt{\log n})$, $\mathbb{E}[|M|] \leq (\frac{c'}{2})^2 n$.*

The two theorems imply that with constant probability, $|L'|, |R'| \geq \frac{c'}{2}n$, and $d(i,j) \geq \Delta$ for all $i \in L'$ and $j \in R'$. We showed that if this is the case, we can then conclude that the algorithm gives us an $O(\sqrt{\log n})$-approximation. Today we turn to the proof of the Structure Theorem.

[0]This lecture is derived from lecture notes of Boaz Barak and David Steurer `http://sumofsquares.org/public/lec-arv.html`.

2 Proof of Structure Theorem

The proof shown in this section is due to Boaz Barak and David Steurer (2016). The original ARV algorithm gives an $O((\log n)^{2/3})$-approximation algorithm and needs another algorithm to reach the guarantee of $O(\sqrt{\log n})$. Later, Lee showed that the original ARV algorithm also gives $O(\sqrt{\log n})$-approximation. Both of these analyses are long and technical. In 2016, Rothvoss gave a somewhat easier proof (`https://arxiv.org/abs/1607.00854`). Very recently Barak and Steurer gave a much easier proof, and this is what we will show today.

Recall the proof ideas we talked about last lecture. We know that $\frac{v}{\|v\|} \cdot r \sim N(0,1)$; from this it is possible to prove a concentration result showing that

$$\Pr[v \cdot r \geq \alpha] \leq \exp\left(-\frac{\alpha^2}{\|v\|^2}\right).$$

Thus

$$\Pr[(v_i - v_j) \cdot r \geq C\sqrt{\ln n}] \leq e^{-\frac{C^2 \ln n}{8}} = \frac{1}{n^{C^2/8}}$$

for any $i, j \in U$, since $\|v_i - v_j\|^2 \leq 8$. Hence, for sufficiently large C, we have

$$(v_i - v_j) \cdot r \leq C\sqrt{\ln n}$$

for all $i, j \in U$ with high probability. Then one can show that

$$\mathbb{E}[\max_{i,j \in U}(v_i - v_j) \cdot r] \leq C\sqrt{\ln n}.$$

For simplicity of notation, we rename $v_i \cdot r$ as X_i. Then,

$$\mathbb{E}[\max_{i,j \in U}(X_i - X_j)] \leq C\sqrt{\ln n}.$$

Next, we will prove the following lemma.

Lemma 4 (Projection Lemma)

$$\frac{\Omega(1)}{\Delta}\left(\frac{\mathbb{E}[|M|]}{n}\right)^3 \leq \mathbb{E}\left[\max_{i,j \in U}(X_i - X_j)\right] \leq C\sqrt{\ln n}.$$

For the rest of the lecture, we will restrict our attention to vertices in U and ignore anything outside of U; we let $n = |U|$, and since $|U| \geq n/2$, this only changes the constants in what we need to prove.

From the Projection Lemma, for the right choice of constants and $\Delta = \Omega(1/\sqrt{\log n})$, we get that

$$\left(\frac{\mathbb{E}[|M|]}{n}\right)^3 \leq \left(\left(\frac{c'}{2}\right)^2\right)^3,$$

which then proves the Structure Theorem. So the Projection Lemma implies the Structure Theorem, and we now turn to proving the Projection Lemma.

Consider a graph $H = (U, E')$ where $E' = \{(i,j) \in U \times U : d(i,j) \le \Delta\}$. Let

$$H(i,k) = \{j \in U : j \text{ can be reached from } i \text{ in at most } k \text{ steps in } H\}.$$

Define

$$Y(i,k) = \max_{j \in H(i,k)} (X_j - X_i)$$

$$\Phi(k) = \sum_{i=1}^{n} \mathbb{E}[Y(i,k)]$$

where i ranges over all starting points. Then,

$$\frac{1}{n}\Phi(k) \le \mathbb{E}\left[\max_{i,j \in U}(X_i - X_j)\right].$$

The idea is that the expectation in the Projection Lemma of $\mathbb{E}\left[\max_{i,j\in U}(X_i - X_j)\right]$ ranges over all $i, j \in U$; what we'll do here is see how large the difference of $X_j - X_i$ can be if they are at most k edges apart in the graph H, and we'll look at what happens as we increment k. That incrementation happens in the following lemma.

Lemma 5 (Chaining Lemma)

$$\Phi(k+1) \ge \Phi(k) + 4\mathbb{E}[|M|] - O(n) \max_{i,j \in H(i,k+1)} \left(\mathbb{E}[(X_i - X_j)^2]\right)^{\frac{1}{2}}$$

We first show that the Chaining Lemma implies the Projection Lemma. For any vector x,

$$\mathbb{E}[(x \cdot r)^2] = \|x\|^2 \mathbb{E}\left[\left(\frac{x}{\|x\|} \cdot r\right)^2\right] = \|x\|^2.$$

since $\frac{x}{\|x\|} \cdot r \sim N(0,1)$. Therefore,

$$\mathbb{E}[(X_i - X_j)^2] = \|v_i - v_j\|^2 \le k\Delta$$

by triangle inequality and the fact that each edge (p,q) in H has $\|v_p - v_q\|^2 \le \Delta$. The Chaining Lemma implies that there exists a constant $\tilde{c}$ such that for any k,

$$\Phi(k+1) \ge \Phi(k) + 4\mathbb{E}[|M|] - \tilde{c}n\sqrt{k\Delta}.$$

Let

$$k_0 = \left(\frac{9}{\tilde{c}^2}\right)\left(\frac{\mathbb{E}[|M|]}{n}\right)^2 \cdot \frac{1}{\Delta}.$$

Then, for any $k \le k_0$, we have

$$\Phi(k+1) \ge \Phi(k) + \mathbb{E}[|M|],$$

which implies that

$$\Phi(k+1) \ge (k+1)\mathbb{E}[|M|].$$

Thus,

$$\begin{aligned}\mathbb{E}[\max_{i,j}(X_i - X_j)^2] &\geq \frac{1}{n}\Phi(k_0) \geq \frac{k_0}{n}\mathbb{E}[|M|] \\ &= \left(\frac{9}{\tilde{c}}\right)^2 \left(\frac{\mathbb{E}[|M|]}{n}\right)^3 \cdot \frac{1}{\Delta}.\end{aligned}$$

Hence, we have shown that the Chaining Lemma implies the Projection Lemma.

Now the remaining step is to prove the Chaining Lemma. We will need the following two probability results for this proof.

Lemma 6 *For any two random variables X and Y,*

$$|\mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]| \leq \sqrt{\mathrm{Var}[X]\,\mathrm{Var}[Y]}$$

Theorem 7 (Borell's Theorem) *If $Z_1, Z_2, \ldots, Z_t$ have mean 0 and are jointly normally distributed, then there exists a constant $\hat{c}$ such that*

$$\mathrm{Var}[\max(Z_1, \ldots, Z_t)] \leq \hat{c}\max(\mathrm{Var}[Z_1], \ldots, \mathrm{Var}[Z_t]).$$

Note that in Borell's Theorem, there's no dependence on the number of variables t. The reason why Borell's Theorem is useful is that for fixed i, $(X_i - X_j)$ for some $j \in H(i,k)$ has mean 0 and are (jointly) normally distributed, so $\mathrm{Var}[X_j - X_i] = \mathbb{E}[(X_j - X_i)^2]$. Borell's Theorem then says that

$$\begin{aligned}\mathrm{Var}[Y(i,k)] = \mathrm{Var}\left[\max_{j \in H(i,k)}(X_j - X_i)\right] &\leq \hat{c}\max_{j \in H(i,k)}\mathrm{Var}[X_j - X_i] \\ &= \hat{c}\max_{j \in H(i,k)}\mathbb{E}[(X_j - X_i)^2] \\ &= \hat{c}\max_{j \in H(i,k)}\|v_j - v_i\|^2 \\ &\leq \hat{c} \cdot k\Delta\end{aligned}$$

We can now turn to the proof of the Chaining Lemma.

Proof of Chaining Lemma: If $(i,j) \in E'$, then $H(j,k) \subseteq H(i,k+1)$, so if $Y(j,k) = X_h - X_j$ where $h \in H(j,k)$, then

$$Y(i,k+1) \geq X_h - X_i = Y(j,k) + X_j - X_i.$$

Thus, if $(i,j) \in M$,

$$Y(i,k+1) \geq Y(j,k) + 2 \tag{1}$$

since $X_i \leq -1$ and $X_j \geq 1$ given that (i,j) is in the matching.

Let N be an arbitrary pairing of vertices not in M. Then, for any $(i,j) \in N$,

$$\frac{1}{2}Y(i,k+1) + \frac{1}{2}Y(j,k+1) \geq \frac{1}{2}Y(i,k) + \frac{1}{2}Y(j,k). \tag{2}$$

Now we want to add both sides over all $(i,j) \in M \cup N$, take expectations and get Φ. Unfortunately, if we take an expectation, there will be a coefficient in front of $Y(i,k)$ of the probability that i is in the matching. To get around this issue, we introduce some new random variables. Let

$$L_i = \begin{cases} 1 & \text{if } i \text{ is matched in } M,\ i \in L, \\ 0 & \text{if } i \text{ is matched in } M,\ i \in R, \\ \frac{1}{2} & \text{otherwise} \end{cases}$$

and

$$R_i = \begin{cases} 1 & \text{if } i \text{ is matched in } M,\ i \in R, \\ 0 & \text{if } i \text{ is matched in } M,\ i \in L, \\ \frac{1}{2} & \text{otherwise.} \end{cases}$$

Note that $\mathbb{E}[L_i] = \mathbb{E}[R_i] = \frac{1}{2}$ since the probability that i is in the matching when $i \in L$ is the same that i is in the matching when $i \in R$. Adding both sides of (1) and (2) over M and N, we get

$$\sum_{i=1}^{n} Y(i,k+1)L_i \geq \sum_{j=1}^{n} Y(j,k)R_j + 2|M|. \tag{3}$$

By Lemma 6,

$$\begin{aligned} |\ \mathbb{E}[Y(i,k+1)L_i] - \mathbb{E}[Y(i,k+1)]\mathbb{E}[L_i]\ | &\leq \sqrt{\operatorname{Var}[Y(i,k+1)]\operatorname{Var}[L_i]} \\ &\leq \hat{c} \max_{j \in H(i,k+1)} \left(\mathbb{E}[(X_j - X_i)^2]\right)^{\frac{1}{2}}. \end{aligned}$$

Similarly, we have

$$|\ \mathbb{E}[Y(j,k)R_i] - \mathbb{E}[Y(j,k)]\mathbb{E}[R_i]\ | \leq \hat{c} \max_{i \in H(j,k)} \left(\mathbb{E}[(X_j - X_i)^2]\right)^{\frac{1}{2}}.$$

Taking expectation of both sides of (3), we get

$$\frac{1}{2}\Phi(k+1) \geq \frac{1}{2}\Phi(k) + 2\mathbb{E}[|M|] - \hat{c}n \cdot \max_{i,j \in H(i,k+1)} \left(\mathbb{E}[(X_i - X_j)^2]\right)^{\frac{1}{2}},$$

which implies that

$$\Phi(k+1) \geq \Phi(k) + 4\mathbb{E}[|M|] - 2\hat{c}n \cdot \max_{i,j \in H(i,k+1)} \left(\mathbb{E}[(X_i - X_j)^2]\right)^{\frac{1}{2}}.$$

□

We have now proven the Chaining Lemma, which implies Projection Lemma. As we claimed above, for the right choice of constants and $\Delta = \Omega(1/\sqrt{\log n})$, we get that

$$\left(\frac{\mathbb{E}[|M|]}{n}\right)^3 \leq \left(\left(\frac{c'}{2}\right)^2\right)^3,$$

which then proves the Structure Theorem.

Research Questions:

- Is there an easier proof? Or a Cheeger-like proof? Recall the connection to the Cheeger-like inequality over flow packings.
- Can this proof be extended to non-uniform sparsest cuts, where for each pair of (s_i, t_i), there is a demand d_i and
$$\rho(S) = \frac{|\delta(S)|}{\sum_{i:(s_i,t_i)\in\delta(S)} d_i}?$$
The sparsest cut problem corresponds to there being a unit demand between each pair of vertices. For the non-uniform case, it is known that there is an $O(\sqrt{\log n}\log\log n)$-approximation algorithm, but it is not known if the extra $\log\log n$ term is necessary.