Note: This is an altered version of the lecture I actually gave, which followed the structure of the Barak-Steurer proof carefully. With the benefit of some hindsight, I think the following rearrangement of the same elements would have been more effective.

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 $\leq O(1) \cdot \text{OPT}$. 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 $\Pr[|L|, |R| \geq c'n] \geq c'$.

---

*This lecture is derived from lecture notes of Boaz Barak and David Steurer [http://sumofsquares.org/public/lec-arv.html](http://sumofsquares.org/public/lec-arv.html)*
Theorem 3 (Structure Theorem) For $\Delta = \Omega(1/\sqrt{\log n})$, $E[|M|] \leq (\frac{c}{2})^2 n$.

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 $O(\sqrt{\log n})$-approximation. Today we turn to the proof of the Structure Theorem.

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 $\frac{v \cdot r}{\|v\|^2} \sim N(0,1)$; 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

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

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

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

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

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

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. We would like to prove the following lemma.

Lemma 4 There exists a constant $\tilde{c}$ such that for any positive integer $k$,

$$E\left[\max_{i,j \in U} (X_i - X_j)\right] \geq \frac{4k}{n} E[|M|] - \tilde{c}\sqrt{k\Delta}.$$
If we can prove this lemma, we have that with high probability
\[
\frac{4k}{n} \mathbb{E}[|M|] - \tilde{c}\sqrt{k}\Delta \leq C\sqrt{\ln n},
\]
or
\[
\frac{1}{n} \mathbb{E}[|M|] \leq \frac{C\sqrt{\ln n}}{4k} + \frac{\tilde{c}}{4}\sqrt{\frac{\Delta}{k}}.
\]
So if we set
\[
k = \left(\frac{2}{c'}\right)^2 C\sqrt{\ln n} = O(\sqrt{\ln n}),
\]
and
\[
\Delta = \frac{1}{c'^2 k} = \Omega\left(\frac{1}{\sqrt{\ln n}}\right),
\]
then
\[
\frac{1}{n} \mathbb{E}[|M|] \leq \frac{1}{4} \left(\frac{c'}{2}\right)^2 + \frac{1}{4} \left(\frac{c'}{2}\right)^2 \frac{1}{C\sqrt{\ln n}} \leq \left(\frac{c'}{2}\right)^2,
\]
and we will have proven the Structure Theorem.

How should we prove the lemma? Consider a graph \(H = (U, E')\) where \(E' = \{(i, j) \in U \times U : d(i, j) \leq \Delta\}\). 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,
\[
\frac{1}{n} \Phi(k) \leq \mathbb{E} \left[ \max_{i,j \in U} (X_i - X_j) \right].
\]
So to prove Lemma 4 we’ll instead prove that
\[
\frac{1}{n} \Phi(j) \geq \frac{4k}{n} \mathbb{E}[|M|] - \tilde{c}\sqrt{k}\Delta.
\]
Or rather, we’ll prove the following, which implies Lemma 4

**Lemma 5**
\[
\Phi(k) \geq 4k\mathbb{E}[|M|] - 2\tilde{c}n\sqrt{k}\Delta.
\]

To prove this lemma will need the following probability results.

**Lemma 6** For any two random variables \(X\) and \(Y\),
\[
|\mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]| \leq \sqrt{\text{Var}[X]\text{Var}[Y]}
\]

27-3
Observation 2  For any vector $x$,

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

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

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

Note that in Borell’s Theorem, there’s no dependence on the number of variables $t$. We also observe that

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

by the triangle inequality and the fact that each edge $(p, q)$ in $H$ has $\|v_p - v_q\|^2 \leq \Delta$. 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 that Borell’s Theorem says that

$$\text{Var}[Y(i, k)] = \text{Var} \left[ \max_{j \in H(i, k)} (X_j - X_i) \right] \leq \hat{c} \max_{j \in H(i, k)} \text{Var}[X_j - X_i]$$

$$= \hat{c} \max_{j \in H(i, k)} \mathbb{E}[(X_j - X_i)^2]$$

$$= \hat{c} \max_{j \in H(i, k)} \|v_j - v_i\|^2$$

$$\leq \hat{c} \cdot k\Delta.$$

Now to prove Lemma 5. But before we start, we can reflect a bit on what the lemma actually says. If we think about the expected projections of $X_j - X_i$ as we let $j$ be at most $k$ steps away from $i$, summing over all $i$, we get a constant times $E[|M|]$ for each of the steps; this makes sense, since for any matching edge $(p, q)$, we have that $|X_p - X_q| \geq 2$ since either $X_p \geq 1$ and $X_q \leq -1$ or vice versa, so we pick up that difference for each edge in the matching. However, there is also a correction term that corresponds to the variance. The proof is formalized below.

Proof of Lemma 5: If $(i, j) \in E'$, then $H(j, k-1) \subseteq H(i, k)$, so if $Y(j, k-1) = X_h - X_j$ where $h \in H(j, k)$, then

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

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

$$Y(i, k) \geq Y(j, k-1) + 2$$

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 $(i, j) \in N$,

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

Now we want to add both sides over all $(i, j) \in M \cup N$, take expectations and get $\Phi$. 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 $E[L_i] = E[R_i] = \frac{1}{2}$ 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

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

Similarly, we have that

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

Then by applying induction, we obtain that for $k$ odd

$$\sum_{i=1}^{n} Y(i,k)L_i \geq \sum_{j=1}^{n} Y(j,0)R_j + 2k|M| = 2k|M|$$

and for $k$ even

$$\sum_{i=1}^{n} Y(i,k)L_i \geq \sum_{j=1}^{n} Y(j,0)L_j + 2k|M| = 2k|M|,$$ 

so that

$$\sum_{i=1}^{n} Y(i,k)L_i \geq 2k|M| \quad (4)$$

for any $k$.

By Lemma 6,

$$|\mathbb{E}[Y(i,k)L_i] - \mathbb{E}[Y(i,k)]\mathbb{E}[L_i]| \leq \sqrt{\text{Var}[Y(i,k)] \cdot \text{Var}[L_i]} \leq \sqrt{\epsilon k \Delta}.$$ 

Taking expectation of both sides of (4), we get

$$\frac{1}{2} \Phi(k) \geq 2k\mathbb{E}[^{\cdot}M] - n\sqrt{\epsilon k \Delta},$$

or

$$\Phi(k) \geq 4k\mathbb{E}[^{\cdot}M] - 2n\sqrt{\epsilon k \Delta},$$

as desired. \qed
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.