15 Approximating α-balanced cut with ARV

As described in the previous lecture the following LP gives a relaxation to the c-balanced cut problem.

$$
\begin{aligned}
\text{minimize} \quad & \frac{1}{4}\sum_{i,j} \|v_i - v_j\|^2 \\
\text{subject to} \quad & \|v_i - v_j\|^2 = 1 && \forall i \\
& \|v_i - v_j\|^2 + \|v_j - v_k\|^2 \geq \|v_i - v_k\|^2 && \forall i,j,k \\
& \sum_{i<j} \|v_i - v_j\|^2 \geq cn^2
\end{aligned}
\tag{SDP}
$$

The output of this LP is a set X of n vectors x_i on the unit sphere S^{n-1} for which $\rho(x_i, x_j) := \|x_i - x_j\|_2^2$ is a squared ℓ_2-metric.

Theorem 15.1 (ARV04) *Let (X, ρ) be an n-point squared ℓ_2-metric over vectors from the unit sphere and let $\sum_{x,y \in X} \rho(x,y) \geq c \cdot n^2 > 0$. There exist subsets $L, R \subseteq X$ with $|L|, |R| = \Omega(c \cdot n)$ and $\rho(L, R) \geq \Omega(\frac{1}{\sqrt{\log n}})$.*

We have seen in the last lecture that the above structural theorem about squared ℓ_2-metrics implies a pseudo-approximation to c-balanced separators with an approximation guarantee of $O(\sqrt{\log n})$. In this lecture we prove Theorem 15.1.

We use the following algorithm to construct sets L and R:

Algorithm:

- choose a random hyperplane specified by a normal vector u chosen uniformly at random from S^{n-1}.
- define sets $L' := \{x \in X \mid \langle x, u \rangle \leq -\frac{\sigma}{\sqrt{n}}\}$ and $R' := \{x \in X \mid \langle x, u \rangle \geq \frac{\sigma}{\sqrt{n}}\}$, for a parameter σ which will be specified later.

  This means we choose the sets L' and R' as the set of points that fall on one side of the hyperplane but are not too close to it (at least at distance $\sigma/\sqrt{n}$).
- While there exist a pair x_l, x_r with $x_l \in L'$ and $x_r \in R'$ and $d(x_l, x_r) < \ell$ delete x_l from L' and x_r from R' (here $d(x_l, x_r)$ denotes the Euclidean (not squared) distance between x_l and x_r, and $\ell = \Theta(\frac{1}{\sqrt[4]{\log n}})$ is the Euclidean distance that we want to generate between the sets).

Clearly the above algorithm creates sets L and R that are at Euclidean distance ℓ and hence $\rho(L, R) \geq \Omega(\frac{1}{\sqrt{n}})$. The difficult part is to show that the sets contain a linear number of vertices as required by Theorem 15.1.

Lemma 15.2 *There is a choice of $\sigma = \Theta(c)$ such that the sets L' and R' contain at least $\frac{c}{32}n$ nodes with probability $1/4$.*

Proof. In the analysis of their Maxcut-algorithm, Goemanns and Williamson show that the probability that a random hyperplane seperates x and y is

$$\mathbf{Pr}[\text{a random hyperlane sep. } x,\, y] \geq 0.878 \cdot \frac{\|x-y\|^2}{4} \quad .$$

Therefore the spreading-constraint of our LP ($\sum_{x,y} \|x-y\|_2^2 \geq c \cdot n^2$) gives that the expected number of separated pairs is

$$\mathbf{E}[\text{separated pairs}] \geq (0.878 \cdot \frac{c}{4}) \cdot n^2 \geq \frac{c}{8} \cdot n^2 \quad .$$

Using Markov we get that with probability $\frac{1}{2}$ at least $\frac{c}{16} \cdot n^2$ pairs are separated.

Since $\frac{c}{16} \cdot n^2$ pairs are separated, the smaller side of the hyperplane contains at least $\frac{c}{16} \cdot n$ points. Now, we have to take care of the points that fall into the stripe with width 2σ around the hyperplane. For this we use the following bounds on the length of random projections.

Lemma 15.3 *Let $v \in \mathbb{R}^n$ denote a vector of length ℓ, and let u be a vector chosen uniformly at random from the unit sphere S^{n-1}. Then*

$$\mathbf{Pr}\left[|\langle v, u\rangle| \geq \frac{\sigma}{\sqrt{n}}\right] \leq e^{-\sigma^2/4\ell^2} \text{ and } \mathbf{Pr}\left[|\langle v, u\rangle| \leq \frac{\sigma}{\sqrt{n}}\right] \leq \frac{3\sigma}{\ell} \quad .$$

We only need the second bound here. All vectors in X have unit length. If we choose $\sigma : -\frac{c}{3 \cdot 128}$ the probability of a point falling into the stripe is at most $\frac{c}{128}$. Therefore the expected number of points ending up in the stripe is $\frac{c}{128}n$. By Markov, with probability $3/4$ at most $\frac{c}{32}n$ nodes will end in the stripe.

Now, with probability $1/2$ both sides of the hyperlane contain at least $\frac{c}{16}n$ nodes, and with probability $3/4$ at most $\frac{c}{32}n$ nodes are in the stripe. Consequently, if both events hold (with probability at least $1/4$) both L' and R' contain at least $\frac{c}{32}n$ nodes. ∎

In order to complete the proof of the main theorem we have to show that during the deletion process not too may nodes are deleted from L' and R'. For a specific random direction u, the pairs that are deleted from L' and R' form a partial matching over the node set X. We denote this matching with $M(u)$. If we can prove that with large probability (larger than say $4/5$) the cardinality of this matching is small (less than $\frac{c}{64}n$) the theorem follows.

We will prove this statement by contradicition. Assume that

With probability at least $1/5$ the matching $M(u)$ contains more than $\frac{c}{64}n$ pairs. (Assumption)

Using this assumption we will construct a subset of nodes with certain properties (a so-called core). From the properties of the core we can then deduce a lower bound on the size of the core. We will finally show that a length of $\ell = \Theta(\frac{1}{\sqrt[4]{\log n}})$ (in the above algorithm for constructing L and R) results in a core with size larger than n (if the assumption holds). This shows that the assumption is wrong and hence we can chose $\ell = \Theta(\frac{1}{\sqrt[4]{\log n}})$ such that both L and R contain a linear number of points.

The details are as follows. From the assumption we get that the average probability for a node $x \in X$ to be contained in a matching is large.

$$\sum_{x\in X} \mathbf{Pr}[x \text{ is matched}] \geq \frac{1}{5} \cdot \frac{c}{64} n$$

Now, we construct a sub-set $C \subset X$ in which all nodes are matched with large probability *and* (even stronger) all nodes are matched with large probability to a buddy that is in C. We initialize $C := X$ and then we successively delete vertices from C that are not matched within C with a large probability. Formally, we delete nodes x' for which $\mathbf{Pr}[x' \text{ is matched within } C] \leq \frac{1}{5} \cdot \frac{c}{256}$. Each deletion changes the value of $\sum_{x\in C} \mathbf{Pr}[x \text{ is matched within } C]$ by at most $\frac{1}{5} \cdot \frac{c}{128}$.

A term of $\frac{1}{5} \cdot \frac{c}{256}$ comes from the probability of node x' (which we delete) and further the probability for buddies of x' is reduced (since for those direction for which x' was their buddy they are now not matched within C any more). This adds another loss of $\frac{1}{5} \cdot \frac{c}{256}$ in the total probability.

In total the above procedure only deletes a total probability mass of at most $\frac{1}{5} \cdot \frac{c}{128} n$ and hence guarantees that C is not empty in the end.

We call the set C created by this procedure a core.

Definition 15.4 *A (σ, δ, ℓ)-core is a set $C \subset X$ such that for every node $x \in C$ it holds that for a δ-fraction of directions u*

- *x is matched to buddy $x' \in C$,*
- *x' is close to x: $\|x - x'\|_2 \leq \ell$, and*
- *$|\langle x - x' \rangle, u| \geq \sigma/\sqrt{n}$.*

Observation 15.5 *The set C that we constructed fulfills the above definition with parameters σ,ℓ and $\delta = \frac{1}{5} \cdot \frac{c}{256}$.*

Proof. A matching pair (x, x') is always a pair with large projection $|\langle x - x' \rangle| \geq \frac{2\sigma}{\sqrt{n}}$ because both points lie on opposite sides of the hyperplane and not within the stripe. Further, the distance between x and x' is less than ℓ as otherwise they would not have been matched. ■

The following definition is closely related to a core.

Definition 15.6 *A set C is a (σ, δ, ℓ)-cover for a node x if*

$$\mathbf{Pr}_{u \in S^{n-1}}\left[\exists x' \in C : \langle x - x', u \rangle \geq \frac{\sigma}{\sqrt{n}} \text{ and } \|x - x'\|_2 \leq \ell\right] \geq \delta \ .$$

The definition of a cover is less restrictive as it does not require a matching between nodes x and x', but more restrictive in the sense that it requires $\langle x - x', u\rangle \geq \frac{\sigma}{\sqrt{n}}$ instead of $|\langle x - x', u\rangle| \geq \frac{\sigma}{\sqrt{n}}$. The following lemma gives a lower bound on the size of a cover.

Lemma 15.7 *If a node x is (σ, δ, ℓ)-covered by C, then*

$$|C| \geq \delta \cdot e^{\sigma^2/4\ell^2} \quad .$$

Proof. Consider the projection bounds of Lemma 15.3. For a fixed buddy $x' \in C$

$$\mathbf{Pr}\left[\langle x - x', u\rangle \geq \frac{\sigma}{\sqrt{n}}\right] \leq e^{-\sigma^2/4\ell^2}$$

since the distance to the buddy is at most ℓ. Therefore any fixed buddy can only cover x for an $e^{-\sigma^2/4\ell^2}$-fraction of directions. Since we require x to be covered for at least a δ-fraction of directions we get $|C| \cdot e^{-\sigma^2/4\ell^2} \geq \delta$ which gives the lemma. ■

Note that the above lemma also holds for our (σ, δ, ℓ)-core $|C|$ (with the same proof). Since σ is a constant we can choose $\ell = \Omega(\frac{1}{\sqrt{\log n}})$ and the lemma would give $|C| \geq n$, which is a contradiction. However, we want to obtain a contradiction for $\ell = \Theta(\frac{1}{\sqrt[4]{\log n}})$. For this, one has to do more work.

The idea is to use the core to create a cover with better parameters (parameters that give us a better lower bound). We will do this step by step. We first need some technical lemmas that show trade-offs between the parameters of a cover.[1] The first lemma shows that we can boost the probability of a cover while losing a little bit in the projection-length.

Lemma 15.8 *Let S be a set that is (σ, δ, ℓ)-covered by C. Then S is also $(\sigma - \ell\gamma, 1 - e^{-t^2/2}, \ell)$-covered by C with $\gamma \geq \sqrt{2\log(2/\delta)} + t$.*

The proof of this lemma can be found in [ARV04].

15.1 Length $\Theta(1/\sqrt[3]{\log n})$

We first show how to obtain a contradiction for $\ell = \Theta(\frac{1}{\sqrt[3]{log n}})$.

Lemma 15.9 *For $0 \leq r \leq R$ there eixst a set S_r with $|S_r| \geq (\frac{\delta}{8})^r \cdot |C| \geq (\frac{\delta}{8})^r \cdot |C| \geq (\frac{\delta}{8})^{r+1} \cdot |X|$ that is $(\frac{\sigma}{2}r, 1 - \frac{\delta}{8}, \ell \cdot \sqrt{r})$-covered by C.*

Proof. We prove this by induction

$(r = 0)$ We choose $S_0 = C$. Assume wlog that $\ell \geq \frac{\sigma}{2\sqrt{\ln n}}$ (this is ok since we aim for $\ell = \Theta(1/\sqrt[3]{\log n})$). Then Lemma 15.7 gives that $|C| \geq \delta n \geq \frac{\delta}{8} \cdot |X|$. Further, each set is a $(0, 1, 0)$-cover of itself.

[1] Note that there are three parameters. The length of the projection, the probability and the distance to the covering nodes. We want the projection-length to be large, the probability to be large and the distance to be small.

$(r-1 \to r)$ The induction step takes a set that fulfills the induction hypothesis and constructs a new set that has a larger projection-length (increase by $\sigma/2$) but for which the distance-parameter increases only slightly (the squared distance increases by ℓ^2) We obtain the set S_r by extending S_{r-1} along the core C.

Assume a node $x \in S_{r-1}$ has for some direction u a cover node y in C with $\langle x-y, u\rangle \geq \frac{\sigma(r-1)}{2\sqrt{n}}$ and for direction $-u$ a cover node y' with $\langle x-y', -u\rangle \geq \frac{\sigma(r-1)}{2\sqrt{n}}$. For every node in S_{r-1} this holds for a $(1-\frac{\delta}{2})$-fraction of directions, and the cover nodes are at distance at most $\ell\sqrt{r-1}$ from x.

Now, assume further that x has for direction u also a buddy x' with respect to the core. This means $|\langle x-x', u\rangle| \geq \frac{\sigma}{\sqrt{n}}$, and $\|x-x'\| \leq \ell$. We now assign either y or y' as a cover node for direction u or $-u$ to x'. If $\langle x-x', u\rangle > 0$ we assign y' as a cover node for direction $-u$ and otherwise we assign y' as a cover node for direction u. This gives that the projection length is in the first case $\langle x'-y', -u\rangle = \langle x'-x+x-y', -u\rangle = \langle x-x', u\rangle + \langle x-y', -u\rangle \geq \frac{\sigma}{\sqrt{n}} + \frac{\sigma(r-1)}{2\sqrt{n}} = \frac{\sigma r}{2\sqrt{n}} + \frac{\sigma}{2\sqrt{n}}$. For the second case we obtain the same projection length. For the distance $d(x', y)$ between a node x' and its cover node y we have $(d(x', y))^2 = \|x'-y\|_2^2 \leq \|x'-x\|_2^2 + \|x-y\|_2^2 \leq \ell^2 + (r-1)\ell^2 \leq r\ell^2$, and hence $d(x', y) \leq \sqrt{r}\ell$.

How many cover nodes are assigned by the above process. With probability $(1-\frac{\delta}{2})$ a node x fulfills the first constraint that it has cover nodes for directions u and $-u$ and with probability δ it also has a matching buddy for direction u. Therefore with probability at least $\delta/2$ both events hold and x will assign a cover node to its buddy x'. Hence, in total at least $\frac{\delta}{2} \cdot |S_{r-1}|$ cover nodes are generated. Observe that a node x' may obtain at most two cover nodes for direction u by the process. (One from its buddy in direction u and one from its buddy in direction $-u$). In this case we delete one of the cover nodes, arbitrarily which reduces our number of cover nodes to $\frac{\delta}{4} \cdot |S_{r-1}|$.

This means that on average a node in X is assigned a cover node for a $\frac{\delta \cdot |S_{r-1}|}{4|X|}$-fraction of directions. By Markov, at least $\frac{\delta |S_{r-1}|}{8}$ nodes have at least $\frac{\delta |S_{r-1}|}{8|X|} \geq (\frac{\delta}{8})^r$ directions covered. We define these nodes as the set S_r. Note that $|S_r| \geq \frac{\delta}{8}|S_{r-1}| \geq (\frac{\delta}{8})^r$. So far we have constructed a set S_r that is $(\frac{\sigma}{2}r + \frac{\sigma}{2}, (\frac{\delta}{8})^r, \sqrt{r}\ell)$ covered. We now apply Lemma 15.8 to increase the probability of the cover. For the induction to work we can only afford to lose $\frac{\sigma}{2}$ in the projection length. This means we require that the product of γ and the vector length $\sqrt{r}\ell$ must be less than $\sigma/2$. Further, we want to increase the probability to $1-\delta/4$. This means we need to set $t := \sqrt{2\ln 4/\delta}$. Altogether we have to set

$$\gamma := \sqrt{2\log 2/(\frac{\delta}{8})^r} + \sqrt{2\ln 4/\delta} \leq 2\sqrt{2\log 2/(\frac{\delta}{8})^r} = \Theta(\sqrt{r}) \ .$$

We can ensure the constraint $\sqrt{r}\ell\gamma \leq \sigma/2$ for $r \leq R := O(1/\ell)$.

■

Now we can apply the above lemma for $R = \Theta(1/\ell)$. We obtain a $(\Theta(1/\ell), 1-\delta/2, \sqrt{\ell})$-cover. Now the lower bound of Lemma 15.7 gives that the size of such a cover is $e^{\Theta(1/\ell^3)}$, which means we get a contradiction for $\ell \geq \Omega(\frac{1}{\sqrt[3]{\log n}})$.

References

[ARV04] Sanjeev Arora, Satish Rao, and Umesh Vazirani. Expander flows, geometric embeddings, and graph partitionings. In *Proceedings of the 36th ACM Symposium on Theory of Computing (STOC)*, pages 222–231, 2004.

[Lee05] James R. Lee. On distance scales, embeddings, and efficient relaxations of the cut cone. In *Proceedings of the 16th ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 92–101, 2005.