The Sparsest Cut Problem

**Input:**
- Undirected $G = (V, E)$
- Costs $c_e$ on edge

**Output:** $S \subseteq V$

**Objective:** Minimize $\frac{\sum_{e \in \delta(S)} c(e)}{|S| \cdot |V - S|}$

LP Relaxation

As last time, we have "distance" variables $d(u,v)$ where in the "integral" soln, $d(u,v) = 1$ if $u,v$ are in separate sides of the cut $e$. 

Let $OPT \geq \min \frac{\sum_{e \in E} c(e) d(e)}{\sum_{u,v \in V \times V} d(u,v)}$

"$d$ satisfies $\Delta$-ineq" precisely $|S| \cdot |V - S|$ if $d(u,v) = 1$ A pairs $\{(u,v) \in S \times (V - S)\}$

Ratio $\leq$ Constraint

$OPT \geq \min \sum_{e \in E} c(e) d(e)$

- $\sum_{u,v \in V \times V} d(u,v) = 1$
- $d$ satisfies metric constraints.
ROUNDING - Try 1

- Solve LP
- Pick a node $s$ (arbitrarily for now; we will choose appropriately later)
- Let $R := \max_{v} d(s,v)$
- Sample $\rho \in [0,R]$ u.a.r.
- $S := \{ u \mid d(s,u) \leq \rho \}$

As last time,

$$\mathbb{E} \left[ c(\rho S) \right] = \frac{1}{\rho} \cdot \text{LP}$$

But this time we "also have a denominator!"

Ideally, we would’ve liked to analyze

$$\mathbb{E} \left[ \frac{c(\rho S)}{|S| \cdot |\overline{S}|} \right]$$

This is a rather unwieldy object. Instead let’s look at

$$\frac{\mathbb{E} \left[ c(\rho S) \right]}{\mathbb{E} \left[ |S| \cdot |\overline{S}| \right]}$$

Note these two expressions are NOT the same!

Why is $\odot$ interesting?

\[ \therefore \text{if we show} \quad \frac{\mathbb{E} \left[ c(\rho S) \right]}{\mathbb{E} \left[ |S| \cdot |\overline{S}| \right]} \leq \alpha \cdot \text{LP} \]
Then, by moving terms & LINEARITY OF EXPECTATION,

$$\mathbb{E} \left[ c(\bar{x}_S) - \alpha \cdot \text{LP} \cdot |S| \cdot |\bar{S}| \right] \leq 0$$

⇒ One of the possible cuts that our algorithm returns satisfies

$$c(\bar{x}^*) \leq \alpha \cdot \text{LP} \cdot |S^*| \cdot |\bar{S}^*|$$

⇒ sparsity ($S^*$) ≤ α · LP

Why would we be able to get our hands on $S^*$?

⇒ Only a O(m) different $e$’s are “interesting” & we can deterministically enumerate.

⇒ It suffices to upper bound

$$\frac{\mathbb{E} \left[ c(\bar{x}_S) \right]}{\mathbb{E} \left[ |S| \cdot |\bar{S}| \right]}$$

The numerator we already know to be $\text{LP}/R$

- The denominator is again a bit problematic since $|S|$ & $|\bar{S}|$ are not independent (in fact they sum to $n$)

⇒ For now we use a trivial LOWER BOUND on $|S|$, ie, $|S| \geq 1$

$$\mathbb{E} \left[ |S| \cdot |\bar{S}| \right] \geq \mathbb{E} \left[ |S| \right] \quad \text{(#)}$$
The $\mathbb{E} [1 \bar{S}]$ we can calculate.

$$
\mathbb{E} [1 \bar{S}] = \frac{1}{R} \sum_s \left( \text{# of rejections at } d(s,v) > p \right)
$$

\[ \text{this really should be an integration.} \]

\[
= \frac{1}{R} \sum_s \sum_{v : d(s,v) > p} 1 = \frac{1}{R} \sum_{v} \sum_{s \in S} 1
\]

\[
= \frac{1}{R} \cdot \sum_{v} d(s,v)
\]

- What do we know about $d$'s?

\[
\sum_{u,v} d(u,v) = 1
\]

\[
\therefore \exists s \text{ s.t. } \sum_{v} d(s,v) \geq \frac{1}{n}
\]

- If our algorithm starts from this ... 

\[
\mathbb{E} [1 \bar{S}] \geq \frac{1}{n R}
\]

\[
\therefore \mathbb{E} \left[ \mathcal{C}(\bar{S}) \right] \leq n \cdot LP
\]
\[ \frac{L}{\mathbb{E}[|S|]} \leq n \cdot LP \]

\[ \Rightarrow \text{An n-factor algorithm} \]

Try 2

- The algorithm started cutting around a singleton, \( S \), where \( |S| \geq 1 \) was the only thing we could use.

- Modification:
  - Let \( T \subseteq V \) s.t. \( |T| \approx \Theta(n) \); say \( |T| \geq \frac{n}{3} \)
  - \( R := \max_{v} d(T,v) \quad ; \quad g \in \mathbb{R} \cap (0,R] \)
    \[ \min_{u \in T} d(u,v) \]
  - \( S := \{ u \mid d(T,u) \leq g \} \)

Once again:
\[ \mathbb{E}\left[ \text{cost}(gS) \right] \leq \frac{1}{R} \cdot LP \]

- \( |S| \geq |T| \geq \frac{n}{3} \quad \therefore T \subseteq S \)
- What about \( \mathbb{E}[|S|] \)?

As before,
\[ \mathbb{E}[|S|] = \frac{1}{R} \sum \left( \# \text{ of } v \text{'s s.t. } d(T,v) \geq g \right) \]

\[ = \frac{1}{R} \sum_{g} \sum_{u: d(T,u) \geq g} 1 \]
\[
= \frac{1}{R} \sum_{u,v} \frac{1}{d(T,u) + d(T,v)}
\]
\[
\geq \frac{1}{R} \sum_{u,v} \frac{1}{d(T,u) + \diam(T) + d(T,v)}
\]
\[
\geq \frac{1}{R} \sum_{u,v} \frac{1}{(d(T,u) + \diam(T) + d(T,v))}
\]
\[
= \binom{n}{2} \cdot \diam(T) + 2n \cdot \sum_{u} d(T,u)
\]

\[\text{If } \diam(T) \leq \frac{1}{2n^2}, \sum_{u} d(T,u) \geq \frac{1}{4n} \]

Indeed, if such a \( T \) exists, then

\[\text{from } \bigstar \bigstar \text{ we get } \mathbb{E}\left|1\bar{S}\right| \geq \frac{1}{4nR} \]
\[\Rightarrow \sqrt{\mathbb{E}|1S - 1\bar{S}|} \geq \Theta\left(\frac{1}{n}\right) \]
$\Rightarrow E[|S_1 - 1S_1|] \geq \Omega\left(\frac{1}{R}\right)$

$\Rightarrow \frac{E[cost(\alpha S)]}{E[|S_1 - 1S_1|]} = O(L_{LP})$

$\Rightarrow$ An $O(1)$-approx for sparsest cut.

To conclude, we would be happy if there was a $T \subseteq V$ s.t.
- $|T| = \Omega(n)$
- $\delta(T) \leq \frac{1}{2n^2}$

**Low-diameter Decomposition Lemma** (Leighton-Rao '88)

Given any undirected $G = (V, E)$ with costs $c(e)$ on edges and distances $d : V \times V \rightarrow \mathbb{R}_{\geq 0}$ s.t.

$L = \sum_{e} c(e) \cdot d(e) \quad \text{and} \quad \sum_{u \neq v} d_{uv} = 1,$

and given any $R > 0$, there exists a partition $S_1, S_2, \ldots, S_k \subseteq V$ s.t.

1. $\delta(S_i) \leq R$

2. Total cost of cross-edges, i.e. edges with epts in diff $S_i$'s is $\leq O\left(\frac{\log n}{R}\right) \cdot L$