Lecture 3

Let $X$ be an $n$-point set. For any $S \subseteq X$, define

$$1_S(x) = \begin{cases} 
1 & \text{if } x \in S \\
0 & \text{otherwise}
\end{cases}$$

Note that $S$ defines a semi-metric $\rho_S(x,y) = |1_S(x) - 1_S(y)|$, for $x,y \in X$.

### 3.1 When $X$ is on a Line

Let $X = \{x_1, x_2, \ldots, x_n\}$, where $x_i \in \mathbb{R}$ and $x_1 \leq x_2 \leq \cdots \leq x_n$.

**Claim 1.** There exist $\{\lambda_S\}_{S \subseteq X}$, $\lambda_S \geq 0$, such that $|x - y| = \sum_{S \subseteq X} \lambda_S \rho_S(x,y)$ for any $x,y \in S$.

**Proof.** Let $S_i = \{x_1, \ldots, x_i\}$ and $\lambda_{S_i} = |x_i - x_{i+1}|$, $i = 1, \ldots, n-1$. For any other $S \subseteq X$, let $\lambda_S = 0$. Then for any $x_i,x_j \in X$,

$$|x_i - x_j| = \sum_{k=i}^{j-1} |x_k - x_{k+1}| = \sum_{\ell=1}^{n-1} |x_{\ell} - x_{\ell+1}| \cdot \rho_{S_{\ell}}(x_i, x_j) = \sum_{S \subseteq X} \lambda_S \rho_S.$$

### 3.2 MET($X$), CUT($X$), and $M_1(X)$

Define

- $\text{MET}(X) = \{d : X \times X \to \mathbb{R}_+ \mid d \text{ is a semi-metric on } X\}$
- $\text{CUT}(X) = \{d : X \times X \to \mathbb{R}_+ \mid \exists \lambda_S \geq 0 \text{ s.t. } d(x,y) = \sum_{S \subseteq X} \lambda_S \rho_S(x,y), \text{ for } \forall x,y \in X\}$
- $M_1(X) = \{d : X \times X \to \mathbb{R}_+ \mid \exists f : X \to L_1 \text{ s.t. } d(x,y) = \|f(x) - f(y)\|_1, \text{ for } \forall x,y \in X\}$

**Theorem 2.** $\text{CUT}(X) = M_1(X) \subseteq \text{MET}(X)$.

**Proof.** It is easy to see that for any $d \in \text{CUT}(X)$, $d \in \text{MET}(X)$, thus $\text{CUT}(X) \subseteq \text{MET}(X)$.

For any $d \in \text{CUT}(X)$, where $d(x,y) = \sum_{S \subseteq X} \lambda_S \rho_S(x,y)$. Define $f : X \to \ell_1^n$ as follows: $f(x) = \{\lambda_S 1_S(x)\}_{S \subseteq X}$. Thus,

$$\|f(x) - f(y)\|_1 = \sum_{S \subseteq X} |\lambda_S 1_S(x) - \lambda_S 1_S(y)| = \sum_{S \subseteq X} \lambda_S \rho_S(x,y) = d(x,y)$$

Thus, $d \in M_1(X)$ and $\text{CUT}(X) \subseteq M_1(X)$.

On the other hand, for any $d \in M_1(X)$, we know $d(x,y) = \|f(x) - f(y)\|_1$ for some $f : X \to L_1$. Observe that $\|f(x) - f(y)\|_1 = E_f |f(x) - f(y)|$ where $|f(x) - f(y)|$ is the distance over the line. For any fixed $f$, according to Claim 1, there exist $\{\lambda_S^f \geq 0\}_{S \subseteq X}$ such that

$$|f(x) - f(y)| = \sum_{S \subseteq X} \lambda_S^f \rho_S(x,y).$$
Therefore,
\[ d(x, y) = \| f(x) - f(y) \|_1 = E_f | f(x) - f(y) | = E_f \left( \sum_{S \subseteq X} \lambda_S \rho_S(x, y) \right) = \sum_{S \subseteq X} E_f (\lambda_S) \rho_S(x, y) \]

Hence, \( d \in \text{CUT}(X) \), implying \( \text{CUT}(X) = M_1(X) \).

### 3.3 Sparsest Cut

Given a graph \( G = (V, E) \) and \( S \subseteq V \), let \( \alpha(S) = \frac{|E(S, \overline{S})|}{|S| \cdot |\overline{S}|} \), where \( E(S, \overline{S}) \) is the set of edges between \( S \) and \( \overline{S} \). Let
\[
\alpha(G) = \min_{S \subseteq V} \alpha(S) = \min_{S \subseteq V} \frac{|E(S, \overline{S})|}{|S| \cdot |\overline{S}|}
\]

Define the edge expansion of \( G \) by
\[
\Phi(G) = \min_{S \subseteq V, |S| \leq |V|/2} \frac{|E(S, \overline{S})|}{|S|}
\]

It is easy to see that \( \frac{\Phi(G)}{n} \leq \alpha(G) \leq 2 \frac{\Phi(G)}{n} \). Thus, a good approximation of \( \alpha(G) \) gives a good approximation to \( \Phi(G) \) as well.

Let
\[
\alpha'(G) = \min_{d \in \text{MET}(V)} \frac{\sum_{xy \in E} d(x, y)}{\sum_{x,y \in V} d(x, y)}
\]

**Claim 3.** \( \alpha(G) = \alpha'(G) \).

**Proof.** Note that
\[
\alpha(G) = \min_{S \subseteq V} \frac{\sum_{xy \in E} \rho_S(x, y)}{\sum_{x,y \in V} \rho_S(x, y)} \implies \alpha'(G) \leq \alpha(G)
\]

On the other hand, for any \( d \in \text{CUT}(V) \), where \( d(x, y) = \sum_{S \subseteq V} \lambda_S \rho_S(x, y) \), we have
\[
\frac{\sum_{xy \in E} d(x, y)}{\sum_{x,y \in V} d(x, y)} = \frac{\sum_{xy \in E} \sum_{S \subseteq V} \lambda_S \rho_S(x, y)}{\sum_{x,y \in V} \sum_{S \subseteq V} \lambda_S \rho_S(x, y)} = \frac{\sum_{S \subseteq V} \lambda_S \sum_{xy \in E} \rho_S(x, y)}{\sum_{S \subseteq V} \lambda_S \sum_{x,y \in V} \rho_S(x, y)} \geq \min_{S \subseteq V} \lambda_S \sum_{xy \in E} \rho_S(x, y) = \alpha(G)
\]

where the inequality is due to the fact that for any \( a_1, \ldots, a_n, b_1, \ldots, b_n \geq 0 \), we have \( \sum_{i=1}^{n} a_i \geq \min_i \frac{a_i}{b_i} \).

Define
\[
\beta(G) = \min_{d \in \text{MET}(V)} \frac{\sum_{xy \in E} d(x, y)}{\sum_{x,y \in V} d(x, y)}
\]
Theorem 4. $\beta(G) \leq \alpha(G) \leq O(\log n) \cdot \beta(G)$.

Proof. Let $d^*$ be the optimal solution of $\beta(G)$ and $f : (V, d^*) \to L_1$. Observe that

\[
\sum_{xy \in E} \frac{\|f(x) - f(y)\|_1}{d^*(x, y)} \leq \max_{xy \in E} \frac{\|f(x) - f(y)\|_1}{d^*(x, y)} \leq \|f\|_{\text{Lip}}
\]

and

\[
\frac{\sum_{x, y \in V} d^*(x, y)}{\sum_{x, y \in V} \|f(x) - f(y)\|_1} \geq \min_{x, y \in V} \frac{d^*(x, y)}{\|f(x) - f(y)\|_1} \geq \frac{1}{\|f\|_{\text{Lip}} \sum_{x, y \in V} d^*(x, y)}
\]

Therefore,

\[
\alpha(G) = \alpha'(G) = \min_{d \in M_1(V)} \frac{\sum_{xy \in E} d(x, y)}{\sum_{x, y \in V} d(x, y)} \leq \sum_{xy \in E} \frac{\|f(x) - f(y)\|_1}{\sum_{x, y \in V} \|f(x) - f(y)\|_1} \leq \frac{\|f\|_{\text{Lip}} \sum_{xy \in E} d^*(x, y)}{\|f^{-1}\|_{\text{Lip}} \sum_{x, y \in V} d^*(x, y)} = \frac{\|f\|_{\text{Lip}} \|f^{-1}\|_{\text{Lip}} \cdot \beta(G)}{\text{dist}(f) \cdot \beta(G)}
\]

By Bourgain’s theorem, we can choose $f$ to have $\text{dist}(f) = O(\log n)$, hence $\alpha(G) \leq O(\log n) \cdot \beta(G)$. \hfill \Box

It remains to show that $\beta(G)$ can be computed efficiently. Consider the following linear program, where variables are $d(x, y)$ for any $x, y \in V$.

\[
\min \sum_{xy \in E} d(x, y)
\]

s.t. $d(x, y) \geq 0$, for $\forall x, y \in V$

$d(x, y) = d(y, x)$, for $\forall x, y \in V$

$d(x, y) + d(y, z) \geq d(x, z)$, for $\forall x, y, z \in V$

\[
\sum_{x, y \in V} d(x, y) = 1
\]

where the last constraint normalize the denominator of $\beta(G)$ to be one. By solving the above linear program, we have a $O(\log n)$ approximation algorithm to $\alpha(G)$.

3.4 Bourgain’s Theorem is Tight ($p = 1$)

We will construct a set of graphs $G$ for which $\alpha(G) \geq \Omega(\log n) \cdot \beta(G)$. Consider an $r$-regular expander graph $G = (V, E)$, where $r = O(1)$. We know there is constant $\delta > 0$ such that for any $S \subseteq V$, $|S| \leq |V|/2$, $|E(S, \bar{S})| \geq \delta |S|$. Therefore,

\[
\alpha(G) = \min_{S \subseteq V} \frac{|E(S, \bar{S})|}{|S| \cdot |\bar{S}|} \geq \Omega \left( \frac{\delta}{n} \right) = \Omega \left( \frac{1}{n} \right)
\]
Define \( d(x, y) \) to be the shortest path distance between \( x \) and \( y \) in \( G \). Observe that
\[
\sum_{xy \in E} d(x, y) = |E| = nr/2
\]
and
\[
\sum_{x, y \in V} d(x, y) \geq \Omega \left( n^2 \log n \right)
\]
because \( G \) has has \( O(1) \) maximum degree.

Therefore,
\[
\beta(G) = \min_{d \in \text{MET}(V)} \frac{\sum_{xy \in E} d(x, y)}{\sum_{x, y \in V} d(x, y)} \leq \frac{nr/2}{\Omega(n^2 \log n)} = \Omega \left( \frac{1}{n \log n} \right)
\]
which implies that \( \alpha(G) \geq \Omega(\log n) \beta(G) \).

### 3.5 Concurrent Multi-Commodity Flow

Concurrent multi-commodity flow problem is as follows: Given a graph \( G = (V, E) \), with capacity \( c : E \to \mathbb{R}_+ \), there is a set of demand pairs \((s_1, t_1), \ldots, (s_k, t_k)\), where each \( s_i, t_i \in V \) and has a demand \( d_i \in \mathbb{R}_+ \), representing the amount of flow that we are requested to send from \( s_i \) to \( t_i \). The goal is to maximize \( \epsilon(G) > 0 \) such that we can send \( \epsilon(G) \cdot d_i \) flow from \( s_i \) to \( t_i \) for any \( i = 1, \ldots, k \) simultaneously without violating the capacity constraints.

It is easy to see that
\[
\alpha(S) \triangleq \frac{\sum_{xy \in E} c(x, y) \rho_{S}(x, y)}{\sum_{i=1}^{k} d_i \rho_{S}(s_i, t_i)} \geq \epsilon(G)
\]
When \( k = 1 \), the Max-Flow/Min-Cut Theorem says that \( \alpha(G) = \epsilon(G) \). This is also known to hold for the case \( k = 2 \).

By a similar argument as above, one can show that \( \frac{\alpha(G)}{O(\log k)} \leq \epsilon(G) \leq \alpha(G) \) using Bourgain’s theorem and the next claim which follows again from LP duality.

**Theorem 5. (High dimension max-flow min-metric theorem)** Let
\[
\beta(G) = \min_{d \in \text{MET}(V)} \frac{\sum_{xy \in E} c(x, y) d(x, y)}{\sum_{i=1}^{k} d_i d(s_i, t_i)}
\]
Then \( \epsilon(G) = \beta(G) \).

**Open Problem:** What if \( G \) is planar graph? Do we have \( \epsilon(G) \geq \Omega(\alpha(G)) \)?