CS5275 – The Algorithm Designer’s Toolkit (S2 AY2025/26)

Lecture 3:

Expanders – metric embedding

Our goal: Proving Bourgain's theorem

Bourgain's theorem:

- Given any n-point metric space (X, d),
    - there is an embedding $f\colon X \to \mathbb{R}^{O(\log^2 n)}$ such that for any $x, y \in X$,

$$\|f(x) - f(y)\|_1 \leq d(x, y) \in O(\log n) \cdot \|f(x) - f(y)\|_1$$

Algorithm

Bourgain's theorem:

- Given any n-point metric space (X, d),
  - there is an embedding $f\colon X \to \mathbb{R}^{O(\log^2 n)}$ such that for any $x, y \in X$,

$$\|f(x) - f(y)\|_1 \leq d(x, y) \in O(\log n) \cdot \|f(x) - f(y)\|_1$$

Algorithm:

- Randomly sample $O(\log^2 n)$ subsets:

$$A_{i,j} \subseteq X \quad | \quad 1 \leq i \leq c \log n\,, \quad 1 \leq j \leq \log n$$

- Each point in X joins $A_{i,j}$ with probability 2^{-j} independently.
- Set $f_{i,j}(x) = d\left(x, A_{i,j}\right) = \min_{y \in A_{i,j}} d(x, y)$.

$$\|f(x) - f(y)\|_1 = \sum_{i,j} \left|f_{i,j}(x) - f_{i,j}(y)\right|$$

- For simplicity, assume $\log n$ is an integer.
- c is some constant.

Correctness

Bourgain's theorem:
- Given any n-point metric space (X, d),
  - there is an embedding $f\colon X \to \mathbb{R}^{O(\log^2 n)}$ such that for any $x, y \in X$,
$$\|f(x) - f(y)\|_1 \le d(x, y) \in O(\log n) \cdot \|f(x) - f(y)\|_1$$

Algorithm:
- Randomly sample $O(\log^2 n)$ subsets:
$$A_{i,j} \subseteq X \quad | \quad 1 \le i \le c \log n, \quad 1 \le j \le \log n$$
- Each point in X joins $A_{i,j}$ with probability 2^{-j} independently.
- Set $f_{i,j}(x) = d\left(x, A_{i,j}\right) = \min_{y \in A_{i,j}} d(x, y)$.

$$\|f(x) - f(y)\|_1 = \sum_{i,j} \left|f_{i,j}(x) - f_{i,j}(y)\right|$$

Correctness:
With high probability, we have:
- **Part 1:** $\sum_{i,j} \left|f_{i,j}(x) - f_{i,j}(y)\right| \le d(x, y) \cdot c \log^2 n$
- **Part 2:** $\sum_{i,j} \left|f_{i,j}(x) - f_{i,j}(y)\right| \in d(x, y) \cdot \Omega(\log n)$

With a scaling $f_{i,j}(x) \leftarrow f_{i,j}(x) \cdot \frac{1}{c \log^2 n}$,
this proves Bourgain's theorem.

Correctness: Part 1

Without loss of generality, assume $d(x, A_{i,j}) \geq d(y, A_{i,j})$

Triangle inequality

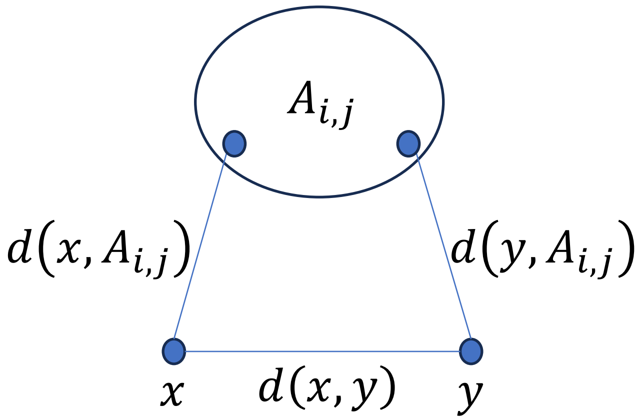
$$
\begin{aligned}
&\left|f_{i,j}(x) - f_{i,j}(y)\right| = \left|d(x, A_{i,j}) - d(y, A_{i,j})\right| \\
&= d(x, A_{i,j}) - d(y, A_{i,j}) \\
&\leq \left(d(x,y) + d(y, A_{i,j})\right) - d(y, A_{i,j}) \\
&= d(x,y)
\end{aligned}
$$

Algorithm:

- Randomly sample $O(\log^2 n)$ subsets:

  $$A_{i,j} \subseteq X \quad | \quad 1 \leq i \leq c \log n, \quad 1 \leq j \leq \log n$$

- Each point in X joins $A_{i,j}$ with probability 2^{-j} independently.
- Set $f_{i,j}(x) = d(x, A_{i,j}) = \min_{y \in A_{i,j}} d(x,y)$.



Correctness:

With high probability, we have:

- **Part 1:** $\sum_{i,j} \left|f_{i,j}(x) - f_{i,j}(y)\right| \leq d(x,y) \cdot c \log^2 n$
- **Part 2:** $\sum_{i,j} \left|f_{i,j}(x) - f_{i,j}(y)\right| \in d(x,y) \cdot \Omega(\log n)$

Correctness: Part 2

Algorithm:

- Randomly sample $O(\log^2 n)$ subsets:
$$A_{i,j} \subseteq X \quad | \quad 1 \le i \le c \log n \,, \quad 1 \le j \le \log n$$
- Each point in X joins $A_{i,j}$ with probability 2^{-j} independently.
- Set $f_{i,j}(x) = d\left(x, A_{i,j}\right) = \min_{y \in A_{i,j}} d(x, y)$.

Correctness:

With high probability, we have:

- **Part 1:** $\sum_{i,j} \left| f_{i,j}(x) - f_{i,j}(y) \right| \le d(x, y) \cdot c \log^2 n$
- **Part 2:** $\sum_{i,j} \left| f_{i,j}(x) - f_{i,j}(y) \right| \in d(x, y) \cdot \Omega(\log n)$

Next

Main tool: Chernoff bound

- Consider the following setting.
  - $X_1, X_2, \dots, X_n$ are independent random variables taking values in $\{0, 1\}$.
  - $X = \sum_{i=1}^{n} X_i$.
  - $\mu = \mathbb{E}[X]$.

Main tool: Chernoff bound

- Consider the following setting.
  - $X_1, X_2, \ldots, X_n$ are independent random variables taking values in $\{0, 1\}$.
  - $X = \sum_{i=1}^{n} X_i$.
  - $\mu = \mathbb{E}[X]$.

Chernoff bounds:

- $\Pr[X \geq (1+\delta)\mu] \leq e^{-\frac{\delta^2 \mu}{3}}$, for $1 \geq \delta \geq 0$.
- $\Pr[X \geq (1+\delta)\mu] \leq e^{-\frac{\delta^2 \mu}{2+\delta}}$, for $\delta \geq 0$.
- $\Pr[X \leq (1-\delta)\mu] \leq e^{-\frac{\delta^2 \mu}{2}}$, for $1 \geq \delta \geq 0$.

https://en.wikipedia.org/wiki/Chernoff_bound

Main tool: Chernoff bound

- Consider the following setting.
  - $X_1, X_2, \ldots, X_n$ are independent random variables taking values in $\{0, 1\}$.
  - $X = \sum_{i=1}^{n} X_i$.
  - $\mu = \mathbb{E}[X]$.

Chernoff bounds:

- $\Pr[X \geq (1+\delta)\mu] \leq e^{-\frac{\delta^2 \mu}{3}}$, for $1 \geq \delta \geq 0$.
- $\Pr[X \geq (1+\delta)\mu] \leq e^{-\frac{\delta^2 \mu}{2+\delta}}$, for $\delta \geq 0$.
- $\Pr[X \leq (1-\delta)\mu] \leq e^{-\frac{\delta^2 \mu}{2}}$, for $1 \geq \delta \geq 0$.

https://en.wikipedia.org/wiki/Chernoff_bound

Key takeaway:

- X is within a constant factor of $\mathbb{E}[X]$ with a probability of $1 - e^{-\Omega(\mathbb{E}[X])}$.
- In particular, $\mathbb{E}[X] \in \Omega(c \log n)$ guarantees a success probability of $1 - n^{-\Omega(c)}$.

Correctness: Part 2

- $B(v, r) =$ the ball of radius r around v.
- $B(v, < r) =$ the <u>open</u> ball of radius r around v.

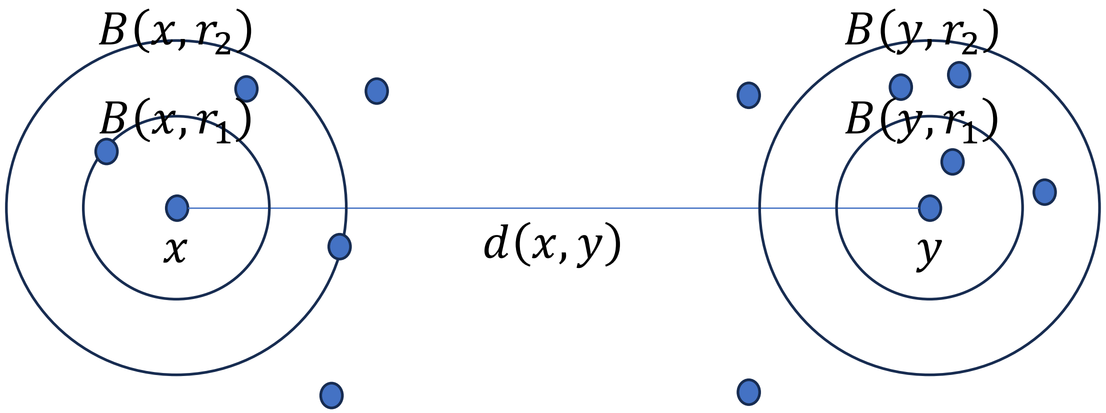
Correctness: Part 2

- $B(v, r)$ = the ball of radius r around v.
- $B(v, < r)$ = the open ball of radius r around v.

Consider two points x and y in X.

r_j = the minimum of:

- Smallest number r such that $\min\{|B(x,r)|, |B(y,r)|\} \geq 2^j$
- $\frac{d(x,y)}{3}$



Correctness: Part 2

- $B(v, r)$ = the ball of radius r around v.
- $B(v, < r)$ = the <u>open</u> ball of radius r around v.

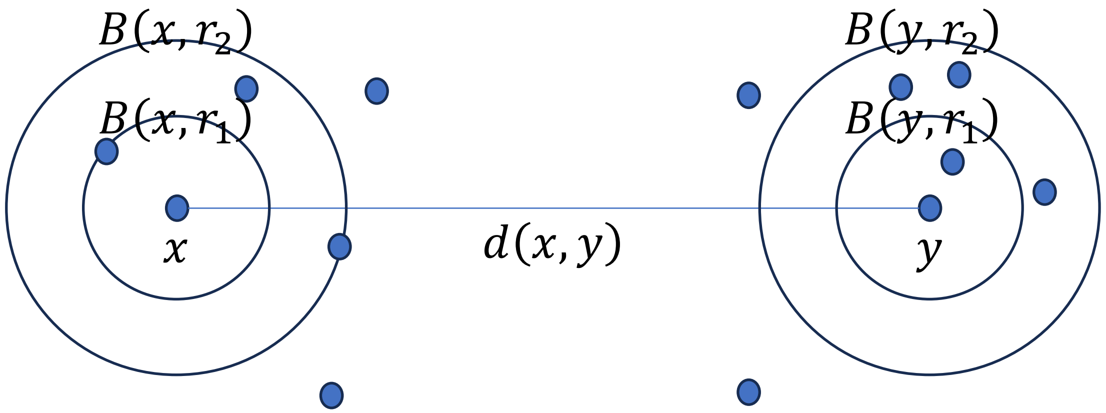
Consider two points x and y in X.

r_j = the minimum of:

- Smallest number r such that $\min\{|B(x,r)|, |B(y,r)|\} \geq 2^j$
- $\frac{d(x,y)}{3}$

Observation 1:
$|B(x, < r_j)| < 2^j$ or $|B(y, < r_j)| < 2^j$

Without loss of generality, we assume this.



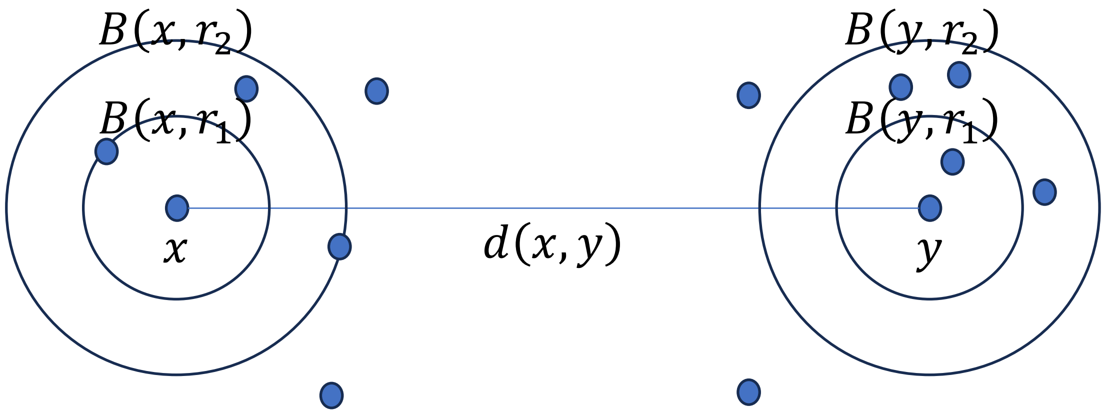
Correctness: Part 2

- $B(v, r)$ = the ball of radius r around v.
- $B(v, < r)$ = the open ball of radius r around v.

Consider two points x and y in X.

r_j = the minimum of:

- Smallest number r such that $\min\{|B(x,r)|, |B(y,r)|\} \geq 2^j$
- $\frac{d(x,y)}{3}$



Observation 1:
$|B(x, < r_j)| < 2^j$ or $|B(y, < r_j)| < 2^j$

Without loss of generality, we assume this.

Observation 2:
If $r_{j-1} < \frac{d(x,y)}{3}$, then
$|B(x, r_{j-1})| \geq 2^{j-1}$ and $|B(y, r_{j-1})| \geq 2^{j-1}$

Correctness: Part 2

- Each point in X joins $A_{i,j}$ with probability 2^{-j} independently.
- Set $f_{i,j}(x) = d(x, A_{i,j}) = \min_{y \in A_{i,j}} d(x,y)$.

- $B(v,r)$ = the ball of radius r around v.
- $B(v,<r)$ = the <u>open</u> ball of radius r around v.

Observation 1:
$|B(x,<r_j)| < 2^j$ or $|B(y,<r_j)| < 2^j$

Without loss of generality, we assume this.

Observation 2:
If $r_{j-1} < \frac{d(x,y)}{3}$, then
$|B(x, r_{j-1})| \geq 2^{j-1}$ and $|B(y, r_{j-1})| \geq 2^{j-1}$

Suppose $r_{j-1} < \frac{d(x,y)}{3}$.
Let $\mathcal{E}_{i,j}$ be the event that:
- $B(x,<r_j) \cap A_{i,j} = \emptyset$
- $B(y, r_{j-1}) \cap A_{i,j} \neq \emptyset$

Correctness: Part 2

- $B(v, r)$ = the ball of radius r around v.
- $B(v, < r)$ = the open ball of radius r around v.

- Each point in X joins $A_{i,j}$ with probability 2^{-j} independently.
- Set $f_{i,j}(x) = d(x, A_{i,j}) = \min_{y \in A_{i,j}} d(x, y)$.

For $n \geq 2$, we have: $\left(1 - \frac{1}{n}\right)^n \geq \left(1 - \frac{1}{2}\right)^2 = \frac{1}{4}$

Observation 1:
$|B(x, < r_j)| < 2^j$ or $|B(y, < r_j)| < 2^j$

Without loss of generality, we assume this.

Observation 2:
If $r_{j-1} < \frac{d(x,y)}{3}$, then
$|B(x, r_{j-1})| \geq 2^{j-1}$ and $|B(y, r_{j-1})| \geq 2^{j-1}$

Suppose $r_{j-1} < \frac{d(x,y)}{3}$.
Let $\mathcal{E}_{i,j}$ be the event that:

- $B(x, < r_j) \cap A_{i,j} = \emptyset$
  - Probability $= \left(1 - \frac{1}{2^j}\right)^{|B(x, < r_j)|} > \left(1 - \frac{1}{2^j}\right)^{2^j} \geq \frac{1}{4}$
- $B(y, r_{j-1}) \cap A_{i,j} \neq \emptyset$
  - Probability $= 1 - \left(1 - \frac{1}{2^j}\right)^{|B(y, r_{j-1})|}$

$\geq 1 - \left(1 - \frac{1}{2^j}\right)^{2^{j-1}} \geq 1 - e^{-1/2} \geq \frac{1}{3}$

Claim 1: $\Pr[\mathcal{E}_{i,j}] > \frac{1}{12}$

$1 + x \leq e^x$

Correctness: Part 2

- Each point in X joins $A_{i,j}$ with probability 2^{-j} independently.
- Set $f_{i,j}(x) = d(x, A_{i,j}) = \min_{y \in A_{i,j}} d(x,y)$.

- $B(v,r)$ = the ball of radius r around v.
- $B(v,< r)$ = the open ball of radius r around v.

Observation 1:
$|B(x,< r_j)| < 2^j$ or $|B(y,< r_j)| < 2^j$

Without loss of generality, we assume this.

Observation 2:
If $r_{j-1} < \frac{d(x,y)}{3}$, then
$|B(x,r_{j-1})| \geq 2^{j-1}$ and $|B(y,r_{j-1})| \geq 2^{j-1}$

Suppose $r_{j-1} < \frac{d(x,y)}{3}$.
Let $\mathcal{E}_{i,j}$ be the event that:
- $B(x,< r_j) \cap A_{i,j} = \emptyset$
- $B(y,r_{j-1}) \cap A_{i,j} \neq \emptyset$

Claim 1: $\Pr[\mathcal{E}_{i,j}] > \frac{1}{12}$
Claim 2: If $\mathcal{E}_{i,j}$ occurs, then:
- $|f_{i,j}(x) - f_{i,j}(y)| \geq r_j - r_{j-1}$

Correctness: Part 2

- Consider the following setting.
  - $X_i \in \{0,1\}$ is the indicator random variable for the event $\mathcal{E}_{i,j}$.
  - $X = \sum_{i=1}^{c \log n} X_i$.
  - $\mu = \mathbb{E}[X] > \frac{c \log n}{12}$.

Claim 1: $\Pr[\mathcal{E}_{i,j}] > \frac{1}{12}$
Claim 2: If $\mathcal{E}_{i,j}$ occurs, then:
- $|f_{i,j}(x) - f_{i,j}(y)| \geq r_j - r_{j-1}$

Correctness: Part 2

- Consider the following setting.
  - $X_i \in \{0,1\}$ is the indicator random variable for the event $\mathcal{E}_{i,j}$.
  - $X = \sum_{i=1}^{c \log n} X_i$.
  - $\mu = \mathbb{E}[X] > \frac{c \log n}{12}$.

$$\Pr\left[X \leq \frac{c \log n}{24}\right] \leq \Pr\left[X \leq \frac{\mu}{2}\right] \leq e^{-\frac{\left(\frac{1}{2}\right)^2 \mu}{2}} \in n^{-\Omega(c)}$$

Chernoff bound:

- $\Pr[X \leq (1-\delta)\mu] \leq e^{-\frac{\delta^2 \mu}{2}}$, for $1 \geq \delta \geq 0$.

Claim 1: $\Pr[\mathcal{E}_{i,j}] > \frac{1}{12}$

Claim 2: If $\mathcal{E}_{i,j}$ occurs, then:

- $|f_{i,j}(x) - f_{i,j}(y)| \geq r_j - r_{j-1}$

Correctness: Part 2

- Consider the following setting.
  - $X_i \in \{0,1\}$ is the indicator random variable for the event $\mathcal{E}_{i,j}$.
  - $X = \sum_{i=1}^{c \log n} X_i$.
  - $\mu = \mathbb{E}[X] > \frac{c \log n}{12}$.

$$\Pr\left[X \le \frac{c \log n}{24}\right] \le \Pr\left[X \le \frac{\mu}{2}\right] \le e^{-\frac{\left(\frac{1}{2}\right)^2 \mu}{2}} \in n^{-\Omega(c)}$$

With probability $1 - n^{-\Omega(c)}$, we have:

$$\sum_i \left|f_{i,j}(x) - f_{i,j}(y)\right| \ge \left(r_j - r_{j-1}\right) \cdot \frac{c \log n}{24}$$

Chernoff bound:

- $\Pr[X \le (1-\delta)\mu] \le e^{-\frac{\delta^2 \mu}{2}}$, for $1 \ge \delta \ge 0$.

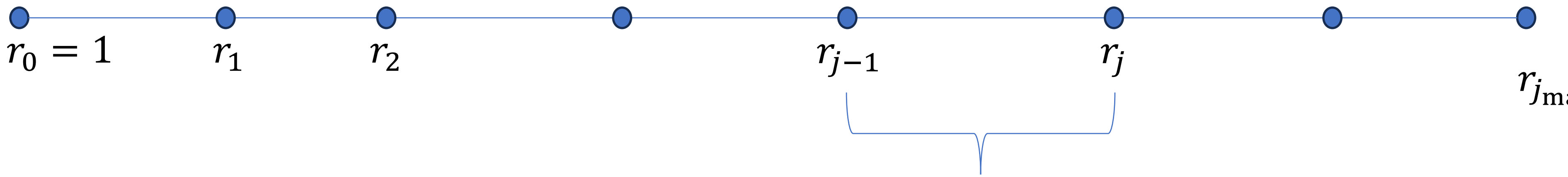
Claim 1: $\Pr\left[\mathcal{E}_{i,j}\right] > \frac{1}{12}$

Claim 2: If $\mathcal{E}_{i,j}$ occurs, then:

- $\left|f_{i,j}(x) - f_{i,j}(y)\right| \ge r_j - r_{j-1}$

Correctness: Part 2

$j_{\max} \leq \log n$

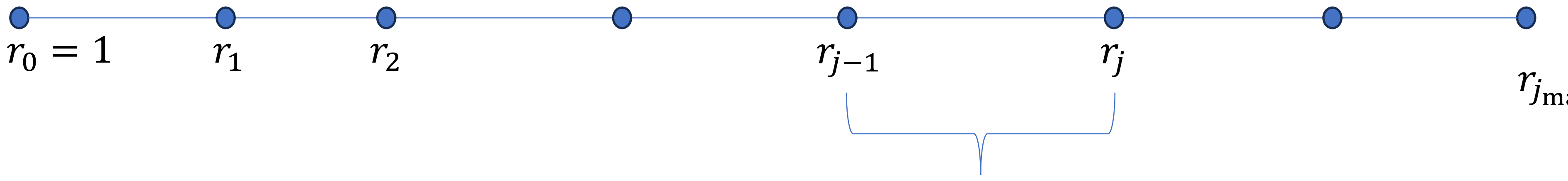


$r_{j_{\max}} = \frac{d(x,y)}{3}$

With probability $1 - n^{-\Omega(c)}$, we have: $\sum_i \left| f_{i,j}(x) - f_{i,j}(y) \right| \geq \left(r_j - r_{j-1} \right) \cdot \frac{c \log n}{24}$

Correctness: Part 2

$j_{\max} \leq \log n$



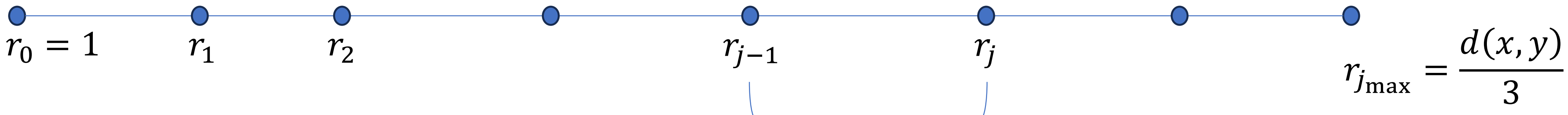
$r_{j_{\max}} = \dfrac{d(x,y)}{3}$

With probability $1 - n^{-\Omega(c)}$, we have: $\sum_i \left| f_{i,j}(x) - f_{i,j}(y) \right| \geq \left(r_j - r_{j-1} \right) \cdot \frac{c \log n}{24}$

With probability $1 - r_{j_{\max}} \cdot n^{-\Omega(c)}$, we have: $\sum_{i,j} \left| f_{i,j}(x) - f_{i,j}(y) \right| \geq \frac{d(x,y)}{3} \cdot \frac{c \log n}{24}$

Correctness: Part 2

$j_{\max} \leq \log n$



With probability $1 - n^{-\Omega(c)}$, we have: $\sum_i \left| f_{i,j}(x) - f_{i,j}(y) \right| \geq (r_j - r_{j-1}) \cdot \frac{c \log n}{24}$

With probability $1 - r_{j_{\max}} \cdot n^{-\Omega(c)}$, we have: $\sum_{i,j} \left| f_{i,j}(x) - f_{i,j}(y) \right| \geq \frac{d(x,y)}{3} \cdot \frac{c \log n}{24}$

With probability $1 - \binom{n}{2} \cdot r_{j_{\max}} \cdot n^{-\Omega(c)}$, we have: $\sum_{i,j} \left| f_{i,j}(x) - f_{i,j}(y) \right| \geq \frac{d(x,y)}{3} \cdot \frac{c \log n}{24}$ for all x and y.

Can make this $1 - \frac{1}{\mathrm{poly}(n)}$ with an arbitrarily large exponent by selecting c to be a sufficiently large constant.

This finishes the proof of correctness.

Summary

Main tool: Chernoff bound.

- $O(\log^2 n)$ subsets are needed because:
  - Need to try all $O(\log n)$ sampling probabilities: $1, \frac{1}{2}, \frac{1}{4}, \dots, \frac{1}{n}$.
  - For each of them, need to repeat for $O(\log n)$ times to get a "with high probability" guarantee.

Summary

Main tool: Chernoff bound.

- $O(\log^2 n)$ subsets are needed because:
  - Need to try all $O(\log n)$ sampling probabilities: $1, \frac{1}{2}, \frac{1}{4}, \ldots, \frac{1}{n}$.
  - For each of them, need to repeat for $O(\log n)$ times to get a "with high probability" guarantee.

If we are happy with an "in expectation" guarantee, then this is not needed!

$$\mathbb{E}[\|f(x) - f(y)\|_1] \leq d(x, y) \in O(\log n) \cdot \mathbb{E}[\|f(x) - f(y)\|_1]$$

Outlook

- We have finished the proof for $O(\log n)$-approximation of conductance.

This allows us to obtain a cut of conductance $O(\Phi(G) \cdot \log n)$ in polynomial time.

- **Next:** A completely different approach to conductance approximation via <u>spectral graph theory</u>.

This allows us to obtain a cut of conductance $O\left(\sqrt{\Phi(G)}\right)$ in polynomial time.

This is better whenever $\Phi(G) \in \omega(1/\log^2 n)$.

References

- **Main reference:**
    - Lecture 2.2 of https://sites.google.com/site/thsaranurak/teaching/Expander
    - Chapter 11 of https://lucatrevisan.github.io/books/expanders-2016.pdf