CS 6815 Pseudorandomness and Combinatorial Constructions **Fall 2018**

Lecture 15: October 18

Instructor: Eshan Chattopadhyay *Scribe: Wei-Kai Lin (wl572), Jyun-Jie Liao (jl3825)*

In this lecture, we will prove that a list-decodable code is also a strong seeded extractor:

Theorem 1. *Let $C : [N] \to [M]^D$ be a $(1 - \frac{1}{M} - \epsilon, L)$ list decodable code. Then $Ext : [N] \times [D] \to [M]$ defined by*

$$Ext(x, y) = C(x)|_y$$

is a strong seeded extractor for min-entropy $k = \log L + \log(1/\epsilon)$ with error $2M\epsilon$.

We have seen that a (k, ϵ) strong seeded extractor $Ext : [N] \times [D] \to [M]$ can be represented by a left-D-regular bipartite graph $Ext = ([N], [M] \times [D])$ such that $x \in [N]$ in the left vertex set is connected to $(y, Ext(x, y))$ in the right vertex set for every edge label $y \in [D]$. For a list decodable code $C : [N] \to [M]^D$, we can similarly define a left-D-regular bipartite graph $C = ([N], [M] \times [D])$ such that a vertex $x \in [N]$ is connected to $(y, C(x)|_y)$ for every $y \in [D]$. We define the following notation for left-regular bipartite graph:

Definition 1. For left-D-regular bipartite graph $G = (L, R)$, $T \subseteq R$ and parameter $\delta \in [0, 1]$, define

$$\mathrm{LIST}_G(T, \delta) = \{x \in L \mid |\Gamma(x) \cap T| \geq \delta D\}.$$

By definition of list-decodable code, we have the following lemma:

Lemma 1. *Let $C : [N] \to [M]^D$ be a $(1 - \delta, L)$ list-decodable code, and $T = \{(y, z_y) \mid y \in [D]\}$ for any $(z_1, z_2, \ldots, z_D) \in [M]^D$. Then $|LIST_C(T, \delta) \leq L|$.*

Now we are ready to prove the theorem.

Proof of Theorem 1. Consider the bipartite graph $C = ([N], [M] \times [D])$. Let X be a subset of $[N]$ of size K. (K will be specified later.) Observe that for the k-source U_X uniformly distributed over X, uniformly random $Y \in [D]$ and every $y \in [D], z \in [M]$,

$$\Pr\left[(Y, C(U_X)|_Y) = (y, z)\right] = \frac{|\Gamma((y, z)) \cap X|}{KD}.$$

Then the statistical distance between $(Y, C(U_X)|_Y)$ and uniform distribution is

$$\sum_{y \in [D], z \in [M]} \max\left(\frac{|\Gamma((y, z)) \cap X|}{KD} - \frac{1}{MD}, 0\right).$$

For every $y \in [D]$, define $z_y = \arg\max_{z \in [M]} (|\Gamma((y, z)) \cap X|)$. Note that $|\Gamma((y, z_y)) \cap X| \geq K/M$ by averaging. Let $T = \{(y, z_y) \mid y \in [D]\}$ and $\delta = 1/M + \epsilon$. Then

$$\begin{aligned}
\sum_{y \in [D], z \in [M]} \max\left(\frac{|\Gamma((y, z)) \cap X|}{KD} - \frac{1}{MD}, 0\right) &\leq \sum_{y \in [D]} \left(\frac{M \cdot |\Gamma((y, z_y)) \cap X|}{KD} - \frac{1}{D}\right) \\
&\leq \frac{M \cdot (|\mathrm{LIST}_C(T, \delta)| \cdot D + K \cdot \delta D)}{KD} - 1 \\
&= M\epsilon + \frac{ML}{K}
\end{aligned}$$

Choose $K = L/\epsilon$ we can conclude that C is a $(\log L + \log(1/\epsilon), 2M\epsilon)$ extractor. $\square$