CS 6815: Pseudorandomness and Combinatorial Constructions **Fall 2022**

Lecture 15: October 13, 2022

Lecturer: Eshan Chattopadhyay *Scribe: Kevin Yu*

Randomness extractors have many applications. Usually sources of randomness in nature are defective. We need to somehow produce an i.i.d. uniform sample of bits from this source.

1 Weak Sources

The following are models of weak sources. Here weak and defective are synonyms. Both are used in literature.

Definition 1.1. *The min-entropy of a distribution X is*

$$H_\infty(X) = \min_{x \in sup(x)} \log \frac{1}{\Pr[X = x]} \tag{1}$$

This is equivalent to the following statement. If $H_\infty(x) \geq k$, $\forall x,\ \Pr[X = x] \leq 2^-k$.

Definition 1.2. *A (n, k)-Source is a distribution on $\{0,1\}^n$ with min-entropy $\geq k$*

Observe that $k \leq n$ since $\sum_{x \in \{0,1\}^n} p_x = 1$. Thus, you must have at least one probability that is greater than or equal to 2^{-n} (equal case is when you have a uniform distribution).

Example 1.3. *A distribution X where x is 0^n with probability 1 has min-entropy $H_\infty(x) = 0$*

Example 1.4. *A distribution X that is uniform on $\{0,1\}^n$ has min-entropy $H_\infty(x) = n$*

Definition 1.5. *The Shannon Entropy of a distribution X is*

$$H(X) = \mathbb{E}\left[\log(\frac{1}{\Pr[X = x]})\right] = -\sum_{x \in sup(x)} \Pr(X = x) \log(\Pr[X = x]) \tag{2}$$

Example 1.6. *Consider the following distribution X*

$$x = 0^n \quad w.p.\ \frac{1}{2}$$
$$x = U_{\{0,1\}^n} \quad w.p.\ \frac{1}{2}$$

The min-entropy of X is $H_\infty(x) = 1 - o(1)$. The Shannon entropy of X is $H(x) \approx \frac{n}{2}$

2 Randomness Extractors

Definition 2.1. *A function, $Ext : \{0,1\}^n \to \{0,1\}^n$ is an extractor for a family of distributions χ (on $\{0,1\}^n$) if $\forall X \in \chi$, $Ext(x)$ is almost uniform. More formally,*

$$\forall X \in \chi,\ |Ext(X) - U_m| \leq \epsilon \tag{3}$$

Here it is better to use min-entropy to describe the weak sources because we care about how bad the worst case is. For example recall example 1.6. If we use Shannon entropy, the result does not really indicate the fact that half the time our extractor will produce a the same result when 0^n is inputted.

Remark 2.2. *A dream extractor, $Ext : \{0,1\}^n \to \{0,1\}$ that can work for all (n, k) sources and a sufficiently large k cannot exist, even for $k = n - 1$*

Proof. Suppose there exists such an extractor $Ext : \{0,1\}^n \to \{0,1\}$

Let $S_b = \{x \in \{0,1\}^n : Ext(x) = b\}$.

$$|S_0| + |S_1| = 2^n \tag{4}$$

So there exists b such that $|S_b| \geq 2^{n-1}$. Let distribution X_b be flat on S_b. Clearly, $H_\infty(x_b) \geq n - 1$ but $Ext(x_b) = \{b\}$. □

Definition 2.3. *The statistical distance between two distributions X_1 and X_2 on Ω is*

$$|X_1 - X_2| = \frac{1}{2}||X_1 - X_2|| = \frac{1}{2}\sum_{x\in\Omega} |\Pr[X_1 = x] - \Pr[X_2 = x]| \tag{5}$$

Observation 2.4. *Let $f_s : \Omega \to \{0,1\}$ be an indicator function such that $f_s(x) = 1$ iff $x \in S$. We claim the following.*

$$\begin{aligned} |X_1 - X_2| &= \max_{S\subseteq\Omega} |\Pr[f_s(X_1) = 1] - \Pr[f_s(X_2) = 1]| && (6)\\ &= \max_{S\subseteq\Omega} |\Pr[X_1 \in S] - \Pr[X_2 \in S]| && (7) \end{aligned}$$

Definition 2.5. *X_1 is ϵ-close to X_2 iff*

$$|\Pr[f(X_1) = 1] - \Pr[f(X_2) = 1]| < \epsilon \quad \forall f : \Omega \to \{0,1\} \tag{8}$$