CS59200-PRS: Pseudorandomness Oct 6th, 2025

Lecture 12: Randomness Extractor

Lecturer: Wei Zhan *Scribe: Xiuyu Ye*

1 Randomness Extractor

Recall the definition of ε-recycling.

Definition 1 (ε-recycling)**.** *Fix $d \in \mathbb{N}^+$. A function $H\colon \{0,1\}^\ell \times [d] \to \{0,1\}^\ell$ is ε-recycling if for every $w \in \mathbb{N}^+$ and every function $F\colon \{0,1\}^\ell \to [w]$,*

$$d_{\mathrm{TV}}\left(\left(F(s), s'\right), \left(F(s), H(s,r)\right)\right) \le w\varepsilon,$$

where $s, s' \sim \{0,1\}^\ell$ and $r \sim [d]$.

In the definition of ε-recycling, the function H recycles the randomness seed s to generate a new seed $H(s,r)$ that looks independently random, even when some information about s is known.

We have shown how to construct ε-recycling functions from ε-mixing and thereby expanders are naturally introduced. Here we take another look from a different perspective. We can rewrite

$$d_{\mathrm{TV}}\left(\left(F(s), s'\right), \left(F(s), H(s,r)\right)\right) = \sum_{v \in [w]} \Pr[F(s) = v] \cdot d_{\mathrm{TV}}\left(s', H(s,r) | F(s) = v\right),$$

where $(H(s,r)|F(s) = v)$ is the distribution of $H(s,r)$ conditioned on $F(s) = v$. Given $F(s) = v \in [w]$, we know at most $\log(w)$ bits of information about the seed $s \in \{0,1\}^\ell$, the leftover entropy of s is at least $\ell - \log(w)$. Therefore s is now an imperfect source of randomness, and we want to *extract* randomness that is close to perfect using the function H, with the help of a tiny amount of extra randomness r.

1.1 Definition of Extractors

To extract perfect randomness from an imperfect source, we introduce the following notion of randomness extractor.

Definition 2 (Attempted definition of extractor)**.** *A function $\mathrm{Ext}\colon \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ is a (k,ε)-extractor if for every distribution X over $\{0,1\}^n$ with entropy $H(X) \ge k$, we have*

$$d_{\mathrm{TV}}\left(\mathrm{Ext}\left(X, r\right), U_m\right) \le \varepsilon$$

where r is chosen randomly from $\{0,1\}^d$ and U_m is the uniform distribution over $\{0,1\}^m$.

The Shannon entropy of a distribution X is defined as

$$H(X) \;:=\; -\sum_x \Pr[X=x]\cdot\log\left(\Pr[X=x]\right).$$

Example. Consider the following distribution over $\{0,1\}^n$:

$$X \sim \begin{cases} 00\ldots0 & \text{with probability } 0.99 \\ U_n & \text{with probability } 0.01 \end{cases}.$$

The Shannon entropy of X satisfy that $H(X) \geq \Omega(n)$, but it is generally impossible to extract perfect randomness from X, since the output distribution of $H(X,r)$ will be heavily supported on the set

$$\left\{H(00\ldots0,r) \mid r\in\{0,1\}^d\right\}.$$

Therefore for any $\varepsilon < 0.49$ it is impossible to have an (k,ε)-extractor with $m > d$. In other words, the randomness of such "extractor" would come entirely from the extra randomness r, and hence is trivial and useless.

The reason behind the above example is that when using Shannon entropy as measurement, a small fraction of random instances contribute significantly to the total entropy. To avoid this, we use a new notion of entropy, min-entropy.

Definition 3 (Min-entropy)**.** *For a distribution X, the min-entropy of X is defined as*

$$H_\infty(X) = -\log\left(\max_x \Pr[X=x]\right).$$

Definition 4 ((k,ε)-Extractor)**.** *A function* $\mathrm{Ext}\colon \{0,1\}^n\times\{0,1\}^d \to \{0,1\}^m$ *is a (k,ε)-extractor if for every distribution X over $\{0,1\}^n$ with* **min-entropy** *$H_\infty(X)\geq k$, we have*

$$d_{\mathrm{TV}}\left(\mathrm{Ext}\left(X,r\right),U_m\right)\leq\varepsilon,$$

where r is chosen randomly from $\{0,1\}^d$ and U_m is the uniform distribution over $\{0,1\}^m$.

Using min-entropy in the definition of extractors has several benefits. First, min-entropy is the smallest entropy measure among the Rényi p-entropy family, defined as

$$H_p(X) = -\log\left(\sum_x \Pr[X=x]^p\right)^{\frac{1}{p-1}},$$

which has a monotonic trend

$$H_\infty(X)\leq\ldots\leq H_p(X)\leq\ldots\leq H_1(X)=H(X)\leq H_0(X).$$

Therefore a lower bound on the min-entropy $H_\infty(X)\geq k$ is the strongest assumption on the distribution, as it implies that all Rényi p-entropies are at least k.

The second benefit is that min-entropy has good geometric properties. Note that the condition that $H_\infty(X) \geq k$ is equivalent to that $\Pr[X = x] \leq 2^{-k}$ for all x, which is a linear constrain. Thus all distributions satisfying $H_\infty(X) \geq k$ form a convex polytope, in contrast to the case of Shannon entropy where the set of distributions with $H(X) \geq k$ is a convex set with smooth surface that is hard to describe. Even better, we can actually identify the vertices of the polytope.

Definition 5 (flat k-source)**.** *We call a distribution X with min-entropy $H_\infty(X) \geq k$ a k-source. A flat k-source is a uniform distribution over a support of size 2^k.*

Theorem 1. *For a distribution X, the min-entropy $H_\infty(X) \geq k$ if and only if X is a convex combination of flat k-source.*

Proof. Consider the linear program over $P_x = \Pr[X = x]$ with an arbitrary objective, where the feasible solutions corresponds to the distributions X with $H_\infty(X) \geq k$:

$$\begin{aligned} \text{maximize } & \sum_x c_x P_x \\ \text{s.t. } & \sum_x P_x = 1, \\ & 0 \leq P_x \leq 2^{-k}, \ \forall x. \end{aligned}$$

The maximum is clearly taken when $P_x = 2^{-k}$ for the largest 2^k coefficients c_x. □

1.2 Extractor implies recycling

Back to our initial intuition, we show below that randomness extractors indeed implies recycling, and we can prove a more general statement even when the output length m is not the same as the input length n:

Theorem 2. *If a function $H\colon \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ is a (k,ε)-extractor, then for every $w \in \mathbb{N}^+$ and every function $F\colon \{0,1\}^n \to [w]$,*

$$d_{\mathrm{TV}}\left(\left(F(s), s'\right), \left(F(s), H(s,r)\right)\right) \leq \varepsilon + w \cdot 2^{k-n},$$

where $s \sim \{0,1\}^n$, $s' \sim \{0,1\}^m$ and $r \sim \{0,1\}^d$.

Proof. Consider an arbitrary function $F\colon \{0,1\}^n \to [w]$. If $\Pr[F(s) = v] \geq 2^{k-n}$ where s is chosen randomly from $\{0,1\}^n$, then for every $x \in \{0,1\}^n$,

$$\Pr[s = x | F(s) = v] = \begin{cases} 0 & \text{if } F(x) \neq v \\ \frac{\Pr[s=x]}{\Pr[F(s)=v]} & \text{if } F(x) = v. \end{cases}$$

Therefore, $\Pr[s = x|F(s) = v] \leq 2^{-k}$ and the min-entropy of distribution $(s|F(s) = v)$ satisfies $H_\infty(s|F(s) = v) \geq k$.

Now for $s \sim \{0,1\}^n$, $s' \sim \{0,1\}^m$ and $r \sim \{0,1\}^d$, we have

$$\begin{aligned}
& d_{\mathrm{TV}}\left(\left(F(s), s'\right), (F(s), H(s,r))\right) \\
= & \sum_{v \in [w]} \Pr[F(s) = v] \cdot d_{\mathrm{TV}}\left(s', H(s,r) | F(s) = v\right) \\
= & \sum_{v \in [w]:\ \Pr[F(s)=v] \geq 2^{k-n}} \Pr[F(s) = v] \cdot d_{\mathrm{TV}}\left(s', H(s,r) | F(s) = v\right) \\
& + \sum_{v \in [w]:\ \Pr[F(s)=v] < 2^{k-n}} \Pr[F(s) = v] \cdot d_{\mathrm{TV}}\left(s', H(s,r) | F(s) = v\right) \\
< & \sum_{v \in [w]:\ \Pr[F(s)=v] \geq 2^{k-n}} \Pr[F(s) = v] \cdot d_{\mathrm{TV}}\left(s', H(s,r) | F(s) = v\right) + w \cdot 2^{k-n} && (d_{\mathrm{TV}} \leq 1) \\
\leq & \varepsilon + w \cdot 2^{k-n} && \text{(Definition 4)}
\end{aligned}$$

□

2 Construction of Extractors

2.1 Random Construction

Let $\mathrm{Ext}\colon \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ be a random function, and we will show that Ext is a randomness extractor with high probability, for some properly chosen parameters.

For every distinguisher $A\colon \{0,1\}^m \to \{0,1\}$ and every flat k-source X, the random variables $A(\mathrm{Ext}(x,r))$ indexed by $x \in \operatorname{supp} X$ and $r \in \{0,1\}^d$ are independent random variables with values in $\{0,1\}$, and there are exactly $2^k \cdot 2^d$ of them. Therefore by Chernoff bound,

$$\Pr_{\mathrm{Ext}}\left[\left|\mathop{\mathbb{E}}_{x \sim X, r \sim \{0,1\}^d}[A(\mathrm{Ext}(x,r))] - \mathop{\mathbb{E}}_{u \sim U_m}[A(u)]\right| \geq \varepsilon\right] \leq 2e^{-\frac{1}{3} \cdot 2^{k+d} \varepsilon^2}.$$

Taking union bound over all distinguishers and all flat k-sources, we get

$$\begin{aligned}
\Pr_{\mathrm{Ext}}[\text{Ext is not a } (k,\varepsilon)\text{-extractor}] \leq & 2^{2^m} \cdot \binom{2^n}{2^k} \cdot 2e^{-\frac{1}{3} \cdot 2^{k+d} \varepsilon^2} \\
\leq & 2^{2^m} \cdot \left(\frac{2^n \cdot e}{2^k}\right)^{2^k} \cdot 2e^{-\frac{1}{3} \cdot 2^{k+d} \varepsilon^2} && \text{(Stirling)} \\
= & \exp\left(\ln 2 \cdot (2^m + 1 + (n-k) \cdot 2^k) + 2^k - \frac{1}{3} \cdot 2^{k+d} \cdot \varepsilon^2\right).
\end{aligned}$$

If we let the above probability to be smaller than 1, we can conclude that there exists a (k,ε)-extractor $\mathrm{Ext}\colon \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ with the following output and seed length

$$\begin{cases} m = k + d - 2\log(1/\varepsilon) - O(1) \\ d = \log(n-k) + 2\log(1/\varepsilon) + O(1) \end{cases}$$

and these are the parameters that we will be shooting for, when we construct explicit extractors later on.

2.2 Construction from Hash Functions

We consider the following construction using pairwise independent hash function.

Theorem 3 (Leftover hash lemma)**.** *If $\mathcal{H} = \{h\colon \{0,1\}^n \to \{0,1\}^m\}$ is a pairwise uniform hash function family, then* $\operatorname{Ext}\colon \{0,1\}^n \times \mathcal{H} \to \{0,1\}^m \times \mathcal{H}$ *where* $\operatorname{Ext}(x,h) = (h(x),h)$ *is a (k,ε)-extractor, where $\varepsilon = 2^{(m-k)/2-1}$.*

Proof. Let X be a flat k-source. For $x \sim X$ and $h \sim \mathcal{H}$ we have

$$
\begin{aligned}
&d_{\mathrm{TV}}\left(\operatorname{Ext}(x,h),(U_m,h)\right)\\
&= \mathop{\mathbb{E}}_{h\sim\mathcal{H}}\left[d_{\mathrm{TV}}(h(x),U_m)\right]\\
&= \mathop{\mathbb{E}}_{h\sim\mathcal{H}}\left[\frac{1}{2}\sum_{y\in\{0,1\}^m}\left|\Pr_{x\sim X}[h(x)=y]-\frac{1}{2^m}\right|\right]\\
&= \frac{1}{2}\sum_{y\in\{0,1\}^m}\mathop{\mathbb{E}}_{h\sim\mathcal{H}}\left[\left|\Pr_{x\sim X}[h(x)=y]-\frac{1}{2^m}\right|\right]\\
&\le \frac{1}{2}\sum_{y\in\{0,1\}^m}\mathop{\mathbb{E}}_{h\sim\mathcal{H}}\left[\left(\Pr_{x\sim X}[h(x)=y]-\frac{1}{2^m}\right)^2\right]^{\frac{1}{2}} && \text{(Jenssen's inequality on } f(x)=x^2)\\
&= \frac{1}{2}\sum_{y\in\{0,1\}^m}\mathop{\mathbb{E}}_{h\sim\mathcal{H}}\left[\Pr_{x_1,x_2\sim X}[h(x_1)=h(x_2)=y]-\frac{2}{2^m}\Pr_{x\sim X}[h(x)=y]+\frac{1}{2^{2m}}\right]^{\frac{1}{2}}\\
&\le \frac{1}{2}\sum_{y\in\{0,1\}^m}\mathop{\mathbb{E}}_{h\sim\mathcal{H}}\left[\Pr_{x_1,x_2\sim X}[x_1=x_2,h(x_1)=y]+\frac{1}{2^{2m}}-\frac{2}{2^m}\cdot\frac{1}{2^m}+\frac{1}{2^m}\right]^{\frac{1}{2}}\\
&\qquad\qquad \left(\Pr_{x_1,x_2\sim X}[h(x_1)=h(x_2)=y\,|\,x_1\neq x_2]=\tfrac{1}{2^{2m}} \text{ with pairwise independence}\right)\\
&= \frac{1}{2}\sum_{y\in\{0,1\}^m}\mathop{\mathbb{E}}_{h\sim\mathcal{H}}\left[\Pr_{x_1,x_2\sim X}[x_1=x_2]\cdot\Pr_{x_1\sim X}[h(x_1)=y]\right]^{\frac{1}{2}}\\
&\le \frac{1}{2}\sum_{y\in\{0,1\}^m}\sqrt{\frac{1}{2^k}\cdot\frac{1}{2^m}} && (|\operatorname{supp} X| = 2^k)\\
&= \frac{1}{2}\cdot 2^{(m-k)/2}. && \square
\end{aligned}
$$

If we let $d = \log|\mathcal{H}|$ to be the seed length, then the above extractor has output length $m + d = k + d - 2\log(1/\varepsilon) + 2$, which is optimal compared to the random construction. However, the seed length $d = O(n)$ is exponentially worse than optimal.