Lecture 9: Randomness Extractors

Lecturer: He Sun

1 Definitions

Extractors are functions which can extract random bits from any distribution which contains sufficient randomness and have played a unifying role in the theory of pseudorandomness.

Examples of k-sources:

- k random and independent bits, together with $n-k$ fixed bits (in an arbitrary order). They are called oblivious bit-fixing sources.
- k random and independent bits, and $n-k$ bits that depend arbitrarily on the first k bits. They are called adaptive bit-fixing sources.
- Uniform distribution on $S \subseteq \{0,1\}^n$ with $|S| = 2^k$. These are called flat k-sources.

Proposition 9.1 *Every k-source is a convex combination of flat k-sources (provided that $2^k \in \mathbb{N}$), i. e. $X = \sum_i p_i X_i$ with $0 \le p_i \le 1, \sum p_i = 1$ and all the X_i are flat k-sources.*

Definition 9.2 (deterministic extractors) *Let $\mathcal{C}$ be a class of sources on $\{0,1\}^n$. An ε-extractor for $\mathcal{C}$ is a function $\mathrm{Ext} : \{0,1\}^n \to \{0,1\}^m$ such that for every $X \in \mathcal{C}$, $\mathrm{Ext}(X)$ is "ε-close" to $\mathcal{U}_m$.*

Lemma 9.3 *Let $A(w;r)$ be a randomized algorithm such that $A(w;U_m)$ has error probability at most γ, and let $\mathrm{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ be a (k,ε)-extractor for a class $\mathcal{C}$ of sources on $\{0,1\}^n$. Define $A'(w,r) = A(w, \mathrm{Ext}(x))$. Then for every source $X \in \mathcal{C}$, $A'(w;X)$ has error probability at most $\gamma + \varepsilon$.*

Proposition 9.4 *For any $\mathrm{Ext} : \{0,1\}^n \to \{0,1\}$ there exists an $(n-1)$-source X such that $\mathrm{Ext}(X)$ is constant.*

Proof: Without loss of generality, we assume that $|\mathrm{Ext}^{-1}(0)| \ge |\mathrm{Ext}^{-1}(1)|$. Then $|\mathrm{Ext}^{-1}(0)| \ge 2^{n-1}$. Let X be the uniform distribution on $\mathrm{Ext}^{-1}(0)$. ■

So we require a weaker concept of extractors. That is, we hope that, by using a small number of random bits which is an additional input, the output of extractors is ε-close to uniform distributions.

Definition 9.5 (seeded extractors) *A (k,ε)-extractor is a function $\mathrm{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ such that for every distribution X on $\{0,1\}^n$ with $\mathbf{H}_\infty(X) \ge k$ the distribution $\mathrm{Ext}(X, \mathcal{U}_d)$ is ε-close to $\mathcal{U}_m$.*

Every extractor has five different parameters:

- The length of the source n.
- The output length m.
- The length of seeds d.
- The min-entropy threshold k.
- The error of the extractor ε.

We refer to the ratio k/n as *the entropy rate* of the source X and to the ratio m/k as *the fraction of randomness* extracted by Ext. The goal of constructing good extractors is to minimize d and maximize m.

Applications of extractors include:

- Simulating randomized algorithms using weak random sources.
- Random sampling using few random bits.
- Expanders that beat the eigenvalue bound.
- Explicit constructions of error correcting codes.

As an example, let us see the applications of extractors for simulating randomized algorithms. Assume that algorithm A uses m random bits. Since we do not know how to obtain truly random bits, algorithm A uses the "almost random" strings to instead to purely random ones. That is, the random strings for A come from the output of $\mathrm{Ext}(X, \mathcal{U}_d)$. The randomness used for seeds can be eliminated by running all the possible seeds and taking the majority value. In particular, an explicit extractor with logarithmic seed length can be used to simulate BPP given access to a weak random source of sufficient min-entropy.

Lemma 9.6 *Let $A(w;r)$ be a randomized algorithm such that $A(w;\mathcal{U}_m)$ has error probability at most γ, and let $\mathrm{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ be a (k,ε)-extractor. Define*

$$A' = \mathrm{maj}_{y \in \{0,1\}^d} \left\{ A(w, \mathrm{Ext}(x,y)) \right\}.$$

Then for every k-source X on $\{0,1\}^n$, $A'(x;X)$ has error probability at most $2(\gamma+\varepsilon)$.

2 Extractors as Hash Functions

Throughout the note, capital variables are 2 raised to the power of the corresponding lower variable, e. g. $D = 2^d$.

Definition 9.7 (pairwise independent hash functions) *A family of pairwise hash functions is a set of functions $h : D \to R$ such that for any distinct $x_1, x_2 \in D$ and all (not necessarily distinct) $y_1, y_2 \in R$, it holds that*

$$\Pr[h(x_1) = y_1 \wedge h(x_2) = y_2] = \frac{1}{|R|^2}.$$

Theorem 9.8 (Leftover Hash Lemma) *If $\mathcal{H} = \{h : \{0,1\}^n \to \{0,1\}^m\}$ is a family of pairwise independent hash functions where $m = k - 2\log(1/\varepsilon)$. Then $\mathrm{Ext}(x, h) = (h, h(x))$ is a (k, ε)-extractor.*

Proof: Let X be an arbitrary k-source on $\{0,1\}^n$ and d be the seed length of Ext. Choose H randomly from $\mathcal{H}$. We show that $(X, H(X))$ is ε-close to $\mathcal{U}_d \times \mathcal{U}_m$ in the following three steps.

Step 1: By definition $\mathrm{CP}(H, H(X)) = \Pr[(H, H(X)) = (H', H'(X'))]$, where $(H', H'(X'))$ is independent of and identically distributed to $(H, H(X))$. Because $(H, H(X)) = (H', H'(X'))$ if and only if $H = H'$ and either $X = X'$ or $X \neq X'$ but $H(X) = H(X')$. Therefore

$$\begin{aligned}\mathrm{CP}(H, H(X)) &= \mathrm{CP}(H) \cdot \big(\mathrm{CP}(X) + \Pr[H(X) = H(X') | X \neq X']\big)\\ &\leq \frac{1}{D} \cdot \left(\frac{1}{K} + \frac{1}{M}\right)\\ &= \frac{1+\varepsilon^2}{DM},\end{aligned}$$

where the last equality uses the fact that $m = k - 2\log(1/\varepsilon)$.

Step 2:

$$\begin{aligned}\|(H, H(X)) - \mathcal{U}_d \times \mathcal{U}_m\|^2 &= \mathrm{CP}(H, H(X)) - \mathrm{CP}(\mathcal{U}_d \times \mathcal{U}_m)\\ &\leq \frac{1+\varepsilon^2}{DM} - \frac{1}{DM} = \frac{\varepsilon^2}{DM}\end{aligned}$$

Step 3:

$$\begin{aligned}\Delta((H, H(X)), \mathcal{U}_d \times \mathcal{U}_m) &= \frac{1}{2} \cdot \|(H, H(X)) - \mathcal{U}_d \times \mathcal{U}_m\|_1\\ &\leq \frac{\sqrt{DM}}{2} \cdot \|(H, H(X)) - \mathcal{U}_d \times \mathcal{U}_m\|\\ &\leq \frac{\sqrt{DM}}{2} \cdot \sqrt{\frac{\varepsilon^2}{DM}}\\ &= \frac{\varepsilon}{2}.\end{aligned}$$

Therefore $\mathrm{Ext}(x, h(x))$ is a (k, ε)-extractor. ■

3 Extractors v.s. Expanders

Extractors $\mathrm{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ can be thought as bipartite graphs $G = ([N], [M], E)$ where the nodes of the left side are strings of length n and the nodes of the right side are strings of length m. Every node x on the left side is connected to all 2^d nodes z for which there exists a $y \in \{0,1\}^d$ such that $\mathrm{Ext}(x, y) = z$. By the definition of extractors, for every set $S \subseteq \{0,1\}^n$ of size 2^k on the left hand side and for every set $T \subseteq \{0,1\}^m$ on the right hand side, the number of edges between S and T is close to what one expects in a random graph. More precisely, we have

$$\left| e(S,T) - |S| \cdot |T| \cdot 2^{d-m} \right| \leq \varepsilon,$$

where $e(S,T)$ is the size of $E(S,T)$.

Comments:

- Extractors are unbalanced bipartite graphs.
- The degree of extractors is not constant.

3.1 Extractors ⇒ Expanders

Let $\mathrm{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ be an extractor. Then for any set $S \subseteq [N]$ of size K, we have

$$\Delta(\mathrm{Ext}(U_S, \mathcal{U}_d), \mathcal{U}_m) \leq \varepsilon$$

where U_S is the uniform distribution on S. Let $\mu(S) = |S|/N$ and $\mu(T) = |T|/M$. Then for any set $T \subseteq [M]$, it holds that

$$|\Pr[\mathrm{Ext}(U_S, \mathcal{U}_d) \in T] - \mu(T)| \leq \varepsilon,$$

i. e.,

$$\left| \frac{e(S,T)}{|S| \cdot D} - \mu(T) \right| \leq \varepsilon.$$

We rewrite the inequality above as

$$\left| \frac{e(S,T)}{ND} - \mu(S)\mu(T) \right| \leq \varepsilon\mu(S). \tag{1}$$

Proposition 9.9 Ext *is a (k,ε)-extractor iff the corresponding bipartite graph $G = ([N],[M],E)$ with left-degree D has the property that $\left|\frac{e(S,T)}{ND} - \mu(S)\mu(T)\right| \leq \varepsilon\mu(S)$ for every $S \subseteq [N]$ of size K and every $T \subseteq [M]$.*

3.2 Expanders ⇒ Extractors

Comparing Eq. (1) with the Expander Mixing Lemma, which states that for any graph G with spectral expansion λ and for any sets $S,T \subseteq [N]$, we have

$$\left| \frac{e(S,T)}{N \cdot D} - \mu(S)\mu(T) \right| \leq \lambda\sqrt{\mu(S)\mu(T)},$$

it suffices that G is an extractor if $\lambda \cdot \sqrt{\mu(S)\mu(T)} \leq \varepsilon\mu(S)$ for all $S \subseteq [N]$ of size K and all $T \subseteq [N]$. So it suffices for $\lambda \leq \varepsilon\sqrt{K/N}$.

For the constructions of such expanders, we take an appropriate power of a constant degree expander. Specially, let G_0 be a D_0-regular expander on N vertices with bounded spectral expansion. We take the t-th power of G_0 and let $G = G_0^t$ where $t = O(\log((1/\varepsilon)\sqrt{N/K})) = O(n - k + \log(1/\varepsilon))$.

Theorem 9.10 *For every $n,k \in \mathbb{N}$ and $\varepsilon > 0$, there is an explicit (k,ε)-extractor* $\mathrm{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^n$ *with $d = O(n - k + \log(1/\varepsilon))$.*

3.3 Comparison

Expanders	**Extractors**
Measured by vertex or spectral expansion	Measured by min-entropy/statistical difference
Typically constant degree	Typically logarithmic or poly-logarithmic degree
All sets of size at most K expand	All sets of size exactly (or at least) K expand
Typically balanced	Typically unbalanced, bipartite graphs