Randomness Extractors

October 5, 2017

Deterministic Extractors for Independent Sources. Let $\mathcal{D}$ be a family of distributions over $\{0,1\}^n$ such that every distribution $X = X_1, \ldots, X_n \in \mathcal{D}$ satisfies that $X_1, \ldots, X_n$ are independently distributed and for all $i \in [n]$, $\Pr[X_i = 0] \in [1/3, 2/3]$.

We say $f\colon \{0,1\}^n \to \{0,1\}$ is a deterministic extractor for $\mathcal{D}$ with error ε if for any $X \in \mathcal{D}$, $|\Pr[f(X) = 0] - 1/2|$ is at most ε. The following claim shows parity function is an extractor for $\mathcal{D}$ with exponentially small error in n.

Claim 1. *For any $X \in \mathcal{D}$, $|\Pr[f(X) = 0] - 1/2| \leq (1/2)(1/3)^n$ where $f(x) = x_1 \oplus x_2 \cdots \oplus x_n$.*

Proof. Here is a useful (and easy to verify) trick: for any boolean variable Z, $\Pr[Z = 0] = \mathbb{E}[\frac{1+(-1)^Z}{2}]$. Therefore it suffices to show for any $X \in \mathcal{D}$, $|\mathbb{E}[(-1)^{X_1 \oplus X_2 \cdots \oplus X_n}]| \leq (1/3)^n$. This follows from $X_1, \ldots, X_n$ are independent from each other and $|\mathbb{E}[(-1)^{X_i}]| \leq 1/3$ for any $i \in [n]$. $\square$

Remark 1 (More Bits?). *A natural idea is to divide the coordinates into chunks of size $\sqrt{n}$ and output parity bit with each chunks. This method will produce $\sqrt{n}$ bits with error $2^{-\Omega(\sqrt{n})}$. Can one extract $\Omega(n)$ bits with error $2^{-\Omega(n)}$? The answer is Yes!*

Impossibility of Deterministic Extractors for Unpredicability Souces. Consider a variant of $\mathcal{D}$ where every distribution $X \in \mathcal{D}$ (instead of the independence condition) satisfies that, for all $i \in [n]$, and $x_1, \ldots, x_{i-1} \in \{0,1\}^{i-1}$, $\Pr[X_i = 0 | X_1 = x_1, \ldots, X_{i-1} = x_{i-1}] \in [1/3, 2/3]$. Namely, we relax the indpendence condition to that each coordinate is still hard to predict conditioning on all previous outcomes.

Sources in this family can be thought of generated in the follow way: there is an adversary holding two biased coin C_1, C_2 where C_1 is 1 with probability 2/3 and C_2 is 0 with probability 1/3, and to generate the ith coodinate, the adversary goes over all previsou generated bit then pick[1] C_1 or C_2 then sample from the distribution.

Is parity still a good extractor for this source? Because conditioning on all previous $n-1$ coordiantes in the sample, the last coordinate completely determine the output and the adversary could generate the last coordinate using either D_1 or D_2, adversary can always make f outputting 0 with probability 2/3, which is no better than the error of just outputting the first coordinate. In fact, Santa and Vazirani [SV86] showed any deterministic function cannot do better.

[1] The adversary could use randomness to pick C_1 or C_2, i.e., pick a convex combination of C_1 and C_2

Theorem 1 ([SV86]). *For any $f\colon \{0,1\}^n \to \{0,1\}$, there exists an $X \in \mathcal{D}$, such that $|\Pr[f(X) = 0] - 1/2| \geq 1/6$.*

This impossibility result motivates the study of using small amount of additional randomness (called seed) to extract randomness for larger family of sources.

Remark 2 (Dice v.s. Coins). *What happend if the adversary is holding two dice instead of two coins? Can one extract randomness? The answer is Yes!*

Seeded Extractors. Now we consider functions from $\{0,1\}^n \times \{0,1\}^d$ to $\{0,1\}^m$ and let U_d denote the uniform distribution over d bits. To prevent f simplying outoutting the seed, we say f is (strong)-extractor for $\mathcal{D}$ with error ε only if for any $X \in \mathcal{D}$, the statistical distance between $(U_d, f(X))$ and (U_d, U_m) is at most ε.

Unlike previous family of distributions, now we consider the family of distributions only satisfying certain randomness requirement. We say a distribution X over $\{0,1\}^n$ has min-entropy $H_\infty(X) = k$ if k is the largest number such that for any $a \in \{0,1\}^n$, $\Pr[X = a] \leq 2^{-k}$. We say f is a (k, ε) extractor if f is an extractor for $\mathcal{D} = \{X : H_\infty(X) \geq k\}$. Here are some examples:

- Uniform distribution: $H_\infty(U_m) = m$.
- k-flat sources: Let $S \subset \{0,1\}^n$ be a set size 2^k and let X_S be the uniform distribution over S. $H_\infty(X_S) = k$.

A useful fact is that any distribution with min-entropy k is a convex combination of k-flat sources. It implies that to show f is (k, ε) extractor, it suffices to show f extracts with error at most ε for any k-flat sources.

How many bits can we hope for to extract? Radhakrishnan and Ta-Shma [RT00] showed that we can extract at most $k - 2\log(1/\varepsilon) + O(1)$ bits in the strong extractors. In other words, the error is at least $\Omega(2^{m-k})$. In the following, we show a construction matching this bound (up to constant multiplicative factors).

Leftover Hash Lemma. Leftover Hash Lemma says pair-wise independence hash is a good strong extractor.

Theorem 2. *Let H be a distribution over a family of functions $\{h : \{0,1\}^n \to \{0,1\}^m\}$ such that for any $x \neq x' \in \{0,1\}^n$ and $y, y' \in \{0,1\}^m$,*

$$\Pr_{h \sim H}[h(x) = y, h(x') = y'] = \frac{1}{2^m}.$$

For any distribution X over $\{0,1\}^n$ with $H_\infty(X) = k$, it holds that

$$\Delta((H, H(X)), (H, U_m)) \leq \frac{1}{2} \cdot \sqrt{2^{m-k}}.$$

where Δ is the statistical distance.

So in general given a distribution Y over m bits, how to bound the statistical distance between Y and U_m? Following claim shows if the collision probability of Y is small, then Y is close to the uniform distribution.

Claim 2. $\Delta(Y, U_m) \le \frac{1}{2}\sqrt{2^m \Pr_{y,y' \sim Y}[y = y']^2 - 1}$.

The proof of the claim is by Cauchy-Schwarz inequality and rewriting things. Specifically for our case, we can derive the following lemma via similar proofs.

Lemma 1. *Let H be a distribution over a family of functions $\{h : \{0,1\}^n \to \{0,1\}^m\}$. For any X with $H_\infty(X) = k$,*

$$\Delta((H, H(X)), (H, U_m)) \le \frac{1}{2}\sqrt{2^m \Pr_{h \sim H, x, x' \sim X}[h(x) = h(x')] - 1}.$$

Proof.

$$\begin{aligned}
2\Delta((H, H(X)), (H, U_m)) &= \sum_{h,b} |\Pr[H = h, H(X) = b] - \Pr[H = h, U_m = b]| \\
&= \sum_{h,b} \Pr[H = h] |\Pr[h(X) = b] - \Pr[U_m = b]| \\
&= \sum_{h,b} \sqrt{\Pr[H = h]} \cdot \sqrt{\Pr[H = h]} |\Pr[h(X) = b] - \Pr[U_m = b]| \\
&\le \sqrt{(\sum_{h,b} \Pr[H = h]) \cdot \sum_{h,b} \Pr[H = h](\Pr[h(X) = b] - \Pr[U_m = b])^2} \\
&= \sqrt{2^m \cdot \sum_{h,b} \Pr[H = h](\Pr[h(X) = b] - \Pr[U_m = b])^2}
\end{aligned}$$

where the inequality is by Cauchy-Schwarz inequality. We finish the proof by rewriting $\sum_{h,b} \Pr[H = h](\Pr[h(X) = b] - \Pr[U_m = b])^2$ as follows

$$\begin{aligned}
&\sum_{h,b} \Pr[H = h](\Pr[h(X) = b]^2 - \frac{2 \cdot \Pr[h(X) = b]}{2^m} + \frac{1}{2^{2m}}) \\
&= \sum_{h} \Pr[H = h](\sum_{b} \Pr[h(X) = b]^2 - \frac{2}{2^m} + \frac{1}{2^m}) \\
&= \Pr_{h \sim H, x, x' \sim X}[h(x) = h(x')] - \frac{1}{2^m}.
\end{aligned}$$

□

Given this lemma, it is easy to see Theorem 2. Note that for pairwise independence hash H, it holds that any $x \ne x'$, $\Pr_{h \sim H}[h(x) = h(x')] \le 1/2^m$, and for X with min-entropy at k, it holds that $\Pr_{x', x \sim X}[x' = x] \le 2^{-k}$. Therefore $\Pr_{h \sim H, x, x' \sim X}[h(x) = h(x')] \le 1/2^m + 1/2^k$.

Remark 3. $\{H_A(x) = Ax\}$ *where $A \in \mathbb{Z}_2^{n \times m}$ is construction of pair-wise independence hash functions. Can a family of spares linear transformation be good extractors? The answer is also Yes.*

References

[RT00] Jaikumar Radhakrishnan and Amnon Ta-Shma. Bounds for dispersers, extractors, and depth-two superconcentrators. *SIAM J. Discrete Math.*, 13(1):2–24, 2000.

[SV86] Miklos Santha and Umesh V. Vazirani. Generating quasi-random sequences from semi-random sources. *J. Comput. Syst. Sci.*, 33(1):75–87, 1986.