Computer Science 294 Lecture 14 Notes

Daniel Raban

March 2, 2023

1 Fooling Low-Degree Polynomials, Restriction-Based PRGs, and Fractional PRGs

1.1 Fooling polynomial functions over $\mathbb{F}_2$ with small-biased distributions

Last time, we discussed pseudorandomness for small-biased distributions. One way to think of small-biased distributions is that they fool linear functions over $\mathbb{F}_2$. You might ask if they fool quadratic functions, as well. This would be the case if $L_1(f)$ is small for quadratic functions, but this is not always the case.

Example 1.1. Let

$$\mathrm{IP}_2 = x_1x_2 + x_3x_4 + \cdots + x_{n-1}x_n \pmod 2.$$

Then for all $S \subseteq [n]$, $|\widehat{\mathrm{IP}}(S)| = 2^{-n/2}$, so $L_1(\mathrm{IP}) = 2^n 2^{-n/2} = 2^{n/2}$.

Now let $\mathcal{D}$ be the uniform distribution on $\{x : \mathrm{IP}_2(x) = 0 \pmod 2\}$. You can show that $\mathcal{D}$ is $2^{-n/2}$-biased. However, $\mathcal{D}$ cannot fool the inner product function IP_2.

It turns out that if you take two independent samples from a small-biased distribution and XOR them, you get a distribution which fools quadratic functions.

Theorem 1.1 (Viola)**.** *The sum of any independent d copies of an ε-biased distribution fools degree d polynomials over $\mathbb{F}_2$ with error $9\varepsilon^{1/2^{d-1}}$.*

If we let $\varepsilon = (\delta/9)^{2^{d-1}}$, where $\delta = 9\varepsilon^{1/2^{d-1}}$, then our seed length is $O(d\log(n/\varepsilon)) = O(d\log n + d2^d\log(1/\delta))$. For fixed d, this is great, but this stops being great if you take $d \geq \log n$. The proof uses the discrete derivative.

Definition 1.1. If $f \in \mathbb{F}_2[x_1, \ldots, x_n]$ with $\deg f = d$, the **discrete derivative** with respect to direction $y \in \mathbb{F}_2^n$ is

$$(\Delta_y f)(x) = f(x+y) - f(x).$$

Proposition 1.1. *If* $\deg f = d$,

$$\deg(\Delta_y f) \leq d - 1.$$

Example 1.2. For $f(x) = x_1 x_2$,

$$\Delta_y f(x) = (x_1 + y_1)(x_2 + y_2) - x_1 x_2$$

The proof of Viola's theorem proceeds by case analysis, looking at whether f is biased or unbiased. If bias $f \geq \delta$, then for all x,

$$f(x) \approx f(x) + f(x + y) = \Delta_y f(x).$$

for a random y. If bias $f \leq \delta$, then

$$\text{bias } f = |\,\mathbb{E}_{Y \sim U_n}[(-1)^{f(y)}]|.$$

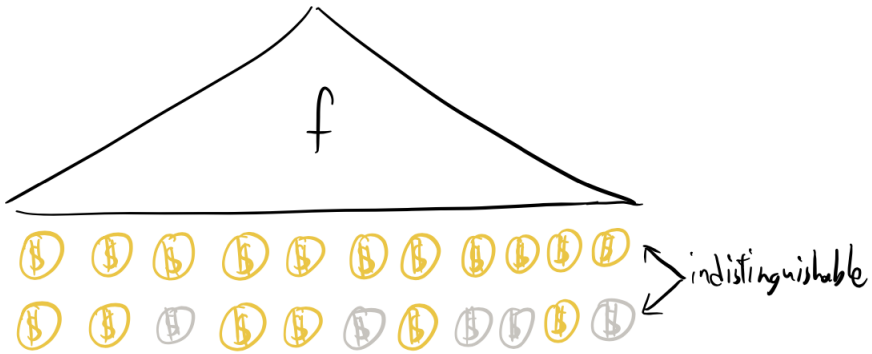
Then you need to argue that

$$\mathbb{E}_{X^{(1)},\ldots,X^{(d)}}[(-1)^{f(X^{(1)}+X^{(2)}+\cdots+X^{(d)})}]$$

is small.

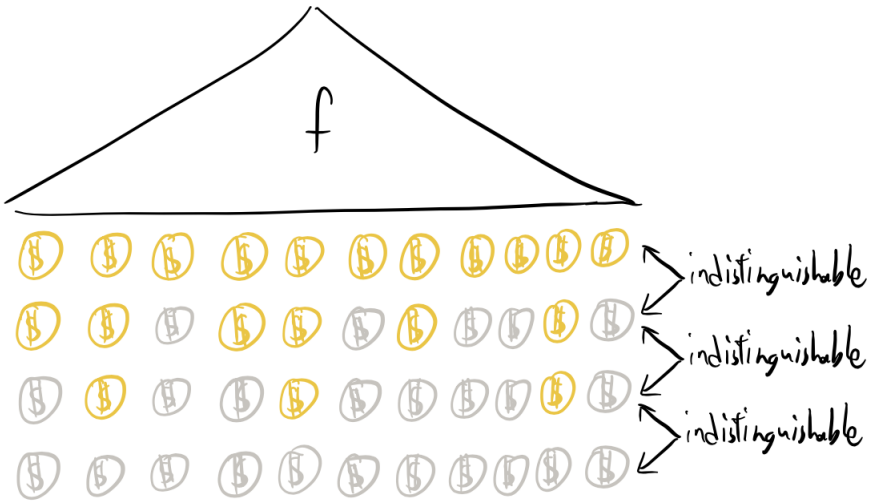
1.2 Restriction-based pseudorandom generation

Here is a thought experiment due to Ajtai and Wigderson. If we use a random restriction, we may get a simplified function which we can more easily fool. If we can fool a randomly restricted function, then the uniformly random and the partially uniformly random inputs should be indistinguishable to f:



Now we recursively replace the actually random bits (gold coins) with pseudorandom bits (silver coins). Then the pseudorandom bits should be indistinguishable from the actually

random bits.



To fool $f \in \mathcal{C}$, it sufficies to fool f under (pseudo)-random restrictions:

- Select $J \subseteq [n]$ pseudo-randomly, with $|J| \approx pn$.
- Select $x \sim \mathcal{D}$, a pseudorandom distribution on $\{\pm 1\}^n$.
- Select $z \sim U_{\overline{J}}$, the uniform distribution on $\overline{J}$, so that

$$\mathbb{E}_{Y \sim U_n}[f(Y)] \approx \mathbb{E}_J[\mathbb{E}_{X \sim \mathcal{D}, Z \sim U_{\overline{J}}}[f_{J,Z}(X)]$$

To get a pseudorandom generator, apply recursion $\frac{1}{p} \log n$ times.

Theorem 1.2 (Ajtai-Wigderson '85)**.** *It is enough for $\mathcal{D}$ to fool $f_{J,z}(x)$ for mosy choices of J, z.*

Theorem 1.3 (GMRTV '12)**.** *It is enough for $\mathcal{D}$ to fool their average*

$$\text{Bias } f(x) = \mathbb{E}_{J,Z \sim U_{\overline{J}}}[f_{J,Z}(x)].$$

Here is the Fourier analytical approach: The Fourier expansion of $f : \{\pm 1\}^n \to \{\pm 1\}$ is

$$f(x) = \sum_{S \subseteq [n]} \widehat{f}(S) \prod_{i \in S} x_I,$$

where

$$\widehat{f}(S) = \mathbb{E}_{X \sim \{\pm 1\}^n} \left[f(X) \prod_{i \in S} X_i \right], \qquad L_1(f) = \sum_{S \subseteq [n]} |\widehat{f}(S)|.$$

In 2013, Reingold, Steinke, and Vadhan considered $L_{1.k}(f) := \sum_{S:|S|=k} |\widehat{f}(S)|$.

Proposition 1.2. *Under p-random restrictions, Fourier coefficients of sets of size k shrink by a p^k factor.*

$$\mathbb{E}[L_{1,k}(\text{Bias } f)] = \sum_{k=0}^{n} p^k L_{1,k}(f).$$

If there exists a t such that for all k, $L_{1,k}(f) \leq t^k$, then picking $p = \frac{1}{2t}$ gives $\mathbb{E}[L_1(\text{Bias } f)] = O(1)$.

This tells us that under p-random restrictions, small-biased distributions fool the restricted function. So if this holds for pseudorandom J as well, this gives a PRG with $O(t \cdot \log^2 n)$ random bits.

1.3 Fractional pseudorandom generators

Assume that there exists a parameter t such that for any $f \in \mathcal{C}$ and for any k,

$$L_{1,k}(f) = \sum_{\substack{S \subseteq [n], \\ |S|=k}} |\widehat{f}(S)| \leq t^k$$

and that $\mathcal{C}$ is closed under restriction. Then the following approach will give a pseudorandom generator with seed-length $O(t^2 \log(n/\varepsilon))$.

The idea of CHHL is to "think inside the box." If $f : \{\pm 1\}^n \to \{\pm 1\}$ has Fourier expansion

$$f(x) = \sum_{S \subseteq [n]} \widehat{f}(S) \prod_{i \in S} x_i,$$

then we can view this as an extension of the function to $f : \mathbb{R}^n \to \mathbb{R}$.

Proposition 1.3. *If we restrict the domain of f to $[-1,1]^n$, then the range is contained in $[-1,1]$.*

Example 1.3. Let $f(x_1, x_2) = \frac{1}{2} - \frac{1}{2}x_1 + \frac{1}{2}x_2 - \frac{1}{2}x_1x_2$.

[insert picture 3]

Proof. If $\mu \in [-1,1]^n$, then we claim that

$$f(\mu) = \mathbb{E}_{X_1,\ldots,X_n}[f(X)],$$

where the X_i are independent with distribution

$$\mathbb{P}(X_i = 1) = \frac{1+\mu_i}{2}, \qquad \mathbb{P}(X_i = -1) = \frac{1-\mu_i}{2}.$$

This follows from the linearity of expectation:

$$\mathbb{E}_{X_1,\ldots,X_n}[f(X)] = \mathbb{E}\left[\sum_{S \subseteq [n]} \widehat{f}(S) \prod_{i \in S} X_i\right]$$

$$
\begin{aligned}
&= \sum_S \widehat{f}(S)\, \mathbb{E}\left[\prod_{i\in S} X_i\right] \\
&= \sum_S \widehat{f}(S) \prod_{i\in S} \mathbb{E}[X_i] \\
&= \sum_S \widehat{f}(S) \prod_{i\in S} \mu_i \\
&= f(\mu). \qquad \square
\end{aligned}
$$

Definition 1.2. An **ε-fractional pseudorandom generator** is a distribution $\mathcal{D}$ over $[-1,1]^n$ (sampleable with s random bits) such that for all $f \in \mathcal{C}$

$$|\,\mathbb{E}_{X\sim\mathcal{D}}[f(X)] - \mathbb{E}_{Y\sim U_n}[f(Y)]| \le \varepsilon.$$

Here we are comparing points inside the box with points on the corners.

Remark 1.1. Using the proof of the previous lemma,

$$\mathbb{E}_{Y\sim U_n}[f(Y)] = f(0).$$

So we can rewrite this condition as

$$|\,\mathbb{E}_{X\sim\mathcal{D}}[f(X)] - f(0)| \le \varepsilon.$$

The issue is that we can always just sample from 0 to get a fractional pseudorandom generator. But we really want to sample from the corners, so CHHL came up with the following condition.

Definition 1.3. An ε-fractional pseudorandom generator $\mathcal{D}$ is **p^2-noticable** if for all i

$$\mathbb{E}_{X\sim\mathcal{D}}[X_i^2] \ge p^2.$$

We will treat the case where X is drawn from a fractional pseudorandom generator over $[-p,p]^n$.

Proposition 1.4. *If we take $\frac{1}{2t}\mathcal{D}$, we get a fractional PRG that fools f with error $O(\varepsilon)$.*

Proof.

$$
\begin{aligned}
\left|\mathbb{E}_{X\sim\mathcal{D}}\left[f\left(\frac{1}{2t}X\right)\right] - f(0)\right| &= \left|\mathbb{E}_{X\sim\mathcal{D}}\left[\sum_S \widehat{f}(S)\prod_{i\in S}\left(\frac{1}{2t}\right)X^i\right] - \widehat{f}(\varnothing)\right| \\
&= \left|\mathbb{E}_{X\sim\mathcal{D}}\left[\sum_{S\neq\varnothing} \widehat{f}(S)\prod_{i\in S}\left(\frac{1}{2t}\right)X^i\right]\right|
\end{aligned}
$$

$$
\begin{aligned}
&\leq \sum_{S \neq \varnothing} |\widehat{f}(S)| \left(\frac{1}{2t}\right)^{|S|} \left| \mathbb{E}_{X \sim \mathcal{D}} \left[\prod_{i \in S} X_i \right] \right| \\
&\leq \varepsilon \sum_{\varnothing \neq S \subseteq [n]} |\widehat{f}(S)| \left(\frac{1}{2t}\right)^{|S|} \\
&\leq \varepsilon \sum_{k=1}^{n} L_{1,k}(f) \left(\frac{1}{2t}\right)^{k} \\
&= \varepsilon \sum_{k=1}^{n} \frac{t^k}{(2t)^k} \\
&\leq \varepsilon.
\end{aligned}
$$

□

What CHHL showed is that if $\mathcal{C}$ is closed under restriction, then we can use random restrictions to get a PRG from a fractional PRG. It is conjectured that low-degree polynomials over $\mathbb{F}_2$ have these desired properties.