Pseudorandom Generators via Polarizing Random Walks (lecture notes)

Course: Analysis of Boolean Functions, Autumn 2025, University of Chicago
Instructor: William Hoza (`williamhoza@uchicago.edu`)

In these lecture notes, we prove the following.

Theorem 0.1 (Fourier growth bounds imply foolability). *For every $n, b \in \mathbb{N}$ and every $\varepsilon \in (0,1)$, there exists an explicit PRG $G: \{\pm 1\}^r \to \{\pm 1\}^n$ with the following properties.*

- *Let $\mathcal{F}$ be a class of Boolean functions $f: \{\pm 1\}^n \to \mathbb{R}$ that is closed under restrictions. Assume that for every $f \in \mathcal{F}$ and every $k \in \mathbb{N}$, we have $L_{1,k}(f) \leq b^k$. Then G fools $\mathcal{F}$ with error ε.*
- *The seed length is $r = \widetilde{O}(b^2 \cdot \log(n/\varepsilon) \cdot \log(1/\varepsilon))$.*

For example, if $\mathcal{F}$ is the class of size-s AC^0_d circuits, then we can take $b = O(\log s)^{d-1}$, so the seed length is $\mathrm{polylog}(ns/\varepsilon)$.

As another example, recall that we proved that width-w oblivious regular ROBPs satisfy $L_{1,k}(f) \leq w^k$. Unfortunately, the class of width-w oblivious regular ROBPs is not closed under restriction. However, we can consider the subclass of width-w oblivious *permutation* ROBPs. These are width-w oblivious ROBPs in which there are no "collisions," i.e., edges with the same label pointing to the same vertex. This class is closed under restrictions, so by Theorem 0.1, we can fool it using a seed of length $\widetilde{O}(w^2 \cdot \log(n/\varepsilon) \cdot \log(1/\varepsilon))$.

1 Fractional PRGs

The proof of Theorem 0.1 is based on the notion of a *fractional PRG*. A fractional PRG is a function $G: \{\pm 1\}^r \to [-1,1]^n$. And what does it mean for a fractional PRG to "fool" a Boolean function? The Fourier expansion is the key to making sense of it.

Definition 1.1 (Multilinear extension). For any $f: \{\pm 1\}^n \to \mathbb{R}$, the *multilinear extension* of f is the function $f: \mathbb{R}^n \to \mathbb{R}$ defined by

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

When $x \in [-1,1]^n$, there is a nice probabilistic interpretation of $f(x)$. We use the following notation.

Definition 1.2 (Product distribution notation). For $x \in [-1,1]^n$, let Π_x be the unique product distribution over $\{\pm 1\}^n$ with expectation x.

Claim 1.3. *Let $f: \{\pm 1\}^n \to \mathbb{R}$, and extend f to $\mathbb{R}^n$ via the Fourier expansion. Then for every $x \in [-1,1]^n$, we have $f(x) = \mathbb{E}_{y \sim \Pi_x}[f(y)]$. More generally, if X has a product distribution over $[-1,1]^n$, then $\mathbb{E}[f(X)] = f(\mathbb{E}[X])$.*

Proof.

$$\mathbb{E}[f(X)] = \sum_{S \subseteq [n]} \widehat{f}(S) \cdot \mathbb{E}\left[\prod_{i \in S} X_i\right] = \sum_{S \subseteq [n]} \widehat{f}(S) \cdot \prod_{i \in S} \mathbb{E}[X_i] = f(\mathbb{E}[X]).$$ □

It follows that if $f: \{\pm 1\}^n \to \{\pm 1\}$, then its multilinear extension maps $[-1,1]^n \to [-1,1]$. Furthermore, $f(0^n) = \mathbb{E}[f]$.

Definition 1.4 (Fractional PRGs). Let $f: \{\pm 1\}^n \to \mathbb{R}$, and extend f to $\mathbb{R}^n$ via the Fourier expansion. Let X be a random variable taking values in $[-1,1]^n$. We say that X *fools* f with error δ if $|\mathbb{E}[f(X)] - \mathbb{E}[f]| \leq \delta$. Let $G: \{\pm 1\}^r \to [-1,1]^n$. We say that G fools f with error δ if $G(U_r)$ fools f with error δ.

For example, if G always outputs 0^n, then G trivially fools every f with error 0. The first step of proving Theorem 0.1 is to construct a fractional PRG that fools $\mathcal{F}$ and that takes values in $\{\pm p\}^n$ for a not-too-small value p.

Lemma 1.5 (Fractional PRG based on Fourier growth bounds)**.** *For every $n, b \in \mathbb{N}$ and every $\delta \in (0,1)$, there exists an explicit fractional PRG $G\colon \{\pm 1\}^r \to \{\pm \frac{1}{2b}\}^n$ with the following properties.*

- *Let $f\colon \{\pm 1\}^n \to \mathbb{R}$. Assume that for every $k \in \mathbb{N}$, we have $L_{1,k}(f) \le b^k$. Then G fools f with error δ.*
- *The seed length of G is $O(\log(1/\delta) + \log\log n)$.*
- *Each individual bit of $G(U_r)$ is uniformly distributed over $\{\pm \frac{1}{2b}\}$.*

Proof. The PRG G samples $X \in \{\pm 1\}^n$ from a k-wise γ-biased distribution and samples $\sigma \in \{\pm 1\}$ uniformly at random, and then it outputs $p\sigma X \in \{\pm p\}^n$, where $k = \lceil \log(2/\delta) \rceil$, $\gamma = \delta/2$, and $p = \frac{1}{2b}$. The seed length is $O(\log(k/\gamma) + \log\log n) = O(\log(1/\delta) + \log\log n)$. Each individual output bit is uniform because of σ. Now let us prove that the generator fools f. We have

$$
\begin{aligned}
|\mathbb{E}[f(p\sigma X)] - \mathbb{E}[f]| = \left| \sum_{S \neq \varnothing} \widehat{f}(S) \cdot \mathbb{E}\left[\prod_{i \in S} (p\sigma X)_i\right] \right| &\le \sum_{S \neq \varnothing} |\widehat{f}(S)| \cdot p^{|S|} \cdot \left| \mathbb{E}\left[\prod_{i \in S} X_i\right] \right| \\
&\le \left(\sum_{d=1}^{k} L_{1,d}(f) \cdot p^d \cdot \gamma \right) + \left(\sum_{d=k+1}^{n} L_{1,d}(f) \cdot p^d \right) \\
&\le \gamma \cdot \sum_{d=1}^{k} (pb)^d + \sum_{d=k+1}^{n} (pb)^d \\
&\le \gamma + 2^{-k} \\
&\le \delta.
\end{aligned}
$$

□

2 Polarizing Random Walks

Theorem 1.5 is the only part of the proof of Theorem 0.1 that uses the assumed $L_{1,k}$ bound. The rest of the proof is a generic transformation from fractional PRGs to non-fractional PRGs. Here's how it works. Let X be a distribution over $\{\pm p\}^n$ that δ-fools $\mathcal{F}$. We construct the following non-fractional PRG:

1. Sample t independent samples from X, say $X^{(1)}, \dots, X^{(t)}$, for a suitable value $t = O(\frac{\log(n/\varepsilon)}{p^2})$.
2. Let $Y^{(0)} = 0^n$, and for $j > 0$ define
$$Y^{(j)} = Y^{(j-1)} + \Delta_{Y^{(j-1)}} \odot X^{(j)},$$
where $\Delta_y := (1 - |y_1|, 1 - |y_2|, \dots, 1 - |y_n|)$ and $\odot$ is coordinatewise multiplication.
3. Output $\operatorname{sign}(Y^{(t)})$, where $\operatorname{sign}(\cdot)$ is applied coordinatewise.

We will prove that $\operatorname{sign}(Y^{(t)})$ fools $\mathcal{F}$ with error $O(\varepsilon \cdot t)$.

The construction can be interpreted as a *pseudorandom walk* through $[-1,1]^n$. We start at 0^n. We use $X^{(j)}$ to decide which direction to move in step j, and the magnitude of the step is determined based on the current location.

The first step of the analysis is to prove that a single step doesn't do much harm.

Lemma 2.1 (One step doesn't do much harm)**.** *Let $f\colon \{\pm 1\}^n \to \mathbb{R}$. Assume that X fools every restriction of f with error δ. Then for every $y \in [-1,1]^n$, we have $|f(y) - \mathbb{E}[f(y + \Delta_y \odot X)]| \le \delta$.*

Proof. Sample $\rho \in \{+1, -1, \star\}^n$ in which the coordinates are independent and

$$\rho_i = \begin{cases} \text{sign}(y_i) & \text{with probability } |y_i| \\ \star & \text{with probability } 1 - |y_i|. \end{cases}$$

For each $x \in [-1,1]^n$, define $\rho \circ x \in [-1,1]^n$ by the rule

$$(\rho \circ x)_i = \begin{cases} x_i & \text{if } \rho_i = \star \\ \rho_i & \text{if } \rho_i \neq \star. \end{cases}$$

Then $\rho \circ x$ is a product distribution, so $\mathbb{E}[f|_\rho(x)] = \mathbb{E}[f(\rho \circ x)] = f(\mathbb{E}[\rho \circ x])$. Now, in each individual coordinate, we have

$$\mathbb{E}[(\rho \circ x)_i] = \text{sign}(y_i) \cdot |y_i| + x_i \cdot (1 - |y_i|) = y_i + x_i \cdot (1 - |y_i|),$$

so $\mathbb{E}[\rho \circ x] = y + \Delta_y \odot x$. Therefore,

$$\begin{aligned} |\,\mathbb{E}[f(y + \Delta_y \odot X)] - f(y)| &= |\,\mathbb{E}[f(y + \Delta_y \odot X)] - f(y + \Delta_y \odot 0^n)| \\ &= |\,\mathbb{E}[f|_\rho(X)] - \mathbb{E}[f|_\rho(0^n)]| \\ &\leq \delta. \end{aligned}$$

□

The next step of the analysis is to show that the random walk *polarizes*, meaning that $Y^{(t)}$ is close to $\{\pm 1\}^n$ with high probability. We will focus on a single coordinate ($n = 1$) and eventually do a union bound.

Lemma 2.2. *Assume $n = 1$. Then* $\Pr[\Delta_{Y^{(t)}} \geq \cdot e^{-tp^2/5}] \leq e^{-tp^2/50}$.

Proof. Since $n = 1$, X is just the uniform distribution over $\{\pm p\}$. So in step j, independently of whatever happened before, there is a 50% chance that we take a "good" step, meaning $X = p \cdot \text{sign}(Y^{(j-1)})$, and there is a 50% chance that we take a "bad" step, meaning $X = -p \cdot \text{sign}(Y^{(j-1)})$. In a "good" step, the distance $\delta_{Y^{(j-1)}}$ decreases by a factor of $1 - p$. In a "bad" step, the distance might increase, but at most it increases by a factor of $1 + p$. By Hoeffding's inequality, except with probability $e^{-tp^2/50}$, the number of good steps is at least $(1/2 - p/10) \cdot t$. In this case,

$$\begin{aligned} \Delta_{Y^{(t)}} \leq (1-p)^{(1/2 - p/10)\cdot t} \cdot (1+p)^{(1/2 + p/10)\cdot t} &= (1 - p^2)^{(1/2 - p/10)\cdot t} \cdot (1+p)^{0.2pt} \\ &\leq (1 - p^2)^{0.4t} \cdot (1+p)^{0.2pt} \\ &\leq e^{-0.4p^2 t} \cdot e^{0.2p^2 t} \\ &= e^{-0.2p^2 t}. \end{aligned}$$

□

The final step is to argue that outputting $\text{sign}(Y^{(t)})$ instead of $Y^{(t)}$ itself doesn't do too much harm.

Lemma 2.3. *Let $f \colon \{\pm 1\}^n \to \{\pm 1\}$, and extend f to $\mathbb{R}^n$ via the Fourier expansion. Then for every $y \in [-1,1]^n$, we have*

$$|f(y) - f(\text{sign}(y))| \leq \|\Delta_y\|_1.$$

Proof.

$$\begin{aligned} |f(y) - f(\text{sign}(y))| &= |\,\mathbb{E}[f(\Pi_y)] - f(\text{sign}(y))| \\ &\leq 2 \cdot \Pr_{x \sim \Pi_y}[x \neq \text{sign}(y)] \\ &\leq 2 \cdot \sum_{i=1}^n \Pr_{x \sim \Pi_y}[x_i \neq \text{sign}(y_i)] \\ &= 2 \cdot \sum_{i=1}^n \frac{1 - |y_i|}{2} \\ &= \|\Delta_y\|_1. \end{aligned}$$

□

Proof sketch of Theorem 0.1. Let X be the fractional PRG from Theorem 1.5 with error δ. Our PRG outputs the corresponding $\mathrm{sign}(Y^{(t)})$. By the analysis above, the error of this PRG is at most

$$t\delta + 2n \cdot e^{-tp^2/50} + n \cdot e^{-tp^2/5} \leq \varepsilon,$$

provided we choose $\delta = \frac{\varepsilon}{2t}$. The seed length is

$$\begin{aligned} t \cdot O(\log(1/\delta) + \log\log n) = O(t\log t + t\log(1/\varepsilon) + t\log\log n) &= \widetilde{O}(p^{-2} \cdot \log(n/\varepsilon) \cdot \log(1/\varepsilon)) \\ &= \widetilde{O}(b^2 \cdot \log(n/\varepsilon) \cdot \log(1/\varepsilon)). \end{aligned}$$

□