Small-Bias Distributions and Almost k-wise Uniformity (lecture notes)

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

In these lecture notes, we will see some applications of Fourier analysis to the problem of constructing *pseudorandom generators* (PRGs). The general definition is as follows.

Definition 0.1 (Fooling)**.** Let D be a distribution over $\{\pm 1\}^n$, let $f\colon \{\pm 1\}^n \to \mathbb{R}$, and let $\varepsilon \in (0,1)$. We say that D fools f with error ε if $|\mathbb{E}_{x\sim D}[f(x)] - \mathbb{E}[f]| \leq \varepsilon$.

Definition 0.2 (PRG)**.** A PRG is a function $G\colon \{\pm 1\}^r \to \{\pm 1\}^n$. We say that G fools $f\colon \{\pm 1\}^n \to \mathbb{R}$ with error ε if $G(U_r)$ fools f with error ε. Here U_r denotes the uniform distribution over $\{\pm 1\}^r$.

We want the *seed length* r and the *error* ε to be as small as possible. We want to fool as rich a class of functions f as possible. The function f represents a "user" of the PRG. For example, f might represent a randomized algorithm, and saying that G fools f means that we can safely use G to simulate the randomized algorithm without significantly distorting its behavior.

We also want the generator to be efficiently computable. We say that G is *explicit* if $G(x)$ can be computed in polynomial time, given x, n, ε, and any parameters that define the class of functions that we are trying to fool (hopefully clear from context).

The PRGs we will present will use finite field arithmetic, so we'll begin by reviewing some basic facts about finite fields.

1 Finite fields

As a reminder, a field is a set equipped with two binary operations, addition and multiplication, such that the following axioms hold: addition and multiplication satisfy the associative, commutative, and distributive laws; there are additive and multiplicative identities (0 and 1); every element x has an additive inverse $-x$; and every nonzero element x has a multiplicative inverse x^{-1}. For example, $\mathbb{R}$, $\mathbb{Q}$, and $\mathbb{C}$ are all infinite fields. We are more interested in finite fields, which are described by the following two theorems.

Theorem 1.1 (Finite field existence)**.** *Let q be a positive integer. There exists a finite field with cardinality q if and only if q is a power of a prime.*

Theorem 1.2 (Finite field uniqueness)**.** *Every two finite fields of the same cardinality are isomorphic.*

In light of Theorems 1.1 and 1.2, for each prime power q, we can sensibly speak of "the" finite field of cardinality q, denoted $\mathbb{F}_q$. We won't fully prove Theorems 1.1 and 1.2, but we'll say a bit more about how to construct $\mathbb{F}_q$. The simplest case is $\mathbb{F}_p$ where p is prime. As a set, we have $\mathbb{F}_p = \{0, 1, \dots, p-1\}$. We add and multiply elements modulo p. (Caution: If m is composite, then the integers modulo m do *not* satisfy the field axioms!)

To construct more fields, we draw inspiration from the complex numbers. Recall that $\mathbb{C}$ can be constructed by starting from $\mathbb{R}$, introducing a new "imaginary unit" i, and declaring that $i^2 = -1$, or equivalently $i^2 + 1 = 0$. It is only natural to try using a different polynomial instead of $i^2 + 1$. Such a construction still gives us a field, provided we work with an *irreducible* polynomial. We say that a polynomial $E(x)$ is *irreducible* if it cannot be factored into two non-constant polynomials.

Theorem 1.3. *For every prime number p and every positive integer d, there exists a degree-d irreducible polynomial $E(x) \in \mathbb{F}_p[x]$.*

Theorem 1.3 can be proven using a counting argument: one shows that a noticeable fraction of all degree-d monic polynomials are in fact irreducible. Such a proof is rather unsatisfying, because it doesn't actually give us any explicit *examples* of irreducible polynomials. Fortunately, it is possible to cook up some very nice explicit examples:

Theorem 1.4. *For every nonnegative integer m, the polynomial $x^{2\cdot 3^m} + x^{3^m} + 1$ is irreducible over $\mathbb{F}_2$.*

We'll skip the proofs of Theorems 1.3 and 1.4. See the textbook "Introduction to Coding Theory" by van Lint if you're curious.

Suppose $E(x)$ is a degree-d polynomial that is irreducible over $\mathbb{F}_p$. Then we can construct $\mathbb{F}_{p^d}$ as $\mathbb{F}_p[x]/E(x)$, i.e., the set of polynomials over $\mathbb{F}_p$ modulo $E(x)$. This is completely analogous to the construction $\mathbb{C} = \mathbb{R}[i]/(i^2+1)$.

For example, suppose d is two times a power of three. In this case, we can explicitly construct $\mathbb{F}_{2^d}$ using Theorem 1.4. As a set, we can take $\mathbb{F}_{2^d} = \{0,1\}^d$. Addition is simply bitwise XOR. Multiplication is only slightly more complicated. We interpret each field element $u \in \{0,1\}^d$ as a vector of coefficients of a polynomial over $\mathbb{F}_2$. To multiply two field elements, first we do ordinary polynomial multiplication, then we repeatedly replace each occurrence of x^d with $x^{d/2}+1$ until the degree is less than d.

2 Limited independence

One of the most "basic" types of PRGs is a PRG for generating k-wise independent bits.

Definition 2.1 (k-wise uniformity)**.** Let D be a distribution over $\{\pm 1\}^n$ and let $k \in [n]$. We say that D is *k-wise uniform* if every k coordinates of D are distributed uniformly over $\{0,1\}^k$. Equivalently, this means that D fools k-juntas with error 0. A *k-wise uniform generator* is a PRG $G\colon \{0,1\}^r \to \{0,1\}^n$ such that $G(U_r)$ is k-wise uniform.

We can construct a k-wise uniform generator using the fact that low degree polynomials have few roots over any field.

Lemma 2.2 (The degree mantra)**.** *Let $\mathbb{F}$ be a field, and let p be a nonzero polynomial over $\mathbb{F}$ of degree at most d. Then p has at most d distinct roots.*

Proof. If $d = 0$, this is obvious. Now suppose $d > 0$. Let a be a root of p. By polynomial long division, we can write $p(x) = (x-a)\cdot q(x) + b$ for some $q(x) \in \mathbb{F}[x]$ and some $b \in \mathbb{F}$. By plugging in $x = a$, we see that $b = 0$, so $p(x) = (x-a)\cdot q(x)$. Every root of p other than a must be a root of q, because in a field, a product of nonzero elements is always nonzero. Finally, $\deg(q) \le d-1$, so by induction, q has at most $d-1$ distinct roots. □

Theorem 2.3. *For every $n, k \in \mathbb{N}$, there exists an explicit k-wise uniform generator $G\colon \{0,1\}^r \to \{0,1\}^n$ with seed length $r = O(k \cdot \log n)$. Furthermore, as a map $G\colon \mathbb{F}_2^r \to \mathbb{F}_2^n$, the generator G is a linear transformation.*

Proof.

- Construction: Let d be a number of the form $2\cdot 3^m$ such that $\log n \le d \le 3\log n$. Since d is two times a power of three, we can construct $\mathbb{F}_{2^d}$ using Theorem 1.4. Let $\mathcal{P}$ be the set of polynomials $p(x) \in \mathbb{F}_{2^d}[x]$ of degree less than k. Let $z_1, z_2, \dots, z_n$ be distinct elements of $\mathbb{F}_{2^d}$. Define $G\colon \mathcal{P} \to \{0,1\}^n$ by

$$G(p) = (p(z_1)_1, \dots, p(z_n)_1),$$

where u_1 denotes the first bit of $u \in \{0,1\}^d = \mathbb{F}_{2^d}$.

- Seed length: A polynomial $p \in \mathcal{P}$ can be represented as a vector of k coefficients, which consists of $r = k \cdot d = O(k \log n)$ bits.

- Correctness: Fix any k distinct coordinates $i_1, i_2, \dots, i_k \in [n]$. Define $H \colon \mathcal{P} \to \mathbb{F}_{2^d}^k$ by

  $$H(p) = (p(z_{i_1}), \dots, p(z_{i_k})).$$

  Then H is injective, because if $H(p) = H(p')$, then $p - p'$ is a polynomial of degree at most $k - 1$ with at least k zeroes, hence $p - p' \equiv 0$ by the degree mantra. Furthermore, the domain and codomain of H are finite sets of equal cardinality. Therefore, H is bijective. So if we pick $p \in \mathcal{P}$ uniformly at random, then $H(p)$ is a uniform random element of $\mathbb{F}_{2^d}^k$. It follows that coordinates $i_1, i_2, \dots, i_k$ of $G(p)$ are distributed uniformly over $\{0,1\}^k$.

- Linearity: Addition in the vector space $\mathbb{F}_2^d$ is just bitwise XOR, the same as addition in the field $\mathbb{F}_{2^d}$. Therefore, $G(p + p') = G(p) + G(p')$.

□

3 Small-bias distributions

The Fourier-analytic perspective is that every Boolean function is a linear combination of character functions. Therefore, from a Fourier-analytic perspective, the most "basic" type of PRG is one that fools character functions.

Definition 3.1 (Bias). Let D be a distribution over $\{\pm 1\}^n$ and let $\varepsilon \in (0,1)$. We say that D is *ε-biased* if D fools all character functions χ_S with error ε. An *ε-biased generator* is a PRG with an ε-biased output distribution.

Another way of understanding the definition is to consider the *probability density function* of D. This is the function $\varphi_D \colon \{\pm 1\}^n \to \mathbb{R}$ defined by

$$\varphi_D(x) = 2^n \cdot \Pr_{y \sim D}[y = x].$$

In other words, its the probability mass function of D, scaled by a factor of 2^n. Every probability density function satisfies $\widehat{\varphi_D}(\varnothing) = 1$.

Claim 3.2. *D is ε-biased if and only if for every nonempty $S \subseteq [n]$, we have $|\widehat{\varphi_D}(S)| \le \varepsilon$.*

Proof.

$$\widehat{\varphi_D}(S) = 2^{-n} \cdot \sum_{x \in \{\pm 1\}^n} \varphi_D(x) \cdot \chi_S(x) = \sum_{x \in \{\pm 1\}^n} \chi_S(x) \cdot \Pr_{y \sim D}[y = x] = \mathop{\mathbb{E}}_{x \sim D}[\chi_S(x)]. \qquad \square$$

We now present an explicit construction of a small-bias generator. The construction once again uses low-degree polynomials over a finite field, just like our k-wise uniform generator, but the details are different.

Theorem 3.3. *For every $n \in \mathbb{N}$ and $\varepsilon \in (0,1)$, there exists an explicit ε-biased generator $G \colon \{0,1\}^r \to \{0,1\}^n$ with seed length $r = O(\log(n/\varepsilon))$.*

Proof. Let d be a number of the form $2 \cdot 3^m$ such that $\log(n/\varepsilon) \le d \le 3\log(n/\varepsilon)$. Since d is two times a power of three, we can construct $\mathbb{F}_{2^d}$ using Theorem 1.4. We think of G as a map $G \colon \mathbb{F}_{2^d} \times \mathbb{F}_{2^d} \to \mathbb{F}_2^n$, so the seed length is $r = 2d = O(\log(n/\varepsilon))$. The generator is given by

$$G(u, v) = (\mathrm{IP}(u^1, v), \mathrm{IP}(u^2, v), \dots, \mathrm{IP}(u^n, v)),$$

where $\mathrm{IP} \colon \{0,1\}^d \times \{0,1\}^d \to \{0,1\}$ is inner product modulo 2, i.e., $\mathrm{IP}(u, v) = \bigoplus_{i=1}^d u_i v_i$. To prove that this works, consider any character function $\chi_a(x)$ where a is a nonzero element of $\mathbb{F}_2^n$. We have $\chi_a(x) = (-1)^{\sum_{i=1}^n a_i x_i}$. Therefore,

$$\chi_a(G(u, v)) = (-1)^{\sum_{i=1}^n a_i \cdot \mathrm{IP}(u^i, v)} = (-1)^{\mathrm{IP}\left(\sum_{i=1}^n a_i \cdot u^i, v\right)} = \chi_{p_a(u)}(v),$$

where $p_a(u) = \sum_{i=1}^n a_i \cdot u^i \in \mathbb{F}_2^d$. Observe that p_a is a polynomial over $\mathbb{F}_{2^d}$ of degree at most n. By the degree mantra, p_a has at most n roots. Therefore,

$$\left|\mathop{\mathbb{E}}_{u,v}[\chi_a(G(u,v))]\right| = \left|\mathop{\mathbb{E}}_{u,v}[\chi_{p_a(u)}(v)]\right| \leq 2^{-k} \cdot \sum_{u \in \mathbb{F}_2^k} \left|\mathop{\mathbb{E}}_{v}[\chi_{p_a(u)}(v)]\right| \leq 2^{-k} \cdot n \leq \varepsilon. \qquad \square$$

4 Almost k-wise uniformity

Theorem 2.3 shows how to sample k-wise uniform bits using $O(k \cdot \log n)$ truly random bits. The factor of $\log n$ in this seed length is perhaps a bit disappointing. Unfortunately, it turns out that $O(k \cdot \log n)$ is optimal, unless $k \approx n$. The good news is that it is possible to improve the seed length, if we are willing to tolerate a small amount of error.

Definition 4.1 (Almost k-wise uniformity)**.** Let D be a distribution over $\{\pm 1\}^n$, let $k \in [n]$, and let $\varepsilon \in (0,1)$. We say that D is *ε-almost k-wise uniform* if D fools $\{0,1\}$-valued k-juntas with error ε. Equivalently, every k coordinates of D are distributed within total variation distance ε of U_k. An *ε-almost k-wise uniform generator* is a PRG $G\colon \{0,1\}^r \to \{0,1\}^n$ such that $G(U_r)$ is ε-almost k-wise uniform.

4.1 Warm-up

As an application of small-bias distributions, we will construct an ε-almost k-wise uniform generator with seed length $O(k + \log(1/\varepsilon) + \log\log n)$. First, as a warm-up, let's construct a generator with seed length $O(k + \log(n/\varepsilon))$. The key is the following lemma.

Lemma 4.2. *Let $f\colon \{\pm 1\}^n \to \mathbb{R}$, and let D be a distribution over $\{\pm 1\}^n$. If D is ε-biased, then D fools f with error $\varepsilon \cdot \hat{\|} f \hat{\|}_1$.*

Proof.

$$\left|\mathop{\mathbb{E}}_{x\sim D}[f(x)] - \mathbb{E}[f]\right| = \left|\sum_{S \subseteq [n]} \widehat{f}(S) \cdot \left(\mathop{\mathbb{E}}_{x \sim D}[\chi_S(x)] - \mathbb{E}[\chi_S]\right)\right| \leq \sum_{S \neq \varnothing} |\widehat{f}(S)| \cdot \left|\mathop{\mathbb{E}}_{x\sim D}[\chi_S(x)]\right| \leq \varepsilon \cdot \hat{\|} f \hat{\|}_1. \qquad \square$$

Corollary 4.3 (Warm-up)**.** *For every $n, k \in \mathbb{N}$ and $\varepsilon \in (0,1)$, there is an explicit ε-almost k-wise uniform PRG $G\colon \{\pm 1\}^r \to \{\pm 1\}^n$ with seed length $r = O(k + \log(n/\varepsilon))$.*

Proof. Let G be an $(\varepsilon \cdot 2^{-k/2})$-biased generator. Then for any k-junta $f\colon \{\pm 1\}^n \to [-1,1]$, we have

$$\begin{aligned} |\mathbb{E}[f(G(U_r))] - \mathbb{E}[f]| &\leq \varepsilon \cdot 2^{-k/2} \cdot \hat{\|} f \hat{\|}_1 && \text{by the lemma} \\ &\leq \varepsilon \cdot 2^{-k/2} \cdot 2^{k/2} \cdot \|f\|_2 && \text{by Exercise 3a} \\ &\leq \varepsilon. && \square \end{aligned}$$

4.2 The improved PRG

To improve the seed length, we will *compose* an exact k-wise uniform generator with a small-bias generator. For the analysis, it is helpful to define a notion of *k-wise small-bias distributions.*

Definition 4.4 (k-wise small-bias)**.** Let D be a distribution over $\{\pm 1\}^n$, let $k \in [n]$, and let $\varepsilon \in (0,1)$. We say that D is *k-wise ε-biased* if, for every $S \subseteq [n]$ with $|S| \leq k$, the distribution D fools the character function χ_S with error ε. We say that a PRG $G\colon \{\pm 1\}^r \to \{\pm 1\}^n$ is a *k-wise ε-biased generator* if $G(U_r)$ is k-wise ε-biased.

Theorem 4.5. *For every $n, k \in \mathbb{N}$ and $\varepsilon \in (0,1)$, there exists an explicit k-wise ε-biased generator $G\colon \{0,1\}^r \to \{0,1\}^n$ with seed length $r = O(\log(k/\varepsilon) + \log\log n)$.*

Proof. Let $M \in \mathbb{F}_2^{n \times r_0}$ be the matrix representation of the k-wise uniform generator from Theorem 2.3. The generator G samples an ε-biased seed $u \in \mathbb{F}_2^{r_0}$, then outputs Mu. To prove that this works, consider any character function $\chi_a \colon \mathbb{F}_2^n \to \{\pm 1\}$, where $a \in \mathbb{F}_2^n$ and a has Hamming weight at most k. Then

$$\chi_a(Mu) = (-1)^{\sum_{i=1}^n a_i \cdot (Mu)_i} = (-1)^{\sum_{i=1}^n a_i \cdot \sum_{j=1}^{r_0} M_{ij} u_j} = (-1)^{\sum_{j=1}^{r_0} \left(\sum_{i=1}^n a_i \cdot M_{ij}\right) \cdot u_j} = \mathbb{E}[\chi_b(u)],$$

where $b_j = \sum_{i=1}^n a_i \cdot M_{ij}$ for each $j \in [r_0]$. Since u is sampled from an ε-biased distribution, we have $|\mathbb{E}[\chi_b(u)] - \mathbb{E}[\chi_b(u')]| \leq \varepsilon$, where $u' \in \mathbb{F}_2^{r_0}$ is selected uniformly at random. Therefore,

$$|\mathbb{E}[\chi_a(Mu)] - \mathbb{E}[\chi_a(Mu')]| \leq \varepsilon.$$

Finally, χ_a is a k-junta, so k-wise uniformity tells us $\mathbb{E}[\chi_a(Mu')] = \mathbb{E}[\chi_a]$. □

Lemma 4.6. *Let $f \colon \{\pm 1\}^n \to \mathbb{R}$, and let D be a distribution over $\{\pm 1\}^n$. If D is k-wise ε-biased and $\deg(f) \leq k$, then D fools f with error $\varepsilon \cdot \hat{\|f\|}_1$.*

Proof.

$$\left| \mathop{\mathbb{E}}_{x \sim D}[f(x)] - \mathbb{E}[f] \right| = \left| \sum_{S \subseteq [n]} \widehat{f}(S) \cdot \left(\mathop{\mathbb{E}}_{x \sim D}[\chi_S(x)] - \mathbb{E}[\chi_S] \right) \right| \leq \sum_{0 < |S| \leq \deg(f)} |\widehat{f}(S)| \cdot \left| \mathop{\mathbb{E}}_{x \sim D}[\chi_S(x)] \right| \leq \varepsilon \cdot \hat{\|f\|}_1. \quad \square$$

Corollary 4.7. *For every $n, k \in \mathbb{N}$, there exists an explicit ε-almost k-wise uniform PRG $G \colon \{\pm 1\}^r \to \{\pm 1\}^n$ with seed length $r = O(k + \log(1/\varepsilon) + \log\log n)$.*

Proof. Let G be a k-wise $(\varepsilon \cdot 2^{-k/2})$-biased generator. Then for any k-junta $f \colon \{\pm 1\}^n \to [-1, 1]$, we have

$$\begin{aligned} |\mathbb{E}[f(G(U_r))] - \mathbb{E}[f]| &\leq \varepsilon \cdot 2^{-k/2} \cdot \hat{\|f\|}_1 && \text{by the lemma} \\ &\leq \varepsilon \cdot 2^{-k/2} \cdot 2^{k/2} \cdot \|f\|_2 && \text{by Exercise 3a} \\ &\leq \varepsilon. \end{aligned}$$

□

The seed length in Theorem 4.7 is shockingly small. For example, if k and ε are constants, then the generator has *doubly exponential stretch*, i.e., it stretches a seed of length r out to a pseudorandom string of length $n = 2^{2^{\Omega(r)}}$.