COMS 6998: Unconditional Lower Bounds and Derandomization **Spring 2024**

Lecture 7: February 27, 2024

Lecturer: Rocco Servedio *Scribes: Akshat Yaparla, Ashvin Jagadeesan*

1 Lecture Overview

In this lecture, we cover the following ideas. Broadly speaking, we concern ourselves with *k-wise independent random variables* and *ϵ-bias distributions*. Here are the contents in detail.

- k-wise independent/uniform random variables.
  - Pairwise independence
  - Derandomization application: MaxCut
  - Constructing k-wise uniformly random variables
  - Derandomization application: Max-3SAT
  - PRGs for k-juntas and depth-k decision trees
- ϵ-bias distributions
  - PAR_S
  - ϵ-bias random variables
  - ϵ-generator for ϵ-biased random variable
  - Combine k-wise, ϵ-biased random variable

2 k-wise Independent/Uniform Random variables

In this section, we introduce the concept of k-wise independent random variables and k-wise uniform random variables.

Definition 1. *Let $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$ be a set of random variables with support over A. The sequence, $(\mathbf{X}_1, \ldots, \mathbf{X}_n)$, is k-wise independent if for all subsequences $a_{i_1}, \ldots, a_{i_k}$*

where $1 \leq i_1 < i_2 \cdots < i_k \leq n$*, we have that*

$$\Pr\left[\bigwedge_{j\in[k]} (\mathbf{X}_{i_j} = a_{i_j})\right] = \prod_{j\in[k]} \Pr[\mathbf{X}_{i_j} = a_{i_j}]$$

Informally, a set of random variables is k-wise independent if any subset of size at most k is mutually independent. We often think about a set of k-wise independent random variables whose joint distribution is uniform over its joint support. Such a set of random variables are referred to as k-wise uniform.

When we are dealing with $k = 2-$wise independent random variables, we will often refer to them as *two-wise* or *pairwise*-independent random variables. Pairwise independence appears often in the context of hash function families, where the concept is used to analyze the probability of collisions in buckets.

Example 2 (1-wise independence)**.** *Let's consider the set of random variables,* $\{\mathbf{X}_i\}_{i\in[n]} \in \{0,1\}$*. Here,* $\mathbf{X}_1 = 1$ *with probability* $1/2$*. Furthermore,* $\mathbf{X}_j = \mathbf{X}_1$ *for all* $j \in [n] \setminus \{1\}$*. This set of random variables is* 1*-wise independent.*

Example 3 (2-wise independent)**.** *Let's consider the set of random variables,* $\{\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3\}$*. Here,* $\mathbf{X}_1 = 1$ *with probability* $1/2$*.* $\mathbf{X}_2$ *has the same distribution as* $\mathbf{X}_1$*. Furthermore,* $\mathbf{X}_3 = \mathbf{X}_1 \oplus \mathbf{X}_2$*. This set is pairwise independent.*

2.1 Generating n-pairwise uniform bits with small seed length

In this section, we constructively prove that it is possible to generate n-pairwise uniform bits with seed length $k = \lceil \log(n+1) \rceil$ bits. Here is the construction.

For each non-empty $S \subseteq [k]$, let $\mathbf{b}_1, \dots, \mathbf{b}_k$ be independent, uniform random bits over support $\{0,1\}$. Furthermore, let

$$\mathbf{X}_S = \bigoplus_{i\in S} \mathbf{b}_i$$

Note that there are $n = 2^k - 1$ of such subsets of this type. Now, consider any two, non-empty subsets $S_1, S_2 \subseteq [k], S_1 \neq S_2$. Next, consider any element $(\alpha, \beta) \in \{0,1\}^2$. Without loss of generality, assume that $S_1 \setminus S_2 \neq \emptyset$. With this set up, we can state that

$$\Pr_{\mathbf{b}_1,\dots,\mathbf{b}_k}[\mathbf{X}_{S_1} = \alpha, \mathbf{X}_{S_2} = \beta] = \Pr[\mathbf{X}_{S_2} = \beta] \cdot \Pr[\mathbf{X}_{S_1} = \alpha \mid \mathbf{X}_{S_2} = \beta]$$

We can fix the outcomes of each coin flip except for that of the last outcome of S_2. Let the outcome of the last coin flip be denoted as $\mathbf{b}_j$. We can now say that the single outcome of $\mathbf{b}_j$ causes $\mathbf{X}_{S_2}$ to be either 0 or 1 with equal probability. Thus, $\Pr[\mathbf{X}_{S_2} = \beta] = 1/2$.

Now, to show that $\Pr[\mathbf{X}_{S_1} = \alpha | \mathbf{X}_{S_2} = \beta] = 1/2$, we can note the following. Let $j' \in S_1 \setminus S_2$. We can fix all $\mathbf{b}_i, i \in S_2$ such that $\mathbf{X}_{S_2} = \beta$. We can also fix all $\mathbf{b}_i$ other than $i = j'$ such that $\mathbf{b}_i, i \in S_2$. In the first setting of j', we have that $\mathbf{X}_{S_1} = \alpha$, and in the other setting, we have that $\mathbf{X}_{S_1} \neq \alpha$. By this construction, we have shown that both outcomes are equally likely, and that $\Pr[\mathbf{X}_{s_1} = \alpha \mid \mathbf{X}_{s_2} = \beta] = 1/2$. To that end, we can say that

$$\Pr_{\mathbf{b}_1,\ldots,\mathbf{b}_k}[\mathbf{X}_{S_1} = \alpha, \mathbf{X}_{S_2} = \beta] = \Pr[\mathbf{X}_{S_2} = \beta] \cdot \Pr[\mathbf{X}_{S_1} = \alpha \mid \mathbf{X}_{S_2} = \beta] = 1/4$$

2.2 Derandomization application: MaxCut

The optimization variant of the MaxCut problem is known to be NP-hard. It is described as follows. Given a graph, $G = (V, E)$, as input, find a partition of the vertex set $V = X \cup Y$ such that $X \cap Y = \emptyset$ and the number of edges crossing X to Y is maximized.

There is a well-known randomized algorithm that 1/2-approximates the optimal solution. It is given as follows.

Algorithm 1 MaxCut$(G = (V, E))$

```
X = ∅
for v ∈ V do
    Toss an unbiased coin, and set its value to b
    if b = 1 then
        X = X ∪ {v}
Y := V − X
return (X, Y)
```

For ease of analysis, let $\mathbf{E} = (X, Y)$ be the set of edges that crosses from a vertex $u \in X$ to leads to a vertex $u \in Y$. Furthermore, let OPT be the size of the optimal cut. Note that we can compute the following expectation.

$$\mathbb{E}[|\mathbf{E}(X,Y)|] = \sum_{\{u,v\}\in E} \Pr[\{u,v\} \text{ crosses cut}]$$

Since each edge has a probability of 1/2 of crossing the cut, due to linearity of expectation, we can say that

$$= \sum_{e\in E} \frac{1}{2} = \frac{1}{2} \cdot |\mathbf{E}| \geq \frac{1}{2} \cdot OPT$$

2.3 Constructing General k-wise Uniformly Random Variables

Our goal is to construct some general k-wise uniform random variables $\mathbf{X} = (\mathbf{X}_1, \ldots, \mathbf{X}_n)$ over the finite field $\mathbb{F}$ that contains n elements. Pick k variables $\mathbf{c}_0, \ldots, \mathbf{c}_{k-1}$ independently and uniformly from $\mathbb{F}$, requiring $k \log n$ bits of randomness. We want to view these $\mathbf{c}_i$ as coefficients of a univariate polynomial over $\mathbb{F}$. That is, for $\alpha \in \mathbb{F}$, let

$$\mathbf{X}_\alpha := \sum_{i=0}^{k-1} \mathbf{c}_i \alpha^i = p_c(\alpha).$$

Claim 4. *$\mathbf{X}_\alpha$ is a k-wise uniform random variable over $\mathbb{F}$.*

Proof. We get k-wise independence by using the Lagrange interpolation, which states that for any desired $\{a_i\}_{i\in[k]} \in \mathbb{F}$ and for any distinct $\{\alpha_i\}_{i\in[k]} \in \mathbb{F}$ there exists a unique set of coefficients $(c_0, \ldots, c_{k-1})$ such that

$$\mathbf{X}_\alpha = p_c(\alpha) = \sum_{i=1}^{k} a_i \cdot \frac{\prod_{j\neq i}(\alpha - \alpha_j)}{\prod_{j\neq i}(\alpha_i - \alpha_j)},$$

with

$$X_{\alpha_i} = p(\alpha_i) = a_i$$

for all $i = 1, \ldots, k$. We then have that our uniform distribution over $(\mathbf{c}_0, \ldots, \mathbf{c}_{k-1})$ induces uniformity over $(a_1, \ldots, a_k)$, meaning

$$\Pr\left[\bigwedge_{i\in[k]} (\mathbf{X}_{\alpha_i} = a_i)\right] = \prod_{i=1}^{k} \Pr[\mathbf{X}_{\alpha_i} = a_i] = \frac{1}{|\mathbb{F}|^k}.$$

Hence, $\mathbf{X}_\alpha$ is a k-wise uniform random variable over $\mathbb{F}$ of size n. ∎

2.4 Official Homework Problem

Let $\mathbb{F}$ be a field with $|\mathbb{F}| = n = 2^j$ and let $i \leq j$. Show how to generate n elements $\mathbf{X_1}, \ldots, \mathbf{X_n}$ of $\{0, 1, \ldots, 2^i - 1\}$ which are k-wise uniform using kj independent uniform random bits. Use the construction presented in the previous section to help you do so.

2.5 Derandomization Application: Max3SAT

Max3SAT is another well known NP-Hard optimization problem. Given a 3CNF instance $\phi = \bigwedge_{i=1}^{m} C_i$, where each $C_i = (l_{i_1} \vee l_{i_2} \vee l_{i_3})$, find an assignment that satisfies a maximal number of clauses. Here, there are m clauses and n variables.

Here's a simple randomized algorithm that 7/8-approximates the optimal solution.

Algorithm 2 Max3SAT(ϕ)

Let $b = \perp^n$ be a string, where $\perp$ is a special character.
for $i \in [n]$ **do**
  $b_i = 1$ with probability $1/2$, and $b_i = 0$ otherwise.
return b

Let $\mathbf{S}$ be a non-negative random variable that counts the number of satisfied clauses. Furthermore, let $C_i(\mathbf{b})$ be a boolean value denoting whether or not C_i is satisfied by assignment $\mathbf{b} \sim \{0,1\}^n$ uniformly. Finally, let OPT denote the maximal number of satisfiable clauses. Given this algorithm, we can say that

$$\mathbb{E}[\mathbf{S}] = \sum_{i=1}^{m} Pr[C_i(\mathbf{b}) = 1] = \frac{7}{8} \cdot m \geq \frac{7}{8} \cdot OPT$$

Say $\mathbf{b}$ was only 3-wise independent. If we can enumerate over all

$$2^{3 \cdot \log n} = \text{poly}(n)$$

strings (as the seed length of our PRG is $k \log n$), we now have a derandomization assumption if we can choose the $\mathbf{b}$ that satisfies the most clauses, that is,

$$\sum_{i=1}^{m} \Pr[C_i(\mathbf{b}) = 1] \geq \frac{7}{8} \cdot m$$

2.6 PRG-type Applications: Fooling Juntas and Decision Trees

Definition 5 (k-junta)**.** *A k-junta over $\{0,1\}^n$ is a function f such that*

$$f(x_1, \ldots, x_n) = g(x_{i_1}, \ldots, x_{i_k})$$

for some other function g and indices $i_1 < \cdots < i_k$.

Definition 6 ($\mathcal{J}_k$)**.**

$$\mathcal{J}_k = \{f \mid f : \{0,1\}^n \rightarrow \{0,1\}, f \text{ is } k\text{-junta}\}$$

Definition 7 ($\mathcal{DT}_k$)**.**

$$\mathcal{DT}_k = \{f \mid f : \{0,1\}^n \rightarrow \{0,1\}, f \text{ is computed by } k\text{-depth Decision Tree}\}$$

Observation 8. *The class of all k-juntas is a strict subset of the class of all k-depth decision trees. That is,*

$$\mathcal{J}_k \subsetneq \mathcal{DT}_k.$$

Corollary 9. *If $\mathbf{X}$ is constructed to be k-wise independent over $\{0,1\}^n$, then $\mathbf{X}$ perfectly fools $\mathcal{J}_k$.*

Proof. Note that the seed length of the PRG used to construct X has $k \log n$. Since a k-junta is essentially a function over k variables, we can capture the scope of all k-juntas by iterating through all seeds. We then output the same values when computing $f(\mathbf{X})$ or $f(\mathbf{U})$. Hence,

$$\mathbb{E}[f(\mathbf{X})] = \mathbb{E}[f(\mathbf{U})],$$

meaning X 0-fools, or perfectly fools, all $f \in \mathcal{J}_k$. ■

Lemma 10. *Since $\mathbf{X}$ perfectly fools $\mathcal{J}_k$, then $\mathbf{X}$ perfectly fools $\mathcal{DT}_k$.*

Proof. Let $f \in \mathcal{DT}_k$. Let L be the set of all 1-leaves of the decision tree that computes f. Then, define f in terms of many f_ℓ:

$$f = \sum_{\ell \in L} f_\ell.$$

Each f_ℓ is also a k-junta as they are at most a conjunction of k variables. Recall the definition of X fooling $\mathcal{J}_k$, with $f_\ell \in \mathcal{J}_k$:

$$\mathbb{E}[f_\ell(\mathbf{X})] = \mathbb{E}[f_\ell(\mathbf{U})].$$

Then,

$$\begin{aligned}
\mathbb{E}[f(\mathbf{X})] &= \mathbb{E}\left[\sum_{\ell \in L} f_\ell(\mathbf{X})\right] \\
\textit{(linearity of expectation)} &= \sum_{\ell \in L} \mathbb{E}[f_\ell(\mathbf{X})] \\
\textit{(0-fooling of } \mathcal{J}_k\textit{)} &= \sum_{\ell \in L} \mathbb{E}[f_\ell(\mathbf{U})] \\
&= \mathbb{E}\left[\sum_{\ell \in L} f_\ell(\mathbf{U})\right] \\
&= \mathbb{E}[f(\mathbf{U})]
\end{aligned}$$

As $f \in \mathcal{DT}_k$, $\mathbf{X}$ then also 0-fools $\mathcal{DT}_k$. ■

We will generalize this result to achieve the triangle inequality.

Lemma 11 (Triangle Inequality)**.** *Let* $f_1, \ldots, f_t : \{0,1\} \to \mathbb{R}$ *and* $\lambda_0, \ldots, \lambda_t \in \mathbb{R}$. *Define*

$$f := \lambda_0 + \sum_{i=1}^{t} \lambda_i f_i(x).$$

If random variable $\mathbf{X}$ ϵ_i*-fools each* f_i *for all* $i \in [t]$*, then* $\mathbf{X}$ *also* ϵ*-fools* f *for*

$$\epsilon = \sum_{i=1}^{t} |\lambda_i| \epsilon_i.$$

Proof. This is a relatively straightforward proof. We first expand out the definition of f, then apply triangle inequality under the usual metric. Note that if $\mathbf{X}$ ϵ_i-fools all f_i,

$$|\mathbb{E}[f_i(\mathbf{X})] - \mathbb{E}[f_i(\mathbf{U})]| \leq \epsilon_i.$$

So,

$$\begin{aligned}
|\mathbb{E}[f(\mathbf{X})] - \mathbb{E}[f(\mathbf{U})]| &= \left|\sum_{i=1}^{t} \lambda_i \mathbb{E}[f_i(\mathbf{X})] - \sum_{i=1}^{t} \lambda_i \mathbb{E}[f_i(\mathbf{U})]\right| \\
&= \left|\sum_{i=1}^{t} \lambda_i (\mathbb{E}[f_i(\mathbf{X})] - \mathbb{E}[f_i(\mathbf{U})])\right| \\
\textit{(triagle inequality)} &\leq \sum_{i=1}^{t} |\lambda_i| \cdot |\mathbb{E}[f_i(\mathbf{X})] - \mathbb{E}[f_i(\mathbf{U})]| \\
\textit{(}\epsilon_i \textit{ fooling of } f_i\textit{)} &\leq \sum_{i=1}^{t} |\lambda_i| \epsilon_i
\end{aligned}$$

Now define

$$\epsilon = \sum_{i=1}^{t} |\lambda_i| \epsilon_i,$$

implying that

$$|\mathbb{E}[f(\mathbf{X})] - \mathbb{E}[f(\mathbf{U})]| \leq \epsilon,$$

hence showing **X** does indeed ϵ-fool f if it also ϵ_i-fools all f_i. ■

In summary, constructing k-wise independent variables lead to PRG-type applications of perfectly fooling k-juntas and k-depth decision trees. Upon generalizing the realization that a depth k decision tree is the sum of many k-juntas, we come up with a general application of PRGs. This application states that if a certain group of functions are ϵ_i-fooled by some k-wise independent random variable **X**, then we can also ϵ-fool any linear combination of these functions for some ϵ.

3 ϵ-bias Distributions

Definition 12 (PAR)**.** *$PAR = \{p_S\}_{S \subseteq [n]}$, where*

$$\begin{aligned}
p_s &= \sum_{i \in S} x_i \textit{ mod } 2 \\
&= \bigoplus_{i \in S} x_i.
\end{aligned}$$

In other words, PAR is the class of all parities over $x_1, \ldots, x_n \in \{0, 1\}^n$.

Also, let $\chi_S(x) = (-1)^{p_S(x)} = e(p_S(x))$ be the character function, outputting in $\{\pm 1\}$.

Definition 13 (ϵ-biased)**.** *We call Random Variable $\mathbf{X} = (\mathbf{X}_1, \ldots, \mathbf{X}_n)$ **ϵ-biased** if it $\epsilon/2$-fools PAR, hence ϵ-fooling all characters χ_S.*

$$|\mathbb{E}[\chi_S(\mathbf{X})] - \mathbb{E}[\chi_S(\mathbf{U})]| \leq \epsilon,$$

and

$$\mathbb{E}[\chi_S(\mathbf{U})] = \begin{cases} 0 & \text{if } S \neq \emptyset \\ 1 & \text{if } S = \emptyset \end{cases}$$

We will use ϵ-biased Random Variables to fool DEG_d, the class of $\mathbb{F}_2$ polynomials with degree d.

3.1 Constructions of ϵ-biased Random Variables

We begin by generalizing the field extension mentioned in the previous lecture from $\mathbb{F}_4$ to $\mathbb{F}_{2^\ell}$.

Definition 14 ($\mathbb{F}_{2^\ell}$)**.** *Up to an isomorphism, $\mathbb{F}_{2^\ell}$ is an unique field of size 2^ℓ. Recall that $\mathbb{F}_{2^\ell}$ is constructed as the extension field of $\mathbb{F}_2$ modulo $p(t)$, where is an irreducible $\mathbb{F}_2$ polynomial of degree ℓ. That is,*

$$\mathbb{F}_{2^\ell} = \frac{\mathbb{F}_2[t]}{p(t)}.$$

Fact 15. *$\mathbb{F}_{2^\ell}$ contains 2^ℓ elements, and any polynomial of degree at least ℓ can be reduced to one of these elements.*

Our standard construction of $(\mathbb{F}_2)^\ell$ is an ℓ-length vector of $\mathbb{F}_2$ values. Let $bij : \mathbb{F}_{2^\ell} \rightarrow (\mathbb{F}_2)^\ell$ be a linear bijection, meaning

$$bij(x + y) = bij(x) + bij(y).$$

For $\ell = \log(n/\epsilon)$, define a generator $G : (\mathbb{F}_{2^\ell})^2 \rightarrow \{0, 1\}^n$ as

$$G(x, y) = (r_0, \ldots, r_{n-1}), r_i = \langle bij(y), bij(x^i)\rangle \bmod 2.$$

Here, x^i refers to the ith power of $x \in \mathbb{F}_{2^\ell}$, so $x^i \in \mathbb{F}_{2^\ell}$, $bij(x^i), bij(y) \in (\mathbb{F}_2)^\ell$, and for $a, b \in (\mathbb{F}_2)^\ell$, the inner product is defined as

$$\langle a, b\rangle = \sum_{i=0}^{l} a_i b_i \bmod 2.$$

Lemma 16. *G as defined above is an ϵ-biased generator. That is for $\ell = \log(n/\epsilon)$, $\mathbf{U} \sim \{0,1\}^{2\ell}$, $G(\mathbf{U})$ is an ϵ-biased random variable.*

Proof. We need to show that G can fool all parities, meaning for all nonzero $\alpha \in \{0,1\}^n$,

$$\left| \Pr_{r\sim G(\mathbf{U})}\left[\sum_{i=0}^{n-1} \alpha_i r_i \equiv 1 \bmod 2\right] - \frac{1}{2}\right| \leq \frac{\epsilon}{2}.$$

From the definition of G,

$$\begin{aligned}
\Pr_{r\sim G(\mathbf{U})}\left[\sum_{i=0}^{n-1} \alpha_i r_i \equiv 1 \bmod 2\right] &= \Pr_{x,y\sim \mathbb{F}_{2^l}}\left[\sum_{i=0}^{n-1} \alpha_i \langle bij(y), bij(x^i)\rangle \equiv 1 \bmod 2\right] \\
&= \Pr_{x,y}\left[\left\langle bij(y), \sum_{i=0}^{n-1} \alpha_i bij(x^i)\right\rangle \equiv 1 \bmod 2\right] \\
&= \Pr_{x,y}\left[\langle bij(y), bij(p_\alpha(x))\rangle \equiv 1 \bmod 2\right],
\end{aligned}$$

where $p_\alpha(x) =$ deg-$(n-1)$ $\mathbb{F}_{2^l}$ polynomial $\sum_{i=0}^{n-1} \alpha_i x^i$. Define event E to occur if

$$\langle bij(y), bij(p_\alpha(x))\rangle \equiv 1 \bmod 2.$$

Now perform casework on $p_\alpha(x)$.

$$\Pr_y[E] = \begin{cases} 0 & \text{if } p_\alpha(x) = 0 \\ \frac{1}{2} & \text{if } p_\alpha(x) \neq 0 \end{cases}$$

Condition using outcomes of $p_\alpha(x)$.

$$\begin{aligned}
\Pr_{r\sim G(\mathbf{U})}\left[\sum_{i=0}^{n-1} \alpha_i r_i \equiv 1 \bmod 2\right] &= \Pr_y[E] \\
&= \Pr_y[E|p_\alpha(x) = 0] \cdot \Pr_x[p_\alpha(x) = 0] + \Pr_y[E|p_\alpha(x) \neq 0] \cdot \Pr_x[p_\alpha(x) \neq 0] \\
&= \frac{1}{2} \cdot \Pr_x[p_\alpha(x) \neq 0]
\end{aligned}$$

as $\Pr_y[E|p_\alpha(x) = 0] = 0, \Pr_y[E|p_\alpha(x) \neq 0] = 1/2$ using the above casework on $p_\alpha(x)$. For a lower bound on $\Pr_x[p_\alpha(x) \neq 0]$, use that p_α is a degree $n-1$ polynomial over $\mathbb{F}_{2^l}$. As $2^l = n/\epsilon$,

$$\begin{aligned}
\Pr_x[p_\alpha(x) \neq 0] &\geq 1 - \frac{n-1}{2^l} \\
&\geq 1 - \epsilon.
\end{aligned}$$

Hence,

$$\Pr_{r\sim G(\mathbf{U})}\left[\sum_{i=0}^{n-1}\alpha_i r_i \equiv 1 \bmod 2\right] \leq \frac{1}{2} - \frac{\epsilon}{2},$$

and

$$\left|\Pr_{r\sim G(\mathbf{U})}\left[\sum_{i=0}^{n-1}\alpha_i r_i \equiv 1 \bmod 2\right] - \frac{1}{2}\right| \leq \frac{\epsilon}{2}.$$

Hence, G is an ϵ-biased generator. ■

3.2 Application to Coding Theory

Distributions over $(\mathbb{F}_2)^n$ that are k-wise independent and ϵ-biased are closely related to linear codes over $\mathbb{F}_2^n$. There is motivation from coding theory to give ϵ-biased distributions that have a smaller seed length than $2\ell = 2\log(n/\epsilon)$.

3.3 Project Topic

The information-theoretic best possible seed length for an ϵ-biased distribution is

$$\log n + 2\log\left(\frac{1}{\epsilon}\right) - \log\log\left(\frac{1}{\epsilon}\right).$$

Can this be achieved with an explicit construction? Ta-Shma '17 showed a bound of

$$\log n + 2\log\left(\frac{1}{\epsilon}\right) + \tilde{O}\left(\log^{2/3}\left(\frac{1}{\epsilon}\right)\right)$$

4 k-wise ϵ-biased Random Variables

Definition 17. *Random Variable $\mathbf{X} = (\mathbf{X}_1, \ldots, \mathbf{X}_n)$ is k-wise ϵ-biased if it ϵ fools all χ_S with $|S| \leq k$.*

Lemma 18. *There exists explicit k-wise ϵ-biased $\mathbf{X} = (\mathbf{X}_1, \ldots, \mathbf{X}_n) = G(\mathbf{U})$ with seed length*

$$\log k + \log\left(\frac{1}{\epsilon}\right) + \log\log n$$

Proof. Let $G : \{0,1\}^s \to \{0,1\}^n$ be a k-wise uniform generator that is a linear transformation viewed as $\mathbb{F}_2^s \to \mathbb{F}_2^n$. Let $\mathbf{Y}$ be an ϵ-biased distribution over $\{0,1\}^s$. Overall, $\mathbf{X} = G(\mathbf{Y})$, with output in $\{0,1\}^n$. Hence, the seed length is

$$2\log\left(\frac{s}{\epsilon}\right) = O\left(\log k + \log\left(\frac{1}{\epsilon}\right) + \log\log n\right)$$

We want to show $G(\mathbf{Y})$ ϵ-fools parities of size at most k. Let $S \subset [n]$, $|S| \leq k$. This parity is

$$p_S(x) = \sum_{i \in S} x_i, x \in (\mathbb{F}_2)^s.$$

Let $M \in \mathbb{F}_2^{n \times s}$ be a matrix acting as a linear transformation for G. Denote M_j to be the jth row of M. Then,

$$G(\mathbf{Y}) = (\langle M_1, \mathbf{Y}\rangle, \ldots, \langle M_n, \mathbf{Y}\rangle) \in \mathbb{F}_2^n.$$

For $\mathbf{Y} \in \mathbb{F}_2^s$, we have

$$\begin{aligned} p_S(G(\mathbf{Y})) &= \sum_{i \in S} \langle M_i, \mathbf{Y}\rangle \\ &= \sum_{i \in S} \sum_{j=1}^{s} M_{ij} \mathbf{Y}_j \\ &= \sum_{j=1}^{s} \left(\sum_{i \in S} M_{ij}\right) \mathbf{Y}_j. \end{aligned}$$

Notice how this is a PAR over $\mathbf{Y} = (\mathbf{Y}_1, \ldots, \mathbf{Y}_s)$. Since $\mathbf{Y}$ is ϵ-biased,

$$|\mathbb{E}[p_S(G(\mathbf{Y})] - \mathbb{E}[p_S(G(\mathbf{U})]| \leq \frac{\epsilon}{2}.$$

Since p_s is a k-junta and G is k-wise uniform, we have

$$\mathbb{E}[p_S(G(\mathbf{U}))] = \mathbb{E}[p_S(G(\mathbf{U}_n))].$$

So,

$$|\mathbb{E}[p_S(G(\mathbf{Y}))] - \mathbb{E}[p_S(G(\mathbf{U}_n)]| \leq \frac{\epsilon}{2},$$

meaning $\mathbf{Y}$ $(\epsilon/2)$-fools p_S. ∎