CS59200-PRS: Pseudorandomness Sep 3, 2025

Lecture 3: k-wise Independence and Fourier Analysis

Lecturer: Wei Zhan *Scribe: Arvind Ramaswami*

1 Error Reduction by k-wise Independence

Suppose we have a randomized algorithm $A(x, r) \in \{0, 1\}$ (r is a random m-bit string) that is correct w.p. $\geq 1/2 + \varepsilon$. We want to reduce the error by repetition: We run $A(x, r_1), \ldots, A(x, r_t)$ with different randomness and take the majority vote of the outputs.

If $r_1, r_2, \ldots, r_t$ are mutually independent, we can use Chebyshev's inequality

$$\Pr[|X - \mathbb{E}[X]| \geq \alpha] \leq \frac{\mathbf{Var}[X]}{\alpha^2}$$

to bound the error rate. Let $X_i = A(x, r_i) \in [0, 1]$, and $X = \frac{1}{t}\sum_i X_i$, then

$$\mathbf{Var}[X] = \frac{1}{t^2}\sum_i \mathbf{Var}[X_i] \leq \frac{1}{t}\left(\frac{1}{4} - \varepsilon^2\right)$$

and the majority vote is only wrong when $|X - \mathbb{E}[X]| \geq \varepsilon$, so the error probability is

$$\Pr[|X - \mathbb{E}[X]| \geq \varepsilon] \leq \frac{1/4 - \varepsilon^2}{t\varepsilon^2} \leq \frac{1}{4t\varepsilon^2}.$$

If we want constant error with independent randomness, we need:

- $O(1/\varepsilon^2)$ repetitions;
- $O(m/\varepsilon^2)$ random bits.

And if we want $1/\text{poly}(n)$ error, we need

- $O(\text{poly}(n)/\varepsilon^2)$ repetitions;
- $O(m \cdot \text{poly}(n)/\varepsilon^2)$ random bits.

Note that by Chernoff bound, we can actually get better bounds for $1/\text{poly}(n)$ error:

- $O(\log(n)/\varepsilon^2)$ repetitions;
- $O(m \cdot \log(n)/\varepsilon^2)$ random bits.

1.1 k-wise Independent Chebyshev

Theorem 1. *If $X_1, \ldots, X_t \in [0,1]$ are $2k$-wise independent, for $X = \frac{1}{t}\sum_{i=1}^{t} X_i$,*

$$\Pr[|X - \mathbb{E}[X]| \geq \varepsilon] \leq \left(\frac{k^2}{t\varepsilon^2}\right)^k.$$

Proof. Consider $(X - \mathbb{E}[X])^{2k}$. Markov gives: $\Pr[(X-\mathbb{E}[X])^{2k} \geq \alpha\, \mathbb{E}[(X-\mathbb{E}[X])^{2k}]] \leq \frac{1}{\alpha}$. We can bound $E[(X - E[X])^{2k}]$ by

$$\begin{aligned}
E[(X-E[X])^{2k}] &= \mathbb{E}\left[\frac{1}{t^{2k}} \sum_{i_1,\ldots,i_{2k}=1}^{t} (X_{i_1} - \mathbb{E}[X_{i_1}]) \cdots (X_{i_{2k}} - \mathbb{E}[X_{i_{2k}}])\right] \\
&= \frac{1}{t^{2k}} \sum_{i_1,\ldots,i_{2k}=1}^{t} \mathbb{E}[(X_{i_1} - \mathbb{E}[X_{i_1}]) \cdots (X_{i_{2k}} - \mathbb{E}[X_{i_{2k}}])] \\
&\leq \frac{1}{t^{2k}} \#\{(i_1, i_2, \ldots, i_{2k}) \in [t]^{2k} : \text{each } i \in [t] \text{ appears 0 or} \geq 2 \text{ times}\} \\
&\leq \frac{1}{t^{2k}} \cdot t^k \cdot k^{2k} = \left(\frac{k^2}{t}\right)^k.
\end{aligned}$$

The third line is because when there exists some $i \in [t]$ that appears in $(i_1, i_2, \ldots, i_{2k})$ exactly once, say $i = i_1$, by $2k$-wise independence we have

$$\mathbb{E}[(X_{i_1} - \mathbb{E}[X_{i_1}]) \cdots (X_{i_{2k}} - \mathbb{E}[X_{i_{2k}}])] = \mathbb{E}[X_{i_1} - \mathbb{E}[X_{i_1}]]\, \mathbb{E}[(X_{i_2} - \mathbb{E}[X_{i_2}]) \cdots (X_{i_{2k}} - \mathbb{E}[X_{i_{2k}}])]$$

which is 0 since $\mathbb{E}[X_{i_1} - \mathbb{E}[X_{i_1}]] = 0$. Each of the rest of the terms in the sum is at most 1.

The fourth line is because within such a $2k$-tuple, there are at most k distinct elements. So we can enumerate such tuples by first choose k elements from $[t]$, and then choose each one of $i_1, \ldots, i_{2k}$ from these k elements. Thus we have

$$\begin{aligned}
\Pr[|X - \mathbb{E}[X]| \geq \varepsilon] &= \Pr[|X - \mathbb{E}[X]|^{2k} \geq \varepsilon^{2k}] \\
&\leq \frac{1}{\epsilon^{2k}} E[(X - \mathbb{E}[X])^{2k}] \leq \left(\frac{k^2}{t\varepsilon^2}\right)^k.
\end{aligned}$$

□

By taking $r_1, \ldots, r_t$ to be $2k$-wise independent (via a $2k$-wise uniform hash function with input length $\log t$ and output length m), we can significantly reduce the number of random bits, especially on the dependence with ε. Notice that now the results $X_i = A(x, r_i)$ are also $2k$-wise uniform, so we can use Theorem 1.

For constant error, by using pairwise independence (k=1), we need:

- $O(1/\varepsilon^2)$ repetitions;
- $O(m + \log(1/\varepsilon))$ random bits, which is much less than independent repetitions.

For $1/\text{poly}(n)$ error, using $k = O(\log n)$-wise independence, we need:

- $t = O(k^2/\varepsilon^2) = O(\log^2 n/\varepsilon^2)$ repetitions;
- $O(\log n \cdot (m + \log(1/\varepsilon) + \log\log n))$ random bits (This is because in the k-wise independent hash function, m is the output length, while $\log t = O(\log(1/\varepsilon) + \log\log n)$ is the input length).

2 What does k-wise independence fool?

- degree-k monomials, by definition.
- degree-k polynomials, by linearity. For instance, the polynomial for MAX-CUT:

$$\sum_{(i,j)\in E} X_i(1 - X_j) + X_j(1 - X_i)$$

- In order to get the most general answer, we will use Fourier analysis.

3 Discrete (Boolean) Fourier Analysis.

Given $g\colon \{0,1\}^n \to \{0,1\}$, we want the **multilinear (polynomial) expansion** of g

$$g(x_1, \dots, x_n) = \sum_{S\subseteq[n]} \alpha_S \prod_{i\in S} x_i, \alpha_S \in \mathbb{R}.$$

To prove such an expansion uniquely exists, we can think of the space of all functions $\{0,1\}^n \to \mathbb{R}$ as a linear space on $\mathbb{R}$ of dimension 2^n, and prove linear independence of all monomials $\prod_{i\in S} x_i$.

It is easier with a change of domain, where we look at functions $f\colon \{\pm1\}^n \to \{\pm1\}$ by defining

$$f(x_1, \dots, x_n) = 2g(1/2 + 1/2x_1, \dots, 1/2 + 1/2x_n) - 1.$$

Notice that f has the same degree as g and keeps the same independence between the input coordinates.

Theorem 2 (Fourier expansion)**.** *For $f\colon \{\pm1\}^n \to \mathbb{R}$, we can uniquely write f as a multilinear polynomial*

$$f(x_1, \dots, x_n) = \sum_{S\subseteq[n]} \widehat{f}(S)\chi_S(x_1, \dots, x_n).$$

Here $\chi_S(x_1, \dots, x_n) = \prod_{i\in S} x_i$ is called the characteristic function *on S, and $\widehat{f} : 2^{[n]} \to \mathbb{R}$ gives the* Fourier coefficients *of f.*

To prove the existence and uniqueness, we equip the linear space of all functions $\{\pm 1\}^n \to \mathbb{R}$ with an inner product:

$$\langle f, g \rangle = \mathop{\mathbb{E}}_{X \sim \{\pm 1\}^n} [f(X) \cdot g(X)].$$

Then it suffices to note the following facts.

Fact 1. $\{\chi_S\}$ *forms an orthonormal basis.*

Fact 2. *(Fourier duality)*

$$\widehat{f}(S) = \langle f, \chi_S \rangle = \frac{1}{2^n} \sum_x f(x) \chi_S(x)$$

Fact 3. *(Parseval's identity)*

$$\langle f, g \rangle = \sum_S \widehat{f}(S) \widehat{g}(S)$$

Proof of Fact 3.

$$\begin{aligned}
\langle f, g \rangle &= \mathop{\mathbb{E}}_x \left[\sum_{S_1, S_2} \widehat{f}(S_1) \widehat{g}(S_2) \chi_{S_1}(x) \chi_{S_2}(x) \right] \\
&= \sum_{S_1, S_2} \widehat{f}(S_1) \widehat{g}(S_2) \mathop{\mathbb{E}}_x [\chi_{S_1}(x) \chi_{S_2}(x)]
\end{aligned}$$

where $\mathop{\mathbb{E}}_x[\chi_{S_1}(x)\chi_{S_2}(x)] = \langle \chi_{S_1}, \chi_{S_2} \rangle$ is 1 when $S_1 = S_2$ and 0 otherwise. □

3.1 k-wise Uniformity and Fourier Analysis

We can give a Fourier characterization of k-wise uniformity as follows.

Theorem 3. $p : \{\pm 1\}^n \to \mathbb{R}$ *is a k-wise uniform distribution if and only if* $\widehat{p}(S) = 0$ *for all* $1 \le |S| \le k$ *(note that* $\widehat{p}(\varnothing) = 2^{-n}$*).*

Proof. ($\Longrightarrow$):

$$\begin{aligned}
\widehat{p}(S) &= \frac{1}{2^n} \sum_x p(x) \chi_S(x) \\
&= \frac{1}{2^n} \mathop{\mathbb{E}}_{x \sim p} [\chi_S(x)] \\
&= \frac{1}{2^n} \mathop{\mathbb{E}}_{x \in \{\pm 1\}^n} [\chi_S(x)] = 0 \text{ (since } p \text{ fools degree } k \text{ polynomials)}
\end{aligned}$$

($\Longleftarrow$): For $(b_1, \ldots, b_n) \in \{\pm 1\}^n$, write $b_S = \prod_{i \in S} b_i$ and we have

$$
\begin{aligned}
\Pr_{X\sim p}[X_{i_1}=b_{i_1},\dots,X_{i_k}=b_{i_k}] &= \sum_x p(x)\mathbb{1}_{x_{i_1}=b_{i_1}}\cdot\dots\cdot\mathbb{1}_{x_{i_k}=b_{i_k}} \\
&= \sum_x p(x)(1+x_{i_1}b_{i_1})\cdot\dots\cdot(1+x_{i_k}b_{i_k})\cdot\frac{1}{2^k} \\
&= \sum_x p(x)\cdot\sum_{S\subseteq\{i_1,\dots,i_k\}}\chi_S(x)b_S\cdot\frac{1}{2^k} \\
&= \frac{1}{2^k}\sum_{S\subseteq\{i_1,\dots,i_k\}} b_S\sum_x p(x)\chi_S(x) \\
&= \frac{1}{2^k}\sum_{S\subseteq\{i_1,\dots,i_k\}} b_S\cdot 2^n\cdot\widehat{p}(S) \\
&= \frac{1}{2^k}\cdot b_\varnothing\cdot 2^n\cdot\widehat{p}(\varnothing) \\
&= \frac{1}{2^k}.
\end{aligned}
$$

□

A natural question to ask is: Which functions $f:\{\pm1\}^n\to\mathbb{R}$ are ε-fooled by all k-wise independent distributions, i.e.

$$\left|\mathop{\mathbb{E}}_{X\sim\{\pm1\}^n}[f(X)]-\mathop{\mathbb{E}}_{X\sim p}[f(X)]\right|\le\varepsilon?$$

Here we give a partial answer with Fourier analysis. Notice that the left term equals

$$\frac{1}{2^n}\sum_x f(x)=\widehat{f}(\varnothing).$$

while the right term equals

$$
\begin{aligned}
\sum_x p(x)f(x) &= 2^n\langle p,f\rangle \\
&= 2^n\sum_S\widehat{p}(S)\widehat{f}(S) \\
&= \widehat{f}(\varnothing)+2^n\sum_{S\neq\varnothing}\widehat{p}(S)\widehat{f}(S).
\end{aligned}
$$

Thus, f is ε-fooled by $p \iff \left|\sum_{S\neq\varnothing}\widehat{p}(S)\widehat{f}(S)\right|\le 2^{-n}\cdot\varepsilon$. If p is k-wise independent, the sum is equal to $\left|\sum_{|S|\ge k+1}\widehat{p}(S)\widehat{f}(S)\right|$.

Since p is a distribution, $|\widehat{p}(S)|\le\frac{1}{2^n}$, and thus

$$\left|\sum_{|S|\ge k+1}\widehat{p}(S)\widehat{f}(S)\right|\le\left|\sum_{|S|\ge k+1}\widehat{f}(S)\right|\cdot\frac{1}{2^n}.$$

Therefore, if $\left|\sum_{|S|\geq k+1} \widehat{f}(S)\right| \leq \varepsilon$, then f is ε-fooled by all k-wise uniform distributions. The sum $\left|\sum_{|S|\geq k+1} \widehat{f}(S)\right|$ is called the ℓ_1 Fourier tail.

Proving bounds on the Fourier tail is an active research problem. Most of the time, bounding the ℓ_1 Fourier tail by a small ε is too much to ask for (notice how we simply relaxed $|\widehat{p}(S)|$ to $\frac{1}{2^n}$ which is often a huge loss), and instead bounding the ℓ_2 tail

$$\sum_{|S|\geq k+1} \widehat{f}^2(S)$$

is more achievable and still suffices. For further reading, see e.g. the following works on Fourier tails of constant depth circuits

- Nathan Linial, Yishay Mansour, and Noam Nisan. *Constant depth circuits, Fourier transform, and learnability.*
- Mark Braverman. *Polylogarithmic independence fools* AC^0 *circuits.*
- Avishay Tal. *Tight bounds on the Fourier spectrum of* AC^0.