Lecture 35: Min-Entropy Extraction via Small-bias Masking
Recall

- For a probability distribution $\mathbb{X}$ over $\{0, 1\}^n$, we defined the bias of $\mathbb{X}$ with respect to a linear test $S \in \{0, 1\}^n$ as follows:

$$\text{bias}_{\mathbb{X}}(S) = \mathbb{P}[S \cdot \mathbb{X} = 0] - \mathbb{P}[S \cdot \mathbb{X} = 1]$$

- The probability that two independent samples from $\mathbb{X}$ and $\mathbb{Y}$ turn out to be identical is defined as:

$$\text{col}(\mathbb{X}, \mathbb{Y}) = \frac{1}{N} \sum_{S \in \{0, 1\}^n} \text{bias}_{\mathbb{X}}(S)\text{bias}_{\mathbb{Y}}(S)$$

- $\mathbb{X} \oplus \mathbb{Y}$ is a probability distribution over $\{0, 1\}^n$ such that $\mathbb{P}[\mathbb{X} \oplus \mathbb{Y} = z]$ is the probability that two samples according to $\mathbb{X}$ and $\mathbb{Y}$ add up to $z$:

$$\text{bias}_{\mathbb{X} \oplus \mathbb{Y}} = \text{bias}_{\mathbb{X}}\text{bias}_{\mathbb{Y}}$$
The statistical distance between two probability distributions $X$ and $Y$ over the sample space $\{0, 1\}^n$ is

$$2SD(X, Y) = \sum_{x \in \{0, 1\}^n} |P[X = x] - P[Y = x]|$$

We showed that

$$2SD(X, Y) \leq \ell_2(bias_X - bias_Y)$$
Example 1

- Let $\mathbb{U}$ represent the uniform distribution over the sample space $\{0, 1\}^n$.

- Note that, we have

$$
\text{bias}_\mathbb{U}(S) = \begin{cases} 
1, & \text{if } S = 0 \\
0, & \text{if } S \neq 0
\end{cases}
$$

- In fact, $\text{bias}_X(0) = 1$ for all probability distributions $X$.  

Min-Entropy Extraction
Let $\mathbb{U}_{\langle v \rangle}$, for $v \in \{0, 1\}^n$, represent the uniform distribution over the vector space spanned by $\{v\}$, i.e., the set $\{0, v\}$.

Let $\mathbb{U}_{\langle w \rangle}$, for $w \in \{0, 1\}^n$, represent the uniform distribution over the vector space spanned by $\{w\}$, i.e., the set $\{0, w\}$.

Prove: $\mathbb{U}_{\langle v \rangle} \oplus \mathbb{U}_{\langle w \rangle} = \mathbb{U}_{\langle v, w \rangle}$.

Here, $\mathbb{U}_{\langle v, w \rangle}$ represents the uniform distribution over the set spanned by $\{v, w\}$. If $v = w$, then $\langle v, w \rangle = \{0, v\}$; otherwise $\langle v, w \rangle = \{0, v, w, v + w\}$.

In general, for linearly independent vectors $v_1, v_2, \ldots, v_k \in \{0, 1\}^n$, we have

$$\mathbb{U}_{\langle v_1, \ldots, v_k \rangle} = \mathbb{U}_{\langle v_1 \rangle} \oplus \cdots \oplus \mathbb{U}_{\langle v_k \rangle}$$

So, we conclude that

$$\text{bias}_{\mathbb{U}_{\langle v_1, \ldots, v_k \rangle}} = \text{bias}_{\mathbb{U}_{\langle v_1 \rangle}} \cdots \text{bias}_{\mathbb{U}_{\langle v_k \rangle}}$$
Example 2

- Prove: There exists a subset $T \subseteq \{0, 1\}^n$ of size $2^{n-1}$ such that $\text{bias}_{U_{\langle \nu \rangle}}(S) = 1$ if $S \in T$; otherwise $\text{bias}_{U_{\langle \nu \rangle}}(S) = 0$.
- Think: Which $S$ have $\text{bias}_{U_{\langle \nu \rangle} \oplus U_{\langle w \rangle}}(S) = 0$?
Let $\mathbb{X}$ be a distribution over the sample space $\{0, 1\}^n$.

We say that the distribution $\mathbb{X}$ has min-entropy at least $k$ if it satisfies the following condition. For any $x \in \{0, 1\}^n$, we have

\[ P[\mathbb{X} = x] \leq \frac{1}{2^k} =: \frac{1}{K} \]

This constraint is succinctly represented as $H_\infty(\mathbb{X}) \geq k$.

Intuition: The probability of any element according to the distribution $\mathbb{X}$ is small. So, the outcome of $\mathbb{X}$ is “highly unpredictable.” Furthermore, $\mathbb{X}$ associates non-zero probability to at least $K$ elements in $\{0, 1\}^n$. 

Min-Entropy Extraction
We had seen that the collision probability of a high min-entropy distribution is low.

\[
\text{col}(X, X) = \sum_{x \in \{0,1\}^n} P[X = x]^2 \leq \sum_{x \in \{0,1\}^n} P[X = x] \frac{1}{K} = \frac{1}{K}
\]

This implies that

\[
\sum_{S \in \{0,1\}^n} \text{bias}_X(S)^2 \leq \frac{N}{K}
\]

Or, equivalently, we write

\[
\sum_{S \in \{0,1\}^n : S \neq 0} \text{bias}_X(S)^2 \leq \frac{N}{K} - 1
\]
Recall: Min-Entropy Sources

Succinctly, we write

\[ \ell^*_2(\text{bias}_X) \leq \sqrt{\frac{N}{K}} - 1 \]

Here \( \ell^*_2(f) \) is identical to the definition of \( \ell_2(f) \) except that it excludes \( f(0)^2 \) in the sum.
Let $\mathcal{Y}$ be a distribution over $\{0, 1\}^n$.

We say that $\mathcal{Y}$ is a small-bias distribution if

$$\text{bias}_\mathcal{Y}(S) \leq \varepsilon$$

for all $0 \neq S \in \{0, 1\}^n$.

Prove: A random probability distribution is a small-bias distribution with very high probability.
Let $X$ be a min-entropy source with $H_\infty(X) \geq K$.

Let $Y$ be a small bias distribution such that $\text{bias}_Y(S) \leq \varepsilon$, for all $0 \neq S \in \{0, 1\}^n$.

We want to say that $X \oplus Y$ is very close to the uniform distribution $U$ over the sample space $\{0, 1\}^n$.

$$2\text{SD} (X \oplus Y, U) \leq \ell_2(\text{bias}_{X \oplus Y} - \text{bias}_U)$$

$$= \ell_2^*(\text{bias}_{X \oplus Y} - \text{bias}_U)$$

$$= \ell_2^*(\text{bias}_{X \oplus Y})$$

$$= \ell_2^*(\text{bias}_X \text{bias}_Y)$$

$$\leq \varepsilon \ell_2^*(\text{bias}_X)$$

$$\leq \varepsilon \sqrt{\frac{N}{K}} - 1$$