Lecture 21: 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

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

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

$$\mathrm{Col}(\mathbb{X}, \mathbb{Y}) = \frac{1}{N} \sum_{S \in \{0,1\}^n} \mathrm{Bias}_{\mathbb{X}}(S)\mathrm{Bias}_{\mathbb{Y}}(S)$$

- $\mathbb{X} \oplus \mathbb{Y}$ 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

$$\mathrm{Bias}_{\mathbb{X}\oplus\mathbb{Y}} = \mathrm{Bias}_{\mathbb{X}} \cdot \mathrm{Bias}_{\mathbb{Y}}$$

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

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

We showed that

$$2\mathrm{SD}\left(\mathbb{X},\mathbb{Y}\right) \leqslant \ell_2(\mathrm{Bias}_{\mathbb{X}} - \mathrm{Bias}_{\mathbb{Y}})$$

Example 1

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

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

- In fact, $\mathrm{Bias}_{\mathbb{X}}(0) = 1$ for all probability distributions $\mathbb{X}$

Example 2

- 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

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

- Prove: There exists a subset $T \subseteq \{0,1\}^n$ of size 2^{n-1} such that $\mathrm{Bias}_{\mathbb{U}_{\langle v \rangle}}(S) = 1$ if $S \in T$; otherwise $\mathrm{Bias}_{\mathbb{U}_{\langle v \rangle}}(S) = 0$.
- Think: Which S have $\mathrm{Bias}_{\mathbb{U}_{\langle v \rangle} \oplus \mathbb{U}_{\langle w \rangle}}(S) = 0$?

Recall: Min-Entropy Sources

- 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

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

  This constraint is succinctly represented as $\mathrm{H}_\infty(\mathbb{X}) \geqslant k$
- 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 $\{0,1\}^n$.

Recall: Min-Entropy Sources II

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

$$\mathrm{Col}(\mathbb{X},\mathbb{X}) = \sum_{x\in\{0,1\}^n} \mathbb{P}\left[\mathbb{X}=x\right]^2 \leqslant \sum_{x\in\{0,1\}^n} \mathbb{P}\left[\mathbb{X}=x\right]\frac{1}{K} = \frac{1}{K}$$

This implies that

$$\sum_{S\in\{0,1\}^n} \mathrm{Bias}_{\mathbb{X}}(S)^2 \leqslant \frac{N}{K}$$

Or, equivalently, we write

$$\sum_{S\in\{0,1\}^n:\ S\neq 0} \mathrm{Bias}_{\mathbb{X}}(S)^2 \leqslant \frac{N}{K} - 1$$

Succinctly, we write

$$\ell_2^*(\mathrm{Bias}_{\mathbb{X}}) \leqslant \sqrt{\frac{N}{K} - 1}$$

Here $\ell_2^*(f)$ is identical to the definition of $\ell_2(f)$ except that it excludes $f(0)^2$ in the sum

- Let $\mathbb{Y}$ be a distribution over $\{0,1\}^n$
- We say that $\mathbb{Y}$ is a small-bias distribution if

$$\mathrm{Bias}_{\mathbb{Y}}(S) \leqslant \varepsilon$$

  for all $0 \neq S \in \{0,1\}^n$
- Prove: A random probability distribution is a small-bias distribution with very high probability

Min-Entropy Extraction via Small-bias Masking

- Let $\mathbb{X}$ be a min-entropy source with $\mathrm{H}_\infty(\mathbb{X}) \geqslant k$
- Let $\mathbb{Y}$ be a small bias distribution such that $\mathrm{Bias}_{\mathbb{Y}}(S) \leqslant \varepsilon$, for all $0 \neq S \in \{0,1\}^n$
- We want to say that $\mathbb{X} \oplus \mathbb{Y}$ is very close to the uniform distribution $\mathbb{U}$ over the sample space $\{0,1\}^n$.

$$\begin{aligned}
2\mathrm{SD}\left(\mathbb{X} \oplus \mathbb{Y}, \mathbb{U}\right) &\leqslant \ell_2(\mathrm{Bias}_{\mathbb{X}\oplus\mathbb{Y}} - \mathrm{Bias}_{\mathbb{U}}) \\
&= \ell_2^*(\mathrm{Bias}_{\mathbb{X}\oplus\mathbb{Y}} - \mathrm{Bias}_{\mathbb{U}}) \\
&= \ell_2^*(\mathrm{Bias}_{\mathbb{X}\oplus\mathbb{Y}}) \\
&= \ell_2^*(\mathrm{Bias}_{\mathbb{X}}\mathrm{Bias}_{\mathbb{Y}}) \\
&\leqslant \varepsilon \ell_2^*(\mathrm{Bias}_{\mathbb{X}}) \\
&\leqslant \varepsilon\sqrt{\frac{N}{K} - 1}
\end{aligned}$$