Strong XOR Lemma for Information Complexity

Pachara Sawettamalya, Huacheng Yu

Princeton University

Two-Player Communication Model

Two player, Alice and Bob, holds a private input:

- Alice has $x \in X$ and Bob has $y \in Y$
- Players jointly compute $f(x,y) \in \{0,1\}$
- This is done via exchanging a sequence of messages M; we call $\pi = (M,X,Y)$ a *protocol*.
- Information cost of a protocol $\pi = (M,X,Y)$ is

$$IC(\pi) = I(M : X \mid Y) + I(M : Y \mid X).$$

  This measures amount of information which the protocol reveals about players' inputs

- Information cost of a function f w.r.t. error δ is

$$\mathrm{I}(f,\delta) = \min_{\pi \text{ computes } f \text{ w.p. } 1-\delta} IC(\pi)$$

XOR Lemma for Information Complexity

Denote $f^{\oplus n}: X^n \times Y^n \longrightarrow \{0,1\}$ to be an n-folded XOR of f such that:

$$f^{\oplus n}(x_1y_1, \dots, x_ny_n) := f(x_1,y_1) \oplus \cdots \oplus f(x_n,y_n)$$

Naïve protocol for $f^{\oplus n}$: compute each $f(x_i,y_i)$ in parallel – this costs n times the cost of computing f.

Strong XOR Lemma asks: For what notions of "cost" and error parameters (ρ,ρ') that the naïve protocol is optimal:

$$\mathrm{Cost}\left(f^{\oplus n},\rho'\right) \geq \Omega(n) \cdot \mathrm{Cost}(f,\rho).$$

The error tradeoffs of ρ vs. ρ' is tight when $(\rho,\rho') = \left(\frac{1}{2}+\frac{\alpha}{2},\frac{1}{2}+\frac{\alpha^n}{2}\right)$ for some advantage $\alpha \in [0,1]$. When the "cost" is information:

- [BBCR'10] True for $(\rho,\rho') = \left(\frac{2}{3},\frac{2}{3}\right)$ but not optimal
- False for $(\rho,\rho') = \left(\frac{2}{3},\frac{1}{2}+2^{-n}\right)$
- [This work] True for $(\rho,\rho') = \left(1-n^{-0.1},\frac{2}{3}\right)$ and asymptotically optimal

Main Results

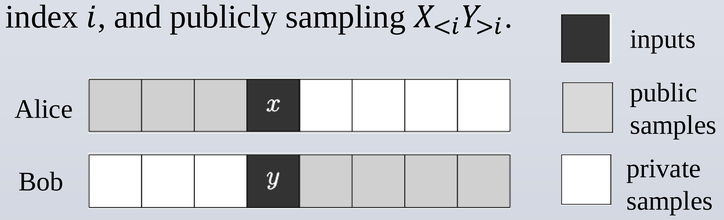
Theorem: For any function $f: X \times Y \longrightarrow \{0,1\}$ and positive integer n, we have:

$$\mathrm{I}\left(f^{\oplus n},\frac{1}{3}\right) \geq \Omega(n) \cdot \left(\mathrm{I}\left(f,\frac{1}{\mathrm{poly}(n)}\right) - o_n(1) - 1\right).$$

Prior Works by [BBCR'10] and [Yu'22]

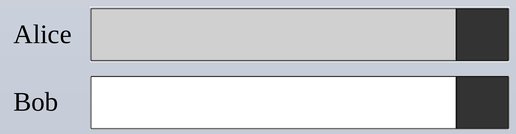
Barak, Braverman, Chen, and Rao (STOC'10) proved an XOR lemma for information in a regime of $\rho = \rho'$.

Let π be a protocol for computing $f^{\oplus n}$ over input distribution μ^n. The protocol π' for computing f is obtained by embedding inputs (x,y) into a random index i, and publicly sampling $X_{<i}Y_{>i}$.

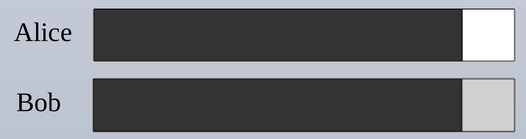


Yu (FOCS'22) gave an alternative view of [BBCR'10] argument by iteratively splitting a protocol π for computing $f^{\oplus n}$ over input dist μ^n into two protocols:

- $\pi^{(n)}$ computes f over input distribution μ



- $\pi^{(<n)}$ computes $f^{\oplus n-1}$over input distribution μ^{n-1}



Lemma: $IC(\pi^{(n)}) + IC(\pi^{(<n)}) = IC(\pi) + O(1)$.

Lemma: There are RVs A,B,Z s.t. $Z = AB$ and

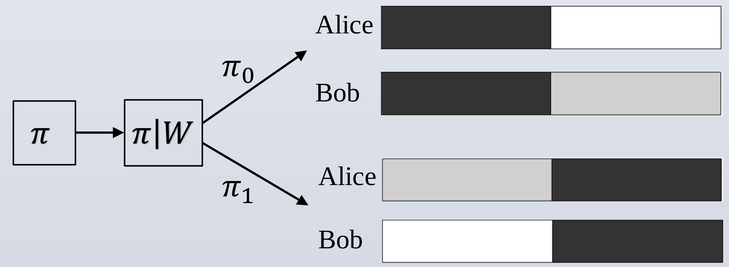
- $adv\left(\pi^{(n)}\right) = \mathbb{E}[A]$
- $adv(\pi) = \mathbb{E}[Z] = \mathbb{E}[AB]$
- $adv\left(\pi^{(<n)}\right) = \mathbb{E}[B]$

Our Approach

Three major changes from [Yu'22]:

- Introduce new parameter "disadvantage" denoted $\varepsilon(\pi) = 1 - \mathrm{adv}(\pi)$ as a proxy to error probability.
- Impose a conditioning event $W = (Z > 0.01)$.
- "Binary" decomposition of protocols.

Staring with a protocol π with $IC(\pi) = I$ and $\varepsilon(\pi) = \varepsilon \approx \Theta(1)$. We decompose π into π_0 and π_1 both computing $f^{\oplus n/2}$ over input distribution $\mu^{n/2}$.



Lemma: $IC(\pi_0) + IC(\pi_1) \leq \exp(\varepsilon(\pi)) \cdot IC(\pi)$
Lemma: $\varepsilon(\pi_0) + \varepsilon(\pi_1) \leq 1.98 \cdot \varepsilon(\pi)$

Proof Sketch of Main Theorem: Recursively apply the decomposition until level $m = \log_2 n$ where we get n protocols for f. Then, the sum of ε's at this level is:

$$\sum_{|S|=m} \varepsilon(\pi_S) \leq 1.98^m \cdot \varepsilon(\pi) < n^{0.99}$$

(*) Assume the *average case*: $\varepsilon(\pi_S) \approx 0.99^{|S|}\varepsilon$. Then,

$$\sum_{|S|=m} IC(\pi_S) \leq IC(\pi) \cdot \exp\left(\sum 0.99^{|S|}\varepsilon\right) = O(I).$$

So, on average π_S has error $n^{-0.01}$ and IC $\approx I/n$. ■

Technical Challenges

- Assumption (*) is not always true.
  Fix: sample S from a more complicated distribution.
- Conditioning induces arbitrary correlation between inputs and messages; thus, π_S is no longer a protocol
  Fix: each conditioning event occurs w.h.p.– thus the correlation is small in expectation. Keep track of this quantity and compensate it by an $o_n(1)$ additive loss.