

Lecture 20

*Instructor: Sruthi Sekar**Scribe: Patil Vipul Sudhir*

1 Recap and Lecture Overview

In last class we discussed about communication complexity. We have seen deterministic communication complexity. Our goal was to minimize communication and for that we have seen Tree view, Communication Matrix, Lower bound for communication complexity. This lecture completes our investigation of communication complexity, focusing on:

- Randomized communication models (public vs private coins)
- Newman's theorem connecting public and private randomness
- Distributional complexity and Yao's Minimax Principle
- Fundamental lower bound for Inner Product mod 2 (IP_2)

2 Randomized Communication Models

2.1 Definitions

2.1.1 Private Coin Protocol

Setting:

- Alice and Bob each have their own private random strings (r_A and r_B), unknown to other.
- They communicate bits to compute $f(x, y)$, using r_A and r_B for randomization.

A protocol where Alice and Bob have independent randomness r_A and r_B . Communication cost measured by worst-case bits over x, y, r_A, r_B .

2.1.2 Public Coin Protocol

Setting:

- Alice and Bob share a common random string r (publicly visible to both).
- They communicate bits deterministically based on x, y, r .

A protocol where Alice and Bob share access to a common random string r . Communication cost is measured by worst-case bits over x, y, r .

2.2 Key Example: Equality Testing

Equality Testing Problem- For inputs $x, y \in \{0, 1\}^n$, define:

$$EQ_n(x, y) = \begin{cases} 1 & \text{if } x = y \\ 0 & \text{otherwise} \end{cases}$$

Let's assume Alice and Bob agree (in advance) on using a binary ECC for n -bit integers, producing $\mathcal{O}(n)$ -bit codewords and with minimum distance S_n (i.e. $[\mathcal{O}(n), n, S_n]$ - asymptotically good code)

2.2.1 Private Coin Protocol

Algorithm 1 Private Coin EQ Protocol

- 1: **Pre-agreement:** Fix ECC $\text{Enc} : \{0, 1\}^n \rightarrow \{0, 1\}^{O(n)}$
 - 2: **Alice:**
 - 3: Compute $\text{Enc}(x)$
 - 4: Pick $k = O(1)$ random indices $i_1, \dots, i_k$
 - 5: Send $(i_1, \text{Enc}(x)_{i_1}), \dots, (i_k, \text{Enc}(x)_{i_k})$ ($O(\log n)$ bits)
 - 6: **Bob:**
 - 7: Compare $\text{Enc}(y)_{i_j}$ with received bits
 - 8: Output 1 if all match, else 0
-

Error Analysis

- If $x = y$: Always correct
- If $x \neq y$: It implies that $\text{Enc}(x)$ and $\text{Enc}(y)$ differ in atleast S_n positions. In this case,
 $\Pr[\text{output} = 1] = (1 - \delta)^k$

For appropriate specific constant δ and k this probability $\leq 1/3$

2.2.2 Public Protocol

Theorem 1. $R_{\varepsilon, \text{pub}}(EQ_n) \leq 3$

Alice and Bob can specifically use a Hamming code as the binary error code.

Algorithm 2 Public Coin EQ Protocol

- 1: **Shared randomness:** Sample random $h : \{0, 1\}^n \rightarrow \{0, 1\}$
 - 2: **Alice:** Send $h(x)$ (1 bit)
 - 3: **Bob:** Output 1 if $h(x) = h(y)$, else 0
-

Error Analysis

- If $x = y$: Always correct
- If $x \neq y$: $\Pr[h(x) = h(y)] = \frac{1}{2}$
- Repeat $O(\log \frac{1}{\delta})$ times to reduce error to δ

3 Connecting Public and Private Coins

Theorem 2 (Newman's Theorem). *Any public coin protocol can be converted to a private coin protocol with at most $+\delta$ -increase in error probability and at most an $O(\log(n) + \log(1/\delta))$ increase in the communication bits (for any $\delta > 0$), i.e.*

For $f : \{0, 1\}^n \times \{0, 1\}^n \rightarrow \{0, 1\}$, For $\varepsilon, \delta > 0$

$$R_{\varepsilon+\delta, \text{priv}}(f) \leq R_{\varepsilon, \text{pub}}(f) + O\left(\log n + \log \frac{1}{\delta}\right)$$

Proof Sketch. 1. Let public protocol use random string r

2. Sample $t = \Theta(n/\delta^2)$ candidate strings

3. Using Chernoff bound + union bound we can show that- For some $t = \Theta(n/\delta^2)$, $\exists r_1, r_2, \dots, r_t$ such that, for every $(x, y) \in \{0, 1\}^n \times \{0, 1\}$, we have $\Pr_{i \sim [t]}[\pi^{Pub}(x, y; r_i) \neq f(x, y)] \leq \varepsilon + \delta$
4. Now both Alice and Bob will get hardcoded private coins generated from public coins. Convert to private coins by having Alice select sample index.

□

4 Distributional Complexity & Yao's Minimax

Definition 3 (Distributional Complexity). $D_\varepsilon^\mu(f)$ = minimum communication cost of deterministic protocol with error $\leq \varepsilon$ under distribution μ or least cost of deterministic protocol π such that $\Pr_{(x,y) \sim \mu}[\pi(x, y) = f(x, y)] \geq 1 - \varepsilon$

Example 4 (Equality (EQ) under uniform μ). For $f = EQ_n$ and uniform distribution μ over $\{0, 1\}^n \times \{0, 1\}^n$:

- $\Pr_{(x,y) \sim \mu}[x = y] = 2^{-n}$
- *Protocol: Always output 0*
- *Achieves error $\varepsilon = 2^{-n}$*
- *Distributional complexity: $D_{1/4}^{unif}(EQ_n) = O(1)$*

Example 5 (Disjointness (DISJ) under uniform μ). For $f = DISJ_n$ and uniform distribution μ over $\{0, 1\}^n \times \{0, 1\}^n$:

- *Most pairs are disjoint: $\Pr[\exists i : x_i = y_i = 1] \approx 1 - e^{-1}$*
- *Protocol: Always output 0*
- *Error $\varepsilon = \Pr[\exists i : x_i = y_i = 1] \approx 1 - e^{-1}$*
- *Distributional complexity: $D_{1/4}^{unif}(DISJ_n) = O(1)$*

Theorem 6. For every distribution μ over $X \times Y$ and every function f ,

$$R_{\varepsilon, pub}(f) \geq D_\varepsilon^\mu(f)$$

Proof. Let π be a public-coin protocol for f with communication cost $c = R_{\varepsilon, pub}(f)$. For each fixed public random string r , π_r is a deterministic protocol. By the correctness guarantee,

$$\forall x, y : \Pr_r[\pi_r(x, y) \neq f(x, y)] \leq \varepsilon$$

Averaging over μ ,

$$\mathbb{E}_{(x,y) \sim \mu} \Pr_r[\pi_r(x, y) \neq f(x, y)] \leq \varepsilon$$

this is equivalent to

$$\Pr_{r, (x,y) \sim \mu}[\pi_r(x, y) \neq f(x, y)] \leq \varepsilon$$

By the probabilistic method, there exists a fixed r^* such that

$$\Pr_{(x,y) \sim \mu}[\pi_{r^*}(x, y) \neq f(x, y)] \leq \varepsilon$$

□

Theorem 7 (Yao's Minimax).

$$R_{\epsilon, \text{pub}}(f) = \max_{\mu} D_{\epsilon}^{\mu}(f)$$

To establish a lower bound $R_{\epsilon, \text{pub}}(f) \geq c$:

1. **Identify Hard Distribution:** Find a distribution μ over inputs where f appears "random"
2. **Analyze Deterministic Protocols:** Show every deterministic protocol with $< c$ bits has error $> \epsilon$ under μ
3. **Apply Minimax:** Conclude no public-coin protocol can beat this worst-case distribution

4.1 Application of Yao's Minimax for $f = IP_2$

$$IP_2(x, S) = \bigoplus_{i \in S} x_i \quad \text{where } x \in \{\pm 1\}^n, S \subseteq [n]$$

Equivalently, using product notation with $\chi_S(x)$ (Alice has $X \in \{-1, +1\}^n$ and Bob has $S \subseteq [n]$):

$$IP_2(x, S) = \prod_{i \in S} x_i = \chi_S(x)$$

Communication matrix for this is Hadamard Matrix (M_{IP_2}).

Theorem 8. $R_{1/4, \text{pub}}(IP_2) \geq \frac{n}{2} - 1$

Proof. By Yao's minmax principle, it suffices to show that for some "hard" distribution μ , $D_{1/4}^{\mu}(IP_2) \geq n/2 - 1$. We will take μ as uniform distribution.

Let π be a deterministic c -bit protocol for IP_2 . We can partition the M_{IP_2} into $\leq 2^c$ combinatorial rectangles (recall it from the last class): $R_1 \times S_1, \dots, R_{2^c} \times S_{2^c}$, where the leaves of protocol π are $\{\pm 1\}$. Each combinatorial rectangle $R_i \times S_i$ corresponds to the output $Z_i \in \{\pm 1\}$ of π .

$$\pi(X, S) = \sum_{i=1}^{2^c} \mathbb{I}[(X, S) \in R_i \times S_i] \cdot Z_i \tag{1}$$

$$= \sum_{i=1}^{2^c} \mathbb{I}[x \in R_i] \mathbb{I}[S \in S_i] \tag{2}$$

Suppose $\Pr_{(X, S) \sim \mu} [\pi(X, S) = IP_2(X, S)] \geq 3/4$

To show $c \geq n/2 - 1$

$$\mathbb{E}_{(X, S) \sim \mu} [\pi(X, S) \cdot IP_2(X, S)] \geq 3/4(+1) + 1/4(-1) = 1/2$$

from above wkt, $\pi(X, S) = \sum_{i=1}^{2^c} \mathbb{I}_{R_i}[x \in R_i] \mathbb{I}[S \in S_i]$, let's put this into above equation-

$$\mathbb{E}_{(X, S)} \left[\sum_{i=1}^{2^c} \mathbb{I}_{R_i}(X) \cdot \mathbb{I}_{S_i}(S) \cdot Z_i \cdot IP_2(X, S) \right] \geq 1/2$$

$$1/2 \leq \sum_{i=1}^{2^c} Z_i \cdot \mathbb{E}_{(X, S)} [\mathbb{I}_{R_i}(X) \cdot \mathbb{I}_{S_i} \cdot IP_2(X, S)]$$

$$1/2 \leq 2^c \cdot \max_i \{ \mathbb{E}_{(X, S)} [\mathbb{I}_{R_i}(X) \cdot \mathbb{I}_{S_i} \cdot IP_2(X, S)] \}$$

$\max_i \{ \mathbb{E}_{(X, S)} [\mathbb{I}_{R_i}(X) \cdot \mathbb{I}_{S_i} \cdot IP_2(X, S)] \}$ captures the local discrepancies of $IP_2(X, S)$ on combinatorial rectangle $R_i \times S_i$ under uniform distribution.

wkt $\forall i, \mathbb{E}_{(X,S)}[\mathbb{I}_{R_i}(X) \cdot \mathbb{I}_{S_i} \cdot IP_2(X,S)] \leq 2^{-n/2}$, so substituting this into above equation we get-

$$1/2 \leq 2^c \cdot 2^{-n/2} \tag{3}$$

$$-1 \leq c - n/2 \tag{4}$$

$$c \geq n/2 - 1 \tag{5}$$