Communication Complexity

Computational Complexity - Sanjeev Arora and Boaz Barak

Presented by Ashish Kumar, Somya Sangal

Computational Complexity Theory - E0-224
Indian Instititute of Science, Banglore

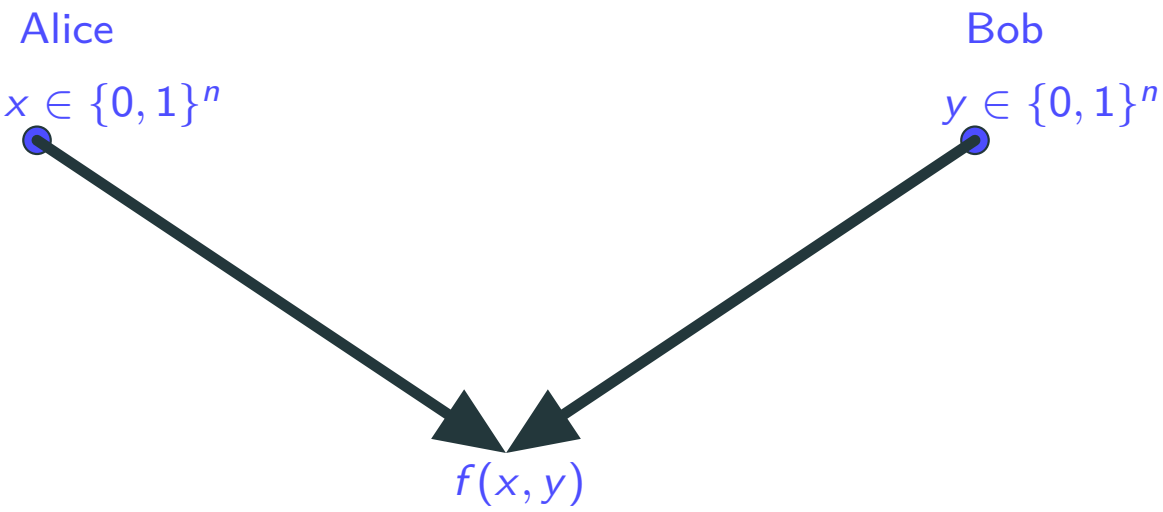
Table of Contents

- Introduction
- Lower bound methods
  - The fooling set method
  - The tiling method
  - The rank method
  - The discrepancy method
- Comparison between lower bounds
- Multiparty communication complexity
- Other Communication models
- Bibliography

Introduction

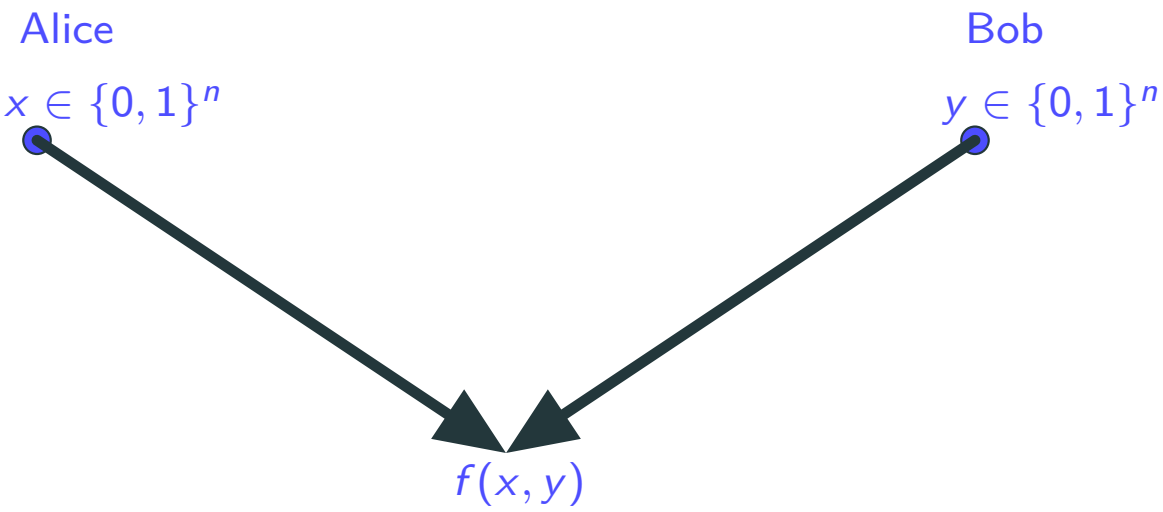
- Two parties, Alice and Bob with **unlimited** computational power.

Introduction



- Two parties, Alice and Bob with **unlimited** computational power.
- Wish to collaboratively compute $f(x,y)$ where $f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$.
- Already agreed upon communication protocol.

Introduction



- Two parties, Alice and Bob with **unlimited** computational power.
- Wish to collaboratively compute $f(x,y)$ where $f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$.
- Already agreed upon communication protocol.
- Cost of the protocol is the number of bits communicated for worst-case choice for x and y.

Definition of two-party communication complexity [1]

- Two parties (players), Alice and Bob.
- Function $f : \{0,1\}^{2n} \to \{0,1\}$.
- A *t-round two-party protocol* Π for computing f which is a sequence of t functions $P_1, ..., P_t : \{0,1\}^* \to \{0,1\}^*$.
- An execution of Π on inputs x, y involves the following: Alice computes $p_1 = P_1(x)$ and sends p_1 to Bob. Bob computes $p_2 = P_2(y, p_1)$ and sends p_2 to Alice. And so on.
- Protocol Π is valid, if $\forall x, y$, last message sent p_t is equal to $f(x, y)$.

Diagram for two-party communication model

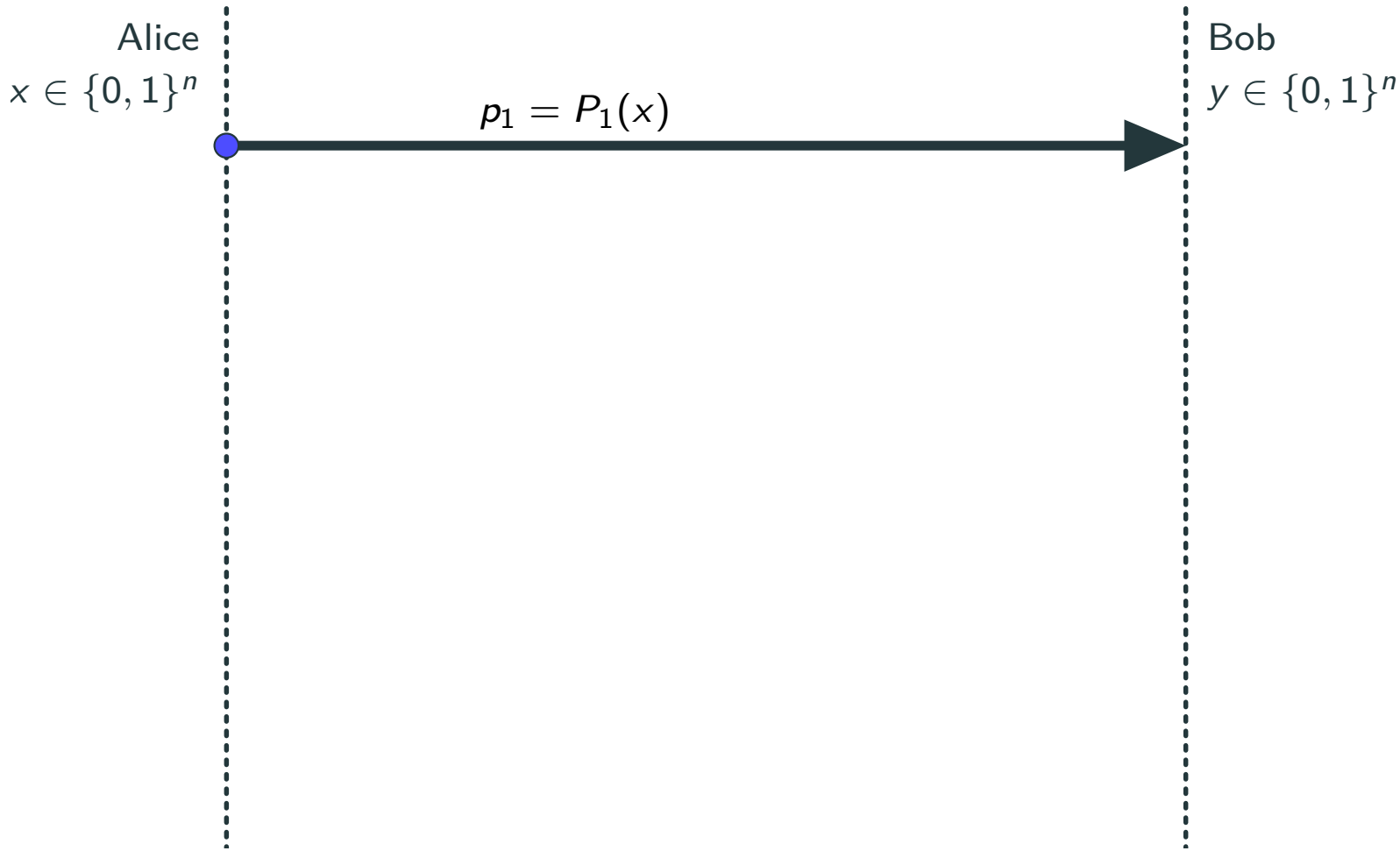


Diagram for two-party communication model

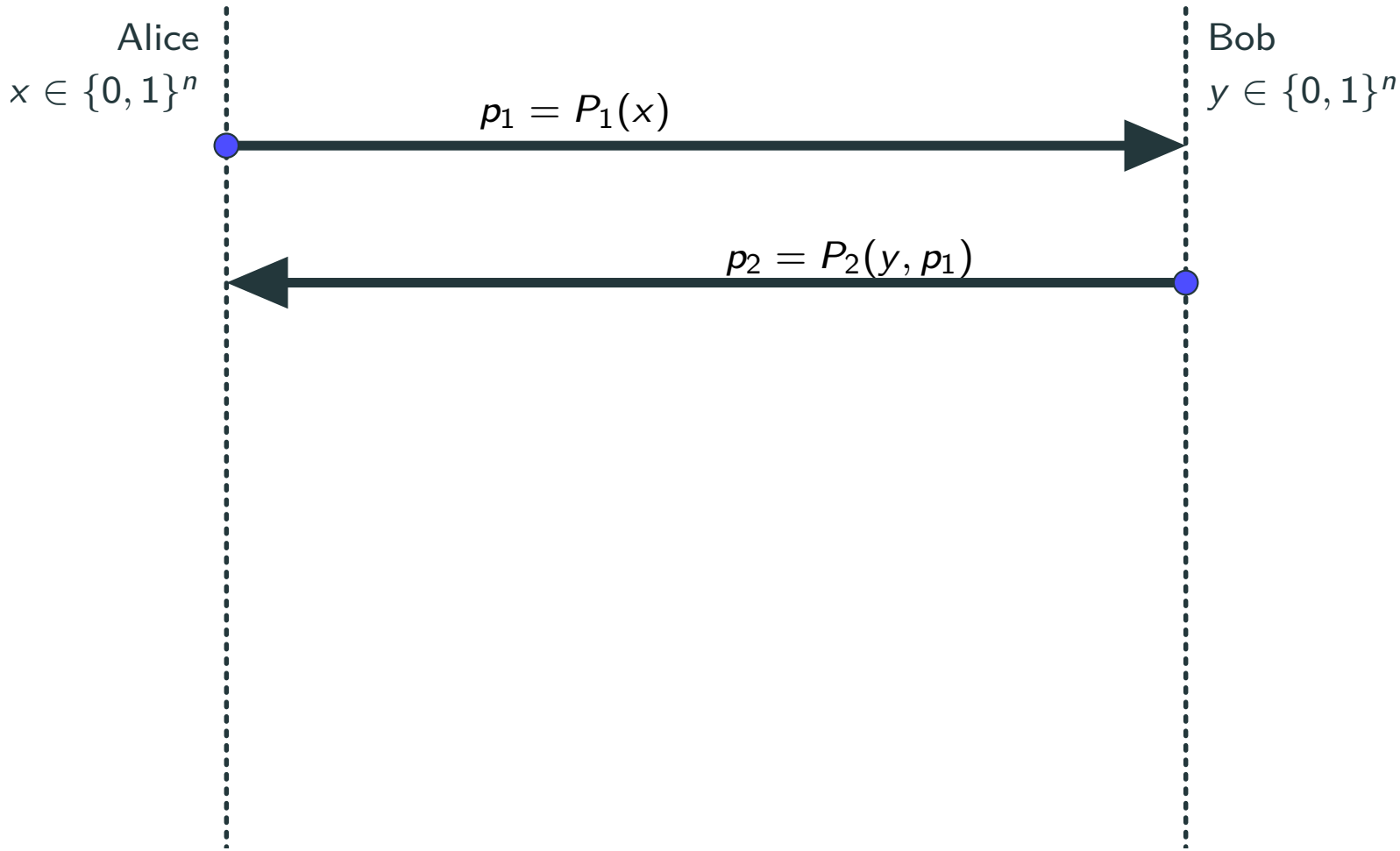


Diagram for two-party communication model

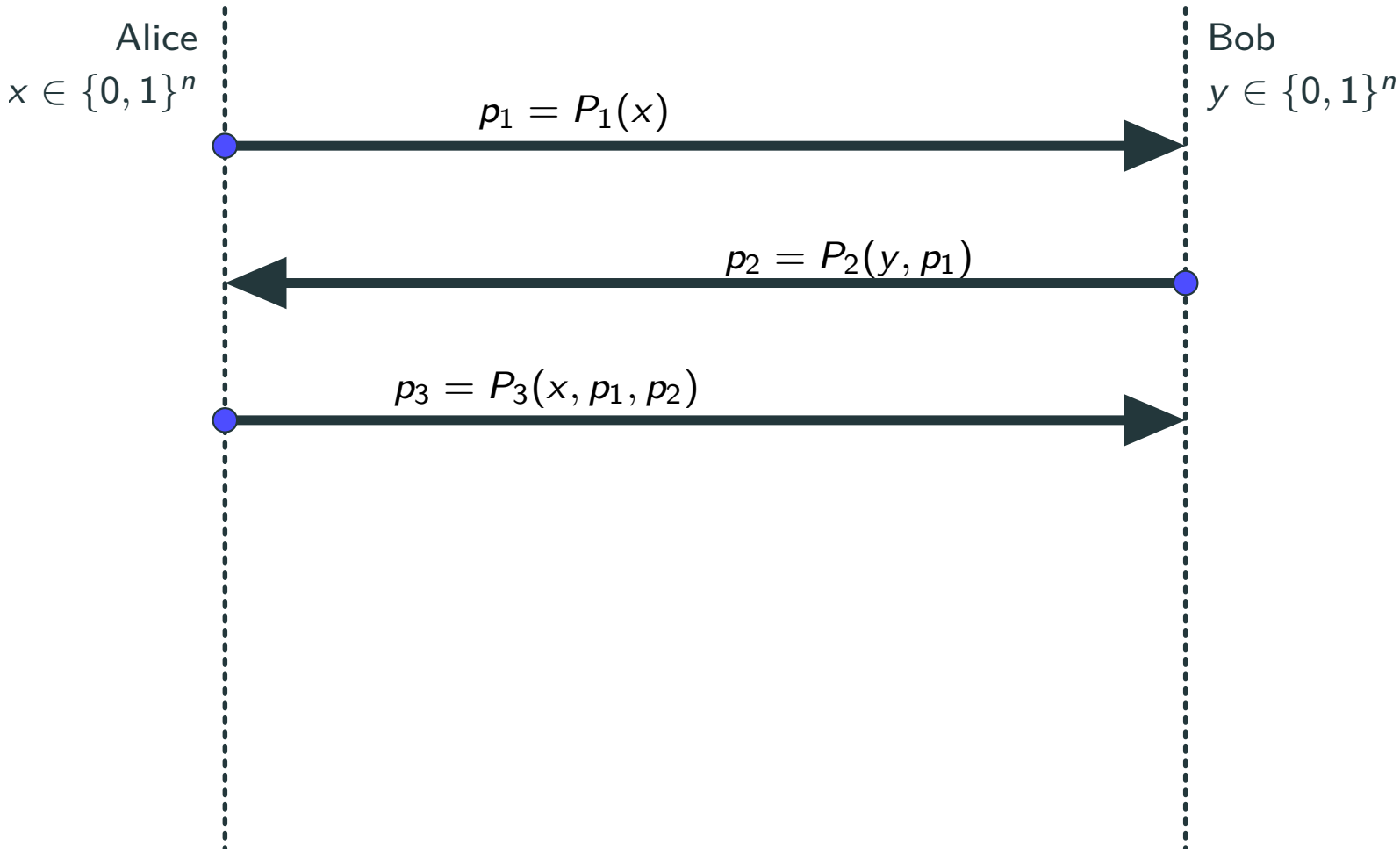
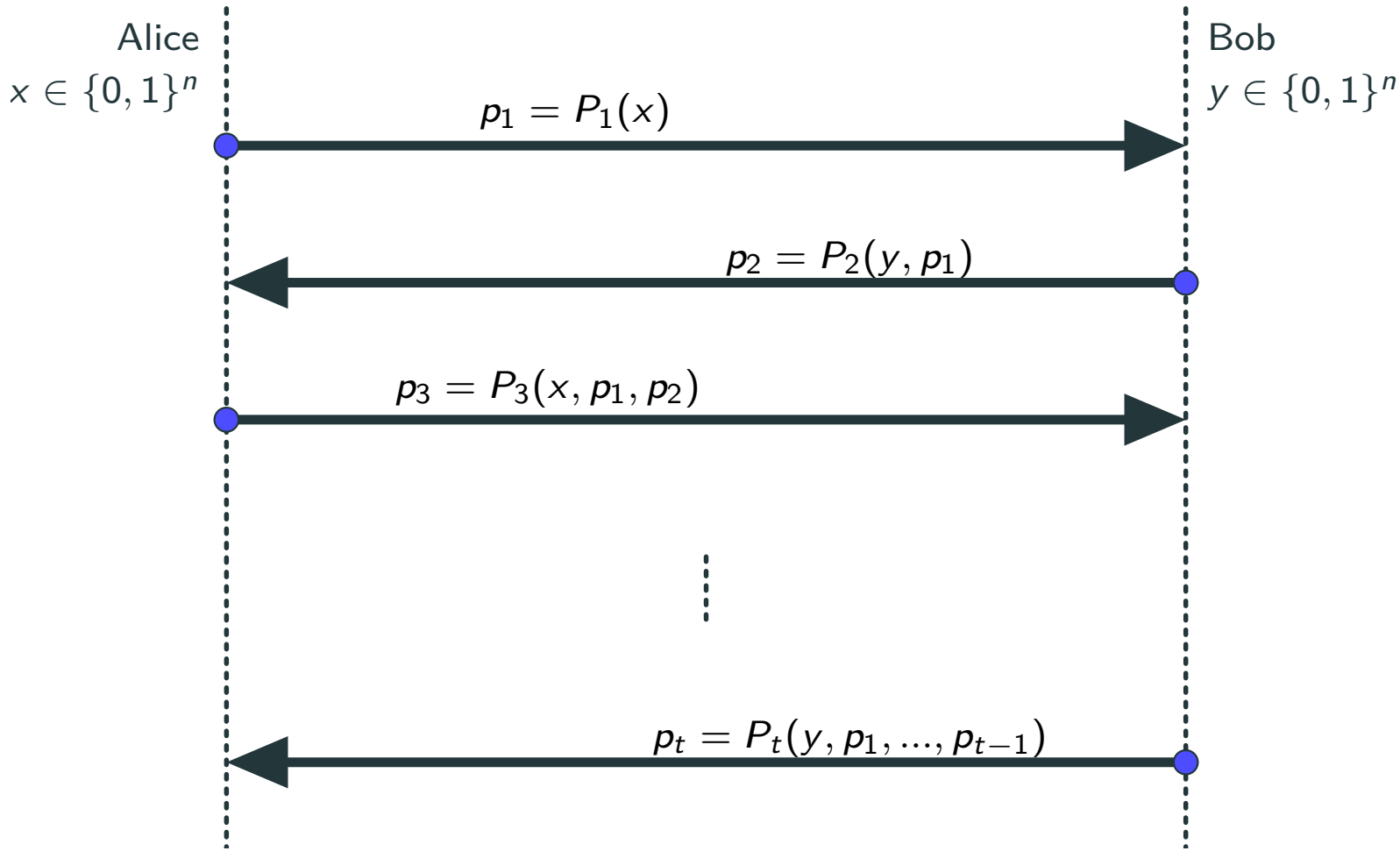


Diagram for two-party communication model



Definition of two-party communication complexity

- *Communication complexity* of Π is maximum number of bits communicated, i.e., $\max(\{|p_1|+\ldots+|p_t|: \forall x, y \in \{0,1\}^n\})$
- *Communication complexity* of f (denoted by $C(f)$), is minimum communication complexity over all valid protocols Π for f.

Definition of two-party communication complexity

- *Communication complexity* of Π is maximum number of bits communicated, i.e., $\max(\{|p_1|+\ldots+|p_t|: \forall x, y \in \{0,1\}^n\})$
- *Communication complexity* of f (denoted by $C(f)$), is minimum communication complexity over all valid protocols Π for f.
- For any function f, $C(f) \leq n+1$. Why?

Examples

- **Parity**: $\mathrm{PARITY}(x, y)$ is parity of all bits in x, y.
  $C(\mathit{PARITY}) = 2$.

Examples

- **Parity**: $\mathrm{PARITY}(x, y)$ is parity of all bits in x, y. $C(PARITY) = 2$.
- **Halting Problem**:
  - Function $H : \{0,1\}^n \times \{0,1\} \to \{0,1\}$, defined as follows. If $x = 1^n$ and $y = code(M)$ for some Turing machine M such that M halts on x, then $H(x, y) = 1$ and otherwise $H(x, y) = 0$.

Examples

- **Parity**: $\mathrm{PARITY}(x, y)$ is parity of all bits in x, y. $C(PARITY) = 2$.
- **Halting Problem**:
  - Function $H : \{0,1\}^n \times \{0,1\} \to \{0,1\}$, defined as follows. If $x = 1^n$ and $y = code(M)$ for some Turing machine M such that M halts on x, then $H(x, y) = 1$ and otherwise $H(x, y) = 0$.
  - $C(H) = 2$.

Examples

- **Parity**: $\mathrm{PARITY}(x, y)$ is parity of all bits in x, y. $C(PARITY) = 2$.
- **Halting Problem**:
  - Function $H : \{0,1\}^n \times \{0,1\} \to \{0,1\}$, defined as follows. If $x = 1^n$ and $y = code(M)$ for some Turing machine M such that M halts on x, then $H(x, y) = 1$ and otherwise $H(x, y) = 0$.
  - $C(H) = 2$.
  - Both parties have unbounded computation power.

Example - Equality

- Equality function will be used as running example through the presentation.

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

Theorem 1

Equality has linear communication complexity $C(EQ) \geq n$.

- We will prove Theorem 1 using various methods.

- Introduction
- Lower bound methods
  - The fooling set method
  - The tiling method
  - The rank method
  - The discrepancy method
- Comparison between lower bounds
- Multiparty communication complexity
- Other Communication models
- Bibliography

- Introduction
- Lower bound methods
  - The fooling set method
  - The tiling method
  - The rank method
  - The discrepancy method
- Comparison between lower bounds
- Multiparty communication complexity
- Other Communication models
- Bibliography

Fooling set method

Observation 1

For any communication protocol for any function, $\forall$ distinct $x, x' \in \{0,1\}^n$ such that the communication pattern is the same for (x, x) and (x', x'), then the output of the protocol will be same for $(x, x), (x, x'), (x', x), (x', x')$.

Proof by induction: Let us assume, Alice has x and Bob has x'. It is given that communication pattern for (x, x) and (x', x') is same. If Alice communicates a bit first, it will be same whether his input in x or x'. If Bob communicates next, it will communicate same bit whether its input is x or x' since he receives the same bit from Alice. And so on. At the end, the Alice and Bob answer on (x, x) must agree with their answer on (x, x'). □

Proof Of $C(EQ) \geq n$ by Fooling Method

- For contradiction let us assume, $C(EQ) = n - 1$. We could have 2^{n-1} different possible communication pattern.
- For $2n-$bit input pair (x, x), we have 2^n possibilities. By pigeonhole principle, $\exists$ distinct pair (x, x) and (x', x') such that their communication pattern is same.
- Using Observation, $EQ(x, x) = 0 = EQ(x, x')$. Contradiction. $\square$

Generalized Fooling Set Method

- Definition of Fooling Set $S \subseteq \{0,1\}^n \times \{0,1\}^n$ corresponding to a value $b \in \{0,1\}$ and function $f : \{0,1\} \times \{0,1\} \to \{0,1\}$
  1. $\forall \langle x, y \rangle \in S$, $f(x,y) = b$
  2. $\forall$ distinct $\langle x, y \rangle, \langle x', y' \rangle \in S$, either $f(x,y') \neq b$ or $f(x',y) \neq b$.

Generalized Fooling Set Method

- Definition of Fooling Set $S \subseteq \{0,1\}^n \times \{0,1\}^n$ corresponding to a value $b \in \{0,1\}$ and function $f : \{0,1\} \times \{0,1\} \to \{0,1\}$
  1. $\forall \langle x, y \rangle \in S$, $f(x,y) = b$
  2. $\forall$ distinct $\langle x, y \rangle, \langle x', y' \rangle \in S$, either $f(x,y') \neq b$ or $f(x',y) \neq b$.

Lemma 2

If f has size-M fooling set then $C(f) \geq \log M$.

Generalized Fooling Set Method

- Definition of Fooling Set $S \subseteq \{0,1\}^n \times \{0,1\}^n$ corresponding to a value $b \in \{0,1\}$ and function $f : \{0,1\} \times \{0,1\} \to \{0,1\}$
  1. $\forall \langle x, y \rangle \in S$, $f(x,y) = b$
  2. $\forall$ distinct $\langle x, y \rangle, \langle x', y' \rangle \in S$, either $f(x,y') \neq b$ or $f(x',y) \neq b$.

Lemma 2

If f has size-M fooling set then $C(f) \geq \log M$.

- We need at least M distinct communication pattern. Why?

Generalized Fooling Set Method

- Definition of Fooling Set $S \subseteq \{0,1\}^n \times \{0,1\}^n$ corresponding to a value $b \in \{0,1\}$ and function $f : \{0,1\} \times \{0,1\} \to \{0,1\}$
  1. $\forall \langle x, y \rangle \in S$, $f(x,y) = b$
  2. $\forall$ distinct $\langle x, y \rangle, \langle x', y' \rangle \in S$, either $f(x,y') \neq b$ or $f(x',y) \neq b$.

Lemma 2

If f has size-M fooling set then $C(f) \geq \log M$.

- We need at least M distinct communication pattern. Why?
- If we have $M-1$ distinct communication pattern then $\exists \langle x, y \rangle, \langle x', y' \rangle \in S$ s.t., their communication pattern is same.
- From observation, $f(x,y') = f(x',y) = b$. Contradiction.
- For M distinct communication pattern we require at least $\log M$ bits. $C(f) \geq \log M$.

□

Example

Disjointness

- x, y are characteristic vectors of subsets of $\{0, ..., n\}$. Thus, $x, y \in \{0, 1\}^n$.
- $\mathrm{DISJ}(x, y) = 1$ if x and y are disjoint, 0 otherwise.
- $C(DISJ) \geq n$

Proof: Following 2^n pairs constitute fooling set:

$$S = \left\{(A, \bar{A}) : A \subseteq \{1, 2, ..., n\}\right\}$$

Thus, $C(DISJ) \geq \log(2^n) = n$. □

- Introduction
- Lower bound methods
  - The fooling set method
  - The tiling method
  - The rank method
  - The discrepancy method
- Comparison between lower bounds
- Multiparty communication complexity
- Other Communication models
- Bibliography

The Tiling Method

- The tiling bound for lower bounds take a more global view of the function f.
- Matrix $M(f)$ of size $2^n \times 2^n$ whose (x, y)th entry has value $f(x, y)$.

Matrix of EQ

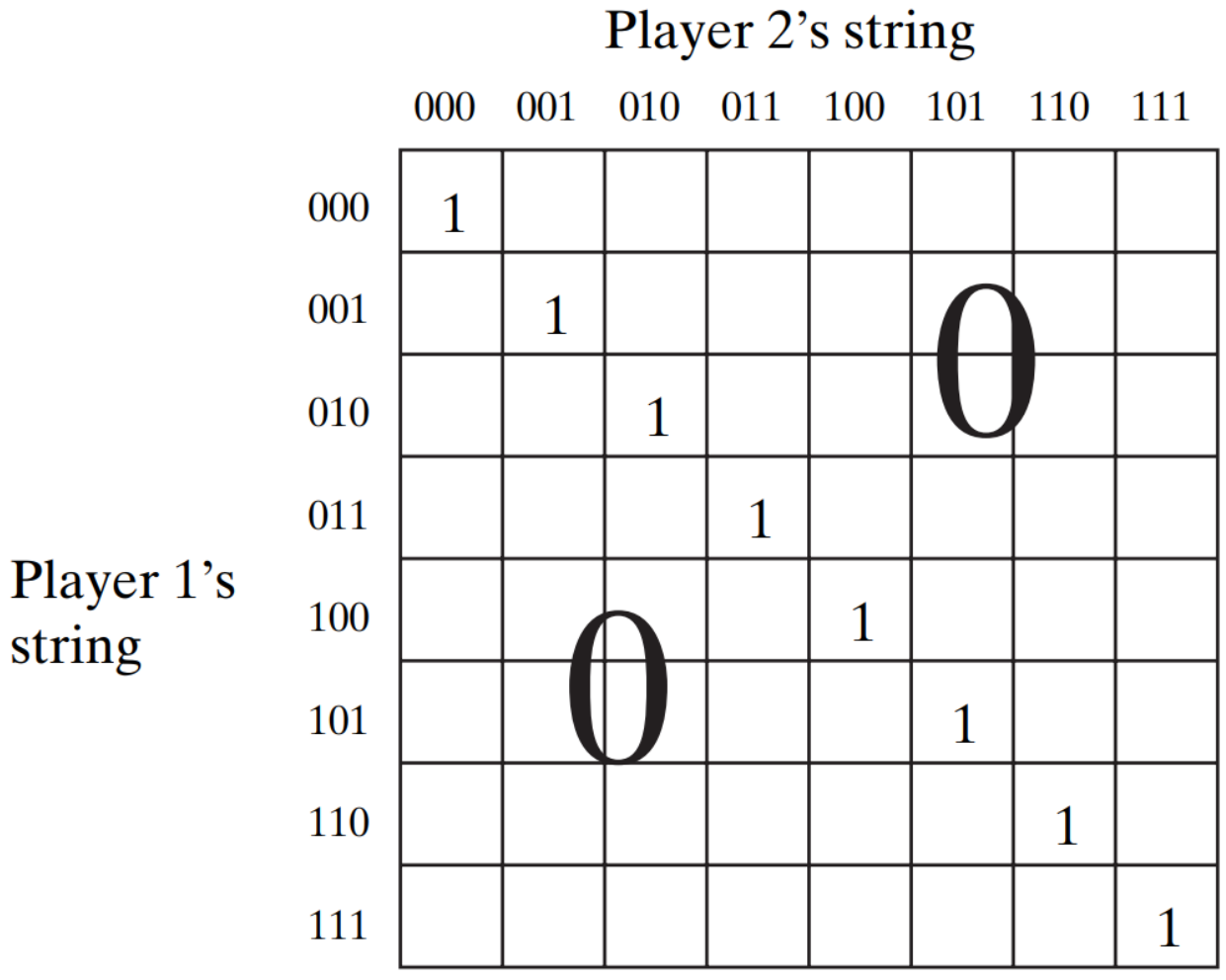


Figure 1: Matrix $M(f)$ for the equality function when the inputs have 3 bits.

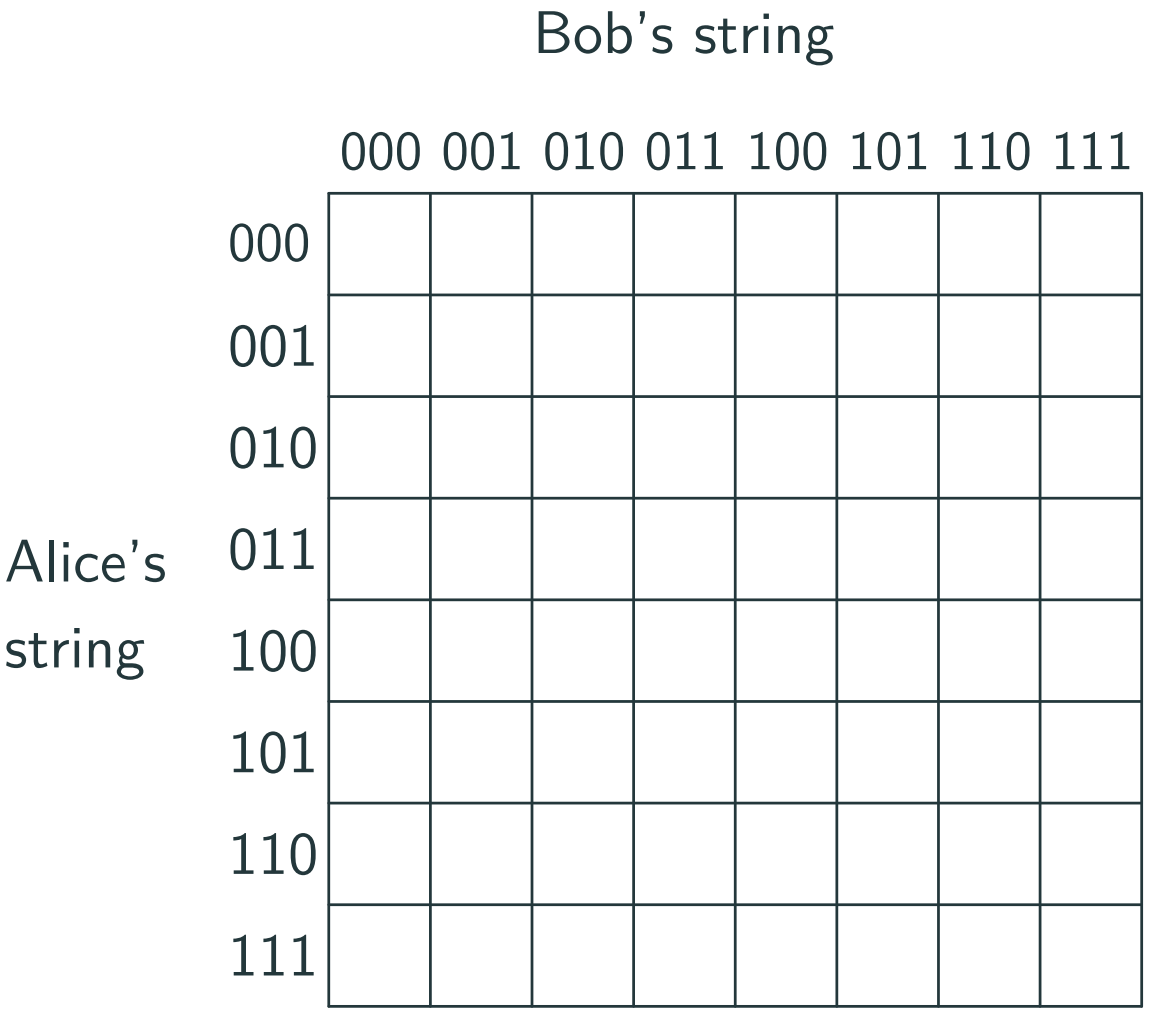
The Tiling Method

- The tiling bound for lower bounds take a more global view of the function f.
- Matrix $M(f)$ of size $2^n \times 2^n$ whose (x, y)th entry has value $f(x, y)$.
- Combinatorial Rectangle: Sub-matrix of M that corresponds to entries in $A \times B$ where $A \subseteq \{0,1\}^n, B \subseteq \{0,1\}^n$
- Monochoromatic rectangle: $A \times B$ is monochromotic if $\forall x \in A, y \in B$, M_{xy} is same.

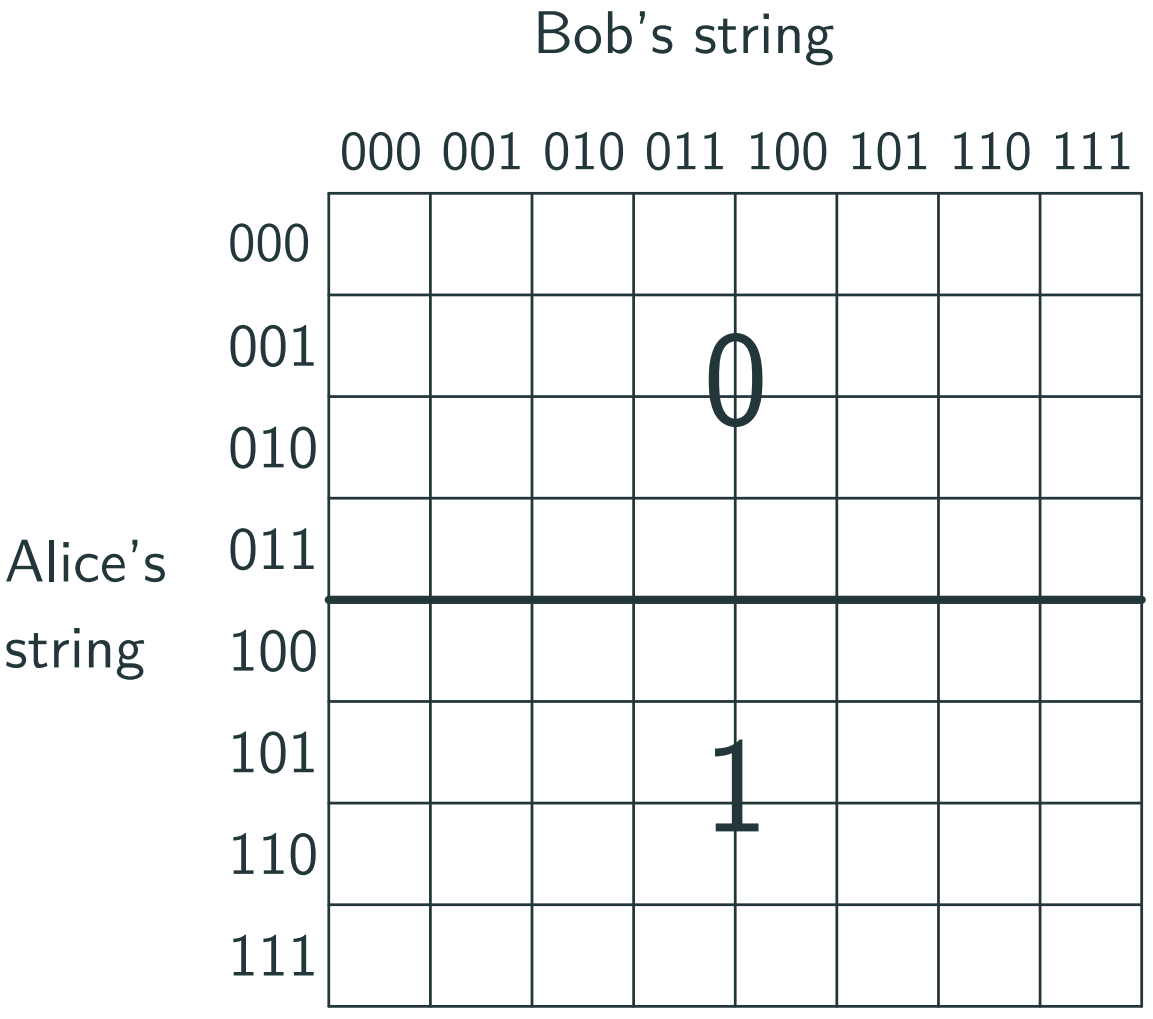
Partition of M based on communication bits

- If Alice sends first bit, $M(f)$ gets partitions into two rectangles of type $A_0 \times \{0,1\}^n, A_1 \times \{0,1\}^n$ where $A_b \subseteq A$ for which bit sent by Alice is b.
- Bob sends next bit, which further partitions two rectangles into smaller rectangles (combinatorial rectangle).
- If total number of bits communicated is k, then matrix gets partitioned into 2^k rectangles.

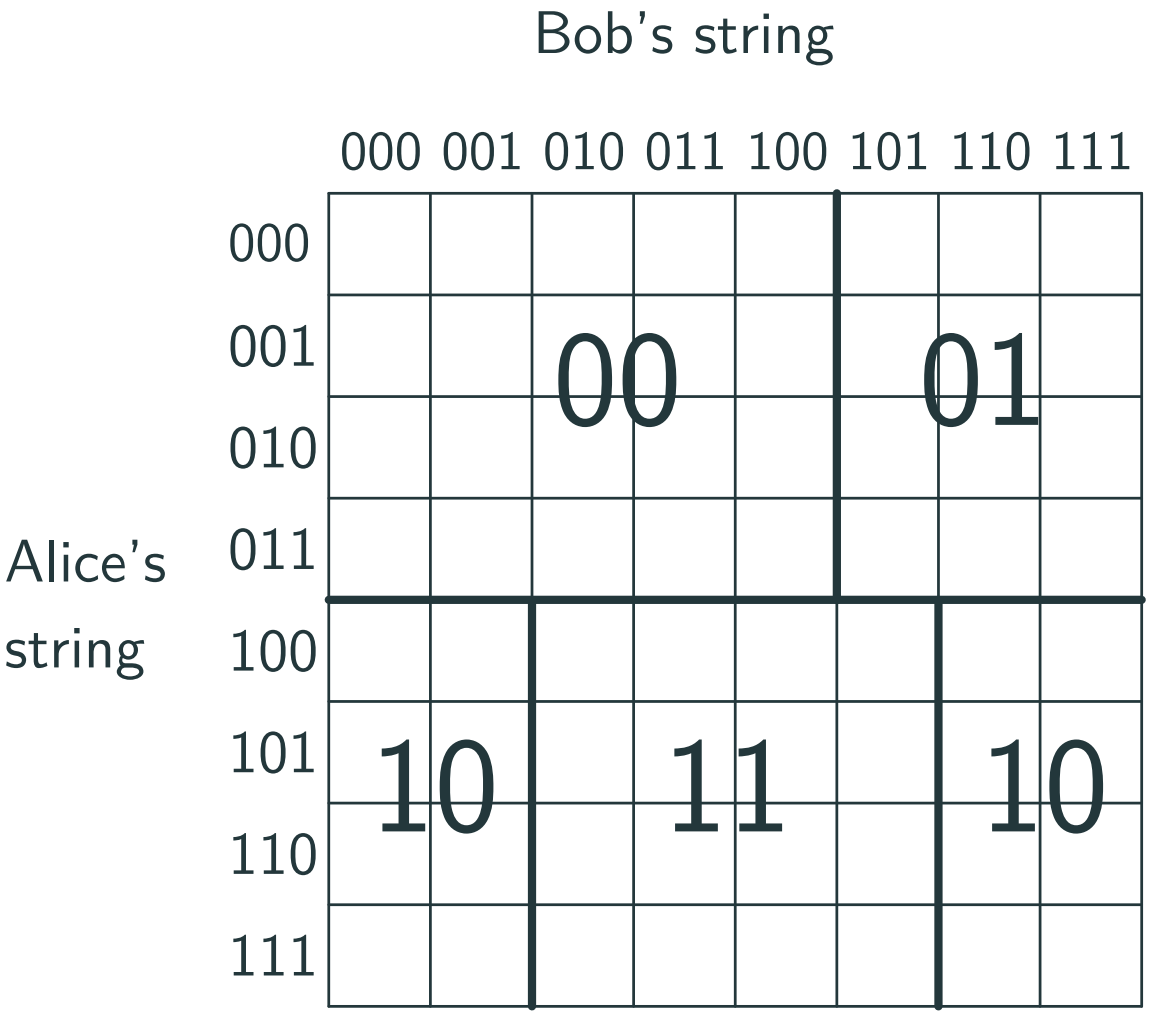
Example of Communication Matrix



Example of Communication Matrix



Example of Communication Matrix



Partition of M based on communication bits

Observation 2

After protocol ends, each partition of matrix will be monochromatic rectangle. Why?

Partition of M based on communication bits

Observation 2

After protocol ends, each partition of matrix will be monochromatic rectangle. Why?

Definition:

- Monochromatic tiling of $M(f)$ is a partition of $M(f)$ into disjoint monochromatic rectangles.
- $\chi(f)$: Minimum of rectangles in any monochromatic tiling of $M(f)$.

Lower and Upper bound using Tiling Method

Theorem 3

$\log_2 \chi(f) \leq C(f) \leq 16(\log_2 \chi(f))^2$

Lower and Upper bound using Tiling Method

Theorem 3

$\log_2 \chi(f) \leq C(f) \leq 16(\log_2 \chi(f))^2$

- Number of rectangles in M at most doubles for each bit communicated.
- Thus, $\chi(f) \leq 2^{C(f)} \Rightarrow \log \chi(f) \leq C(f)$

Relation between fooling set method and tiling method

Lemma 4

If f has a fooling set with m pairs, then $\chi(f) \geq m$.

Relation between fooling set method and tiling method

Lemma 4

If f has a fooling set with m pairs, then $\chi(f) \geq m$.

- For contradiction, $\chi(f) = m - 1$.
- By pigeonhole principle, $\exists(x_1, y_1)$ and (x_2, y_2) as two pairs in fooling set such that they belong to in same monochromatic rectangle, thus $f(x_1, y_1) = f(x_2, y_1) = f(x_1, y_2) = f(x_2, y_2)$.
- Contradiction. □

Relation between fooling set method and tiling method

Lemma 4

If f has a fooling set with m pairs, then $\chi(f) \geq m$.

- For contradiction, $\chi(f) = m - 1$.
- By pigeonhole principle, $\exists(x_1, y_1)$ and (x_2, y_2) as two pairs in fooling set such that they belong to in same monochromatic rectangle, thus $f(x_1, y_1) = f(x_2, y_1) = f(x_1, y_2) = f(x_2, y_2)$.
- Contradiction. □

- $\log M \leq \log \chi(f) \leq C(f)$.
- Thus, Tiling method consumes fooling set method.

- Introduction
- Lower bound methods
  - The fooling set method
  - The tiling method
  - The rank method
  - The discrepancy method
- Comparison between lower bounds
- Multiparty communication complexity
- Other Communication models
- Bibliography

Rank Method

Lemma 5

The rank of an $n \times n$ matrix over a field $\mathbb{F}$, denoted by rank(M) is the minimum value of l such that M can be expressed as

$$M = \sum_{i=1}^{l} B_i$$

where each B_i is an $n \times n$ rank-1 matrix.

Rank Method

Lemma 5

The rank of an $n \times n$ matrix over a field $\mathbb{F}$, denoted by rank(M) is the minimum value of l such that M can be expressed as

$$M = \sum_{i=1}^{l} B_i$$

where each B_i is an $n \times n$ rank-1 matrix.

Proof idea: Matrix $M = \sum_{i=1}^{l} u_i u_i^T$ where $u_i u_i^T$ is a rank-1 matrix and u_is are linearly independent. Now, consider a matrix $Y = \sum_{i=1}^{n} u_i u_i^T$ where n is the dimension of the space and u_is form the basis for it. (using Gram-Schmidt process)

Matrix Y has rank n. Why?

Rank Method

$$Yv = \sum_{i=1}^{n} u_i u_i^T v$$

For the sake of contradiction, let $rank(Y) < n$. This means than $Yv = 0$ for some non-zero v.

From the above equation, $Yv = 0$ only when $u_i^T v = 0$ for all i as u_is are linearly independent.

Now, $v = c_1 u_1 + c_2 u_2 + \cdots + c_n u_n$ where atleast one of $c_1, \cdots, c_n$ is non-zero. But, $u_i^T v = 0$ for all i which is a contradiction. Hence, $rank(Y) = n$.

Rank of $(Y - M)$ can be atmost $n - l$, hence rank of M must be l.

Rank Method

Theorem 6

For every function f, $\chi(f) \geq rank(M(f))$ and hence the communication complexity $C(f) \geq \log_2(rank(M(f)))$

Proof: Consider a deterministic protocol which has communication complexity c. Then there exists a partition of M_f into atmost 2^c monochromatic rectangles. Let these rectangles be $R_1, \cdots, R_l$ for some $l \leq 2^c$.

Rank Method

Theorem 6

For every function f, $\chi(f) \geq rank(M(f))$ and hence the communication complexity $C(f) \geq \log_2(rank(M(f)))$

Proof: Consider a deterministic protocol which has communication complexity c. Then there exists a partition of M_f into atmost 2^c monochromatic rectangles. Let these rectangles be $R_1, \cdots, R_l$ for some $l \leq 2^c$.

Let $\mathbb{R} = \{R_i\}$ be the subset of rectangles for which $f(x, y) = 1$, $\forall (x, y) \in R_i$. For each such rectangle in $\mathbb{R}$, define a matrix M_i as

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

Now, $M_f = \sum_i M_i$

Rank Method

Theorem 7

For every function f, $\chi(f) \geq rank(M(f))$ and hence the communication complexity $C(f) \geq \log_2(rank(M(f)))$

Proof: Consider a deterministic protocol which has communication complexity c. Then there exists a partition of M_f into atmost 2^c monochromatic rectangles. Let these rectangles be $R_1, \cdots, R_l$ for some $l \leq 2^c$.

$$M(f) = \begin{array}{c} \\ \\ \\ \\ \\ \end{array} \begin{array}{c} \mathbf{R_i} \\ \left[\begin{array}{ccccc} 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ - & - & - & - & - \\ - & - & - & - & - \end{array}\right. \end{array}$$

$$M_i = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Rank Method

We know that $rank(A + B) \leq rank(A) + rank(B)$

Hence, $rank(M_f) \leq \sum_i rank(M_i)$

Rank Method

We know that $rank(A+B) \leq rank(A) + rank(B)$

Hence, $rank(M_f) \leq \sum_i rank(M_i)$

Now each M_i has rank atmost 1 since it takes the following form

$$M_i = \begin{bmatrix} 1 & 1 & \cdots & 1 & 0 & \cdots & 0 \\ 1 & 1 & \cdots & 1 & 0 & \cdots & 0 \\ 1 & 1 & \cdots & 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix}$$

Rank Method

We know that $rank(A+B) \leq rank(A) + rank(B)$

Hence, $rank(M_f) \leq \sum_i rank(M_i)$

Now each M_i has rank atmost 1 since it takes the following form

$$M_i = \begin{bmatrix} 1 & 1 & \cdots & 1 & 0 & \cdots & 0 \\ 1 & 1 & \cdots & 1 & 0 & \cdots & 0 \\ 1 & 1 & \cdots & 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix}$$

Hence, $rank(M_f) \leq l$ where l is the number of rectangles that M_f gets partitioned into. Finally,

$$rank(M_f) \leq \chi(f) \leq 2^{C(f)}$$

Rank Method

Ex 1:

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

The matrix M_f for EQ is $2^n \times 2^n$ identity matrix and hence the rank is 2^n. Therefore, $C(EQ) \geq n$

Rank Method

Ex 1:

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

The matrix M_f for EQ is $2^n \times 2^n$ identity matrix and hence the rank is 2^n. Therefore, $C(EQ) \geq n$

Ex 2: Consider the Dot Product function
$\mathbf{DP_n} : \{\mathbf{0}, \mathbf{1}\}^{\mathbf{n}} \times \{\mathbf{0}, \mathbf{1}\}^{\mathbf{n}} \rightarrow \{\mathbf{0}, \mathbf{1}\}$.
$DP(x, y) = \sum_{i=1}^{n} x_i y_i$

Rank Method

Ex 1:

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

The matrix M_f for EQ is $2^n \times 2^n$ identity matrix and hence the rank is 2^n. Therefore, $C(EQ) \geq n$

Ex 2: Consider the Dot Product function
$\mathbf{DP_n} : \{\mathbf{0},\mathbf{1}\}^{\mathbf{n}} \times \{\mathbf{0},\mathbf{1}\}^{\mathbf{n}} \rightarrow \{\mathbf{0},\mathbf{1}\}$.
$DP(x,y) = \sum_{i=1}^{n} x_i y_i$

Observation

The rank of the matrix M_{DP} is atmost $2^n - 1$ over $GF(2)^n$ and hence the communication complexity is atleast n.

Rank Method

Proof : Choice of field is crucial in giving tighter bounds. [2]

Rank Method

Proof: Choice of field is crucial in giving tighter bounds. [2]

- Consider the square of the matrix $N = M_{DP}^2$.
- We know that $(AB)_{ij} = \sum_{k=1}^{n} A_{ik} B_{kj}$, therefore, $N(x, y) = \sum_{z \in \{0,1\}^n} (x \cdot z) \times (y \cdot z)$.
- Now, $N(x, y)$ gives the number of vectors z for which $x \cdot z$ and $y \cdot z$ are both 1.

Rank Method

Proof: Choice of field is crucial in giving tighter bounds. [2]

- Consider the square of the matrix $N = M_{DP}^2$.
- We know that $(AB)_{ij} = \sum_{k=1}^{n} A_{ik} B_{kj}$, therefore, $N(x, y) = \sum_{z \in \{0,1\}^n} (x \cdot z) \times (y \cdot z)$.
- Now, $N(x, y)$ gives the number of vectors z for which $x \cdot z$ and $y \cdot z$ are both 1.
- When $x = y$, the number of such z's is $2^n/2 = 2^{n-1}$ and hence, diagonal entries of N are 2^{n-1}.
- When $x \neq y$, the number of such z's is $2^n/4 = 2^{n-2}$ and hence, off-diagonal entries of N are 2^{n-2}.

Rank Method

Proof: Choice of field is crucial in giving tighter bounds. [2]

- Consider the square of the matrix $N = M_{DP}^2$.
- We know that $(AB)_{ij} = \sum_{k=1}^{n} A_{ik} B_{kj}$, therefore, $N(x, y) = \sum_{z \in \{0,1\}^n} (x \cdot z) \times (y \cdot z)$.
- Now, $N(x, y)$ gives the number of vectors z for which $x \cdot z$ and $y \cdot z$ are both 1.
- When $x = y$, the number of such z's is $2^n/2 = 2^{n-1}$ and hence, diagonal entries of N are 2^{n-1}.
- When $x \neq y$, the number of such z's is $2^n/4 = 2^{n-2}$ and hence, off-diagonal entries of N are 2^{n-2}.
- Hence, $rank(N) = 2^n - 1$ as the first row is always 0.
- Therefore, $rank(M_{DP}) \geq 2^n - 1$ and $CF(DP) \geq n$

- Introduction
- Lower bound methods
  - The fooling set method
  - The tiling method
  - The rank method
  - The discrepancy method
- Comparison between lower bounds
- Multiparty communication complexity
- Other Communication models
- Bibliography

Discrepancy Method

For any 0/1 matrix, consider the map $b \to (-1)^b$.

Definition

The discrepancy of a rectangle $A \times B$ in a matrix M is

$$disc(R_i) = \frac{1}{2^{2n}} \left| \sum_{x \in A, y \in B} M_{xy} \right|$$

The discrepancy of a matrix $M(f)$ with rectangles $R_1, R_2, \cdots, R_l$

$$Disc(f) = \max_i disc(R_i)$$

Discrepancy Method

Lemma 8

The discrepancy of the matrix $M(f)$, denoted by $Disc(f)$, is the maximum discrepancy among all rectangles. The relation of $\chi(f)$ to $Disc(f)$ is

$$\chi(f) \geq \frac{1}{Disc(f)}$$

Discrepancy Method

Lemma 8

The discrepancy of the matrix $M(f)$, denoted by $Disc(f)$, is the maximum discrepancy among all rectangles. The relation of $\chi(f)$ to $Disc(f)$ is

$$\chi(f) \geq \frac{1}{Disc(f)}$$

Proof: If $\chi(f) \leq K$, then there exists a monochromatic rectangle having atleast $\frac{2^{2n}}{K}$ entries. Such a rectangle has discrepancy atleast $1/2^{2n} \times 2^{2n}/K = 1/K$ □

Discrepancy Method

Lemma 8

The discrepancy of the matrix $M(f)$, denoted by $Disc(f)$, is the maximum discrepancy among all rectangles. The relation of $\chi(f)$ to $Disc(f)$ is

$$\chi(f) \geq \frac{1}{Disc(f)}$$

Proof: If $\chi(f) \leq K$, then there exists a monochromatic rectangle having atleast $\frac{2^{2n}}{K}$ entries. Such a rectangle has discrepancy atleast $1/2^{2n} \times 2^{2n}/K = 1/K$ □

Note: The bound is loose. Consider the equality function. The discrepancy for the matrix $M(EQ)$ is atleast $1 - 2^{1-n}$ which gives the lower bound for $\chi(EQ)$ as 2.

Discrepancy Method

Lemma 9

For any symmetric real matrix M, the discrepancy of a rectangle $A \times B$ is atmost $\lambda_{max}\sqrt{|A||B|}/2^{2n}$, where $\lambda_{max}(M)$ is the largest eigenvalue of M.

Discrepancy Method

Lemma 9

For any symmetric real matrix M, the discrepancy of a rectangle $A \times B$ is atmost $\lambda_{max}\sqrt{|A||B|}/2^{2n}$, where $\lambda_{max}(M)$ is the largest eigenvalue of M.

Proof: For a symmetric matrix M, for vectors x, y, we know that $x^T My \leq \lambda_{max}(M)|x \cdot y|$ where $x \cdot y$ denotes the dot product of vectors x and y.

Discrepancy Method

Lemma 9

For any symmetric real matrix M, the discrepancy of a rectangle $A \times B$ is atmost $\lambda_{max}\sqrt{|A||B|}/2^{2n}$, where $\lambda_{max}(M)$ is the largest eigenvalue of M.

Proof: For a symmetric matrix M, for vectors x, y, we know that $x^T My \leq \lambda_{max}(M)|x \cdot y|$ where $x \cdot y$ denotes the dot product of vectors x and y. Also $\mathbf{1}_S$ is the characteristic vector for a subset S.

$$\mathbf{1}_S(i) = \begin{cases} 1 & \text{if } i \in S \\ 0 & \text{otherwise} \end{cases}$$

Hence, $||\mathbf{1}_S|| = \sqrt{\sum_{x \in S} 1^2} = \sqrt{|S|}$.

Discrepancy Method

Lemma 9

For any symmetric real matrix M, the discrepancy of a rectangle $A \times B$ is atmost $\lambda_{max}\sqrt{|A||B|}/2^{2n}$, where $\lambda_{max}(M)$ is the largest eigenvalue of M.

Proof: For a symmetric matrix M, for vectors x, y, we know that $x^T My \leq \lambda_{max}(M)|x \cdot y|$ where $x \cdot y$ denotes the dot product of vectors x and y. Also $\mathbf{1}_S$ is the characteristic vector for a subset S.

$$\mathbf{1}_S(i) = \begin{cases} 1 & \text{if } i \in S \\ 0 & \text{otherwise} \end{cases}$$

Hence, $||\mathbf{1}_S|| = \sqrt{\sum_{x \in S} 1^2} = \sqrt{|S|}$.
Now, for every $A, B \subseteq \{0,1\}^n$, $\sum_{x \in A, y \in B} M_{xy} = \mathbf{1}_A{}^\dagger M \mathbf{1}_B$

Discrepancy Method

The discrepancy of the rectangle $A \times B$ is

$$\frac{1}{2^{2n}}\mathbf{1}_A{}^{\dagger} M \mathbf{1}_B \leq \frac{1}{2^{2n}}\lambda_{max}(M)|\mathbf{1}_A{}^{\dagger}\mathbf{1}_B| \leq \frac{1}{2^{2n}}\sqrt{|A||B|}$$

using the Cauchy-Schwarz inequality. □

Ex: Consider the function $DP(x, y) = x \cdot y = \sum_i x_i y_i \pmod 2$

Discrepancy Method

The discrepancy of the rectangle $A \times B$ is

$$\frac{1}{2^{2n}} \mathbf{1}_A{}^{\dagger} M \mathbf{1}_B \leq \frac{1}{2^{2n}} \lambda_{max}(M) |\mathbf{1}_A{}^{\dagger} \mathbf{1}_B| \leq \frac{1}{2^{2n}} \sqrt{|A||B|}$$

using the Cauchy-Schwarz inequality. □

Ex: Consider the function $DP(x, y) = x \cdot y = \sum_i x_i y_i \pmod 2$

1. Let $M_{xy} = (-1)^{x \cdot y}$. M is the inner product function matrix with entries $\{+1, -1\}$.

Discrepancy Method

The discrepancy of the rectangle $A \times B$ is

$$\frac{1}{2^{2n}} \mathbf{1}_A{}^{\dagger} M \mathbf{1}_B \leq \frac{1}{2^{2n}} \lambda_{max}(M) |\mathbf{1}_A{}^{\dagger} \mathbf{1}_B| \leq \frac{1}{2^{2n}} \sqrt{|A||B|}$$

using the Cauchy-Schwarz inequality. □

Ex: Consider the function $DP(x, y) = x \cdot y = \sum_i x_i y_i \pmod 2$

1. Let $M_{xy} = (-1)^{x \cdot y}$. M is the inner product function matrix with entries $\{+1, -1\}$.
2. Now, for this matrix, every pair of rows (columns) are mutually orthogonal. For distinct rows x and y, $\sum_{z \in \{0,1\}^n} M(x, z).M(y, z)$. For $1/2$ of the vectors z, $M(x, z) = M(y, z)$ resulting in summand $+1$, otherwise summand -1. Hence, any two rows x and y are mutually orthogonal.

3. We will calculate each entry of M^2 as follows:

$$MM^T[i,j] = \sum_{k \in \{0,1\}^n} M(i,k).M(j,k) = \begin{cases} 2^n & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

This is because when $i = j$, the product is 1 for all k and otherwise we have, $\sum_k M(i,k).M(j,k) = 0$ as seen earlier. $M^2 = 2^n I$.

Discrepancy Method

3. We will calculate each entry of M^2 as follows:

$$MM^T[i,j] = \sum_{k \in \{0,1\}^n} M(i,k).M(j,k) = \begin{cases} 2^n & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

   This is because when $i = j$, the product is 1 for all k and otherwise we have, $\sum_k M(i,k).M(j,k) = 0$ as seen earlier. $M^2 = 2^n I$.

4. Hence, the spectral norm of the matrix given by $\sqrt{\lambda_{max}(M^T M)}$ is $\sqrt{2^n} = 2^{n/2}$.

Discrepancy Method

3. We will calculate each entry of M^2 as follows:

$$MM^T[i,j] = \sum_{k \in \{0,1\}^n} M(i,k).M(j,k) = \begin{cases} 2^n & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

   This is because when $i = j$, the product is 1 for all k and otherwise we have, $\sum_k M(i,k).M(j,k) = 0$ as seen earlier. $M^2 = 2^n I$.

4. Hence, the spectral norm of the matrix given by $\sqrt{\lambda_{max}(M^T M)}$ is $\sqrt{2^n} = 2^{n/2}$.

5. Finally, using Lemma $\text{Disc}(A \times B) \leq 2^{-3n/2}\sqrt{|A||B|}$ and the overall discrepancy $Disc(DP) \leq 2^{-n/2}$

Upper Bounding the Discrepancy

Definition

For a ± 1 valued function f,

$$\epsilon(f) = \mathbb{E}_{a_1,a_2,b_1,b_2}\left[\prod_{i=1,2}\prod_{j=1,2} f(a_i, b_j)\right]$$

Upper Bounding the Discrepancy

Definition

For a ± 1 valued function f,

$$\epsilon(f) = \mathbb{E}_{a_1,a_2,b_1,b_2}\left[\prod_{i=1,2}\prod_{j=1,2} f(a_i, b_j)\right]$$

Lemma 10

$$Disc(f) \leq \epsilon(f)^{1/4}$$

Upper Bounding the Discrepancy

Proof: Define g, h to be the characteristic functions of A and B respectively.

Upper Bounding the Discrepancy

Proof: Define g, h to be the characteristic functions of A and B respectively.

Also,

$$\begin{aligned}
\epsilon(f) &= \mathbb{E}_{a_1,a_2}\left[\mathbb{E}_{b_1,b_2}\left[\prod_{i=1,2}\prod_{j=1,2} f(a_i,b_j)\right]\right] \\
&= \mathbb{E}_{a_1,a_2}\left[\left(E_b[f(a_1,b)f(a_2,b)]\right)^2\right] \\
&\geq \mathbb{E}_{a_1,a_2}\left[g(a_1)^2 g(a_2)^2\left(E_b[f(a_1,b)f(a_2,b)]\right)^2\right] \\
&= \mathbb{E}_{a_1,a_2}\left[\left(E_b[g(a_1)g(a_2)f(a_1,b)f(a_2,b)]\right)^2\right] \\
&\geq \mathbb{E}_{a_1,a_2}\left(\left[E_b[g(a_1)g(a_2)f(a_1,b)f(a_2,b)]\right]\right)^2 \\
&\geq \left(\mathbb{E}_{a,b}[f(a,b)g(a)h(b)]\right)^4
\end{aligned}$$

Upper Bounding the Discrepancy

Proof: Define g, h to be the characteristic functions of A and B respectively.

Also,

$$\begin{aligned}
\epsilon(f) &= \mathbb{E}_{a_1,a_2}\left[\mathbb{E}_{b_1,b_2}\left[\prod_{i=1,2}\prod_{j=1,2} f(a_i,b_j)\right]\right] \\
&= \mathbb{E}_{a_1,a_2}\left[(E_b[f(a_1,b)f(a_2,b)])^2\right] \\
&\geq \mathbb{E}_{a_1,a_2}\left[g(a_1)^2 g(a_2)^2 (E_b[f(a_1,b)f(a_2,b)])^2\right] \\
&= \mathbb{E}_{a_1,a_2}\left[(E_b[g(a_1)g(a_2)f(a_1,b)f(a_2,b)])^2\right] \\
&\geq \mathbb{E}_{a_1,a_2}([E_b[g(a_1)g(a_2)f(a_1,b)f(a_2,b)]])^2 \\
&\geq (\mathbb{E}_{a,b}[f(a,b)g(a)h(b)])^4
\end{aligned}$$

Now, $Disc(f) = \mathbb{E}_{a,b\in\{0,1\}^n}[f(a,b)g(a)h(b)]$ □

Upper Bounding the Discrepancy

Ex: Consider the *DP* function $f(x, y) = x \cdot y = \sum_i x_i y_i \pmod 2$.
Now,

$$\epsilon(DP) \leq 2^{-n}$$

Proof:

$$\epsilon(f) = \mathbb{E}_{a_1, a_2}\left[(E_b[f(a_1, b)f(a_2, b)])^2\right]$$

Now, $E[X^2] \geq E[X]^2$

$$\epsilon(f) \leq \mathbb{E}_{a_1, a_2}\left[\left(E_b[f^2(a_1, b)f^2(a_2, b)]\right)\right]$$

Since the *IP* function takes values only ± 1, we have

$$\epsilon(IP) \leq \frac{2^n}{2^{2n}} = 2^{-n}$$

Hence, the $Disc(IP) \leq 2^{-n/4}$

- Introduction
- Lower bound methods
  - The fooling set method
  - The tiling method
  - The rank method
  - The discrepancy method
- Comparison between lower bounds
- Multiparty communication complexity
- Other Communication models
- Bibliography

Comparison between lower bound techniques

- Tiling method provides the strongest lower bound
- The rank method is always atleast as strong as the fooling set method

Observation 4

If a function f has a fooling set of size S, then the rank method can be used to give a lower bound of atleast $\frac{1}{2}\log_2 S$

Lemma 11

Rank of $A \bigotimes B$ over any field $\mathbb{F}$ is the product of ranks of A and B.

Comparison between lower bound techniques

- Tiling method provides the strongest lower bound
- The rank method is always atleast as strong as the fooling set method

Observation 4

If a function f has a fooling set of size S, then the rank method can be used to give a lower bound of atleast $\frac{1}{2}\log_2 S$

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \qquad B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}, \qquad A \otimes \mathrm{B} = \begin{pmatrix} a_{11}b_{11} & a_{11}b_{12} & a_{12}b_{11} & a_{12}b_{12} \\ a_{11}b_{21} & a_{11}b_{22} & a_{12}b_{21} & a_{12}b_{22} \\ a_{21}b_{11} & a_{21}b_{12} & a_{22}b_{11} & a_{22}b_{12} \\ a_{21}b_{21} & a_{21}b_{22} & a_{22}b_{21} & a_{22}b_{22} \end{pmatrix}$$

Comparison between lower bound techniques

Proof: Let $S = \{(x_1, y_1), \cdots, (x_n, y_n)\}$ be the elements of the fooling set. Consider the sub-matrix A of M_f that corresponds to the rows of $x_1, \cdots, x_n$ and columns of $y_1, \cdots, y_n$.

$$A = \begin{array}{c} \\ x_1 \\ \\ x_n \end{array} \begin{array}{c} \begin{array}{ccc} y_1 & & y_n \end{array} \\ \begin{bmatrix} 1 & \cdots & \\ \vdots & \ddots & \vdots \\ & \cdots & 1 \end{bmatrix} \end{array}$$

From the definition of fooling set, $A \odot A^T = I$ where $\odot$ represents element-wise multiplication of A.

$$|S| = rank(I) = rank(A \odot A^T) \leq rank(A \otimes A^T) \leq rank(A)^2$$

$$|S| \leq rank(M_f)^2$$

For fooling set of size $|S|$, rank gives a lower bound of atleast $\log_2(rank(M_f)) \geq \frac{1}{2} \log_2|S|$ [3]

Comparison between lower bound techniques

Low rank conjecture

There is a constant $c > 1$ such that
$C(f) = O(log(rank(M(f)))^c)$ for all f and all input sizes n,
where rank is taken over reals.

Nisan and Widgerson '94

$1/Disc(f) = O(rank(f)^{3/2})$

- Introduction
- Lower bound methods
  - The fooling set method
  - The tiling method
  - The rank method
  - The discrepancy method
- Comparison between lower bounds
- Multiparty communication complexity
- Other Communication models
- Bibliography

Multiparty Communication Complexity

- Many ways to generalize two party communication complexity to multiparty setting.
- Most interesting model is “number on the forehead model”.
- Each player has a string on his head that everyone except him can see.

Multiparty Communication Complexity

- Formally, there are k players and k string $x_1, ..., x_k$. Player i has access to all strings *except* x_i. The players are interested in computing $f(x_1, ..., x_k)$ where $f : (\{0,1\}^n)^k \to \{0,1\}$.
- A protocol for communication if agreed upon by all players beforehand. For communication, players write on a "public blackboard" that all of them can see.
- Last message sent should be $f(x_1, x_2, ..., x_k)$.
- $C_k(f)$ is number of bits exchanged by the best protocol for a function f.
- $C_k(f) \leq n + 1$

Examples of multiparty communication

- Consider a three-party model which computes following function:

$$f(x_1, x_2, x_3) = \bigoplus_{i=1}^{n} \mathrm{maj}(x_{1i}, x_{2i}, x_{3i})$$

where $x_1, x_2, x_3 \in \{0,1\}^n$.

$C_3(f) = 3$. Each player will determine majority by examining information available to her. He writes the parity of majority on blackboard. Final answer is parity of all players bit.

Lower bound for multiparty communication

- Idea corresponding to monochromatic rectangle in two-party model is cylinder intersection in k-party model.
- *Cylinder in dimension i* is a subset S of the inputs such that if $(x_1, ..., x_k) \in S$ then $(x_1, ..., x_{i-1}, x_i', x_{i+1}, ..., x_k) \in S$ for all x_i' also.

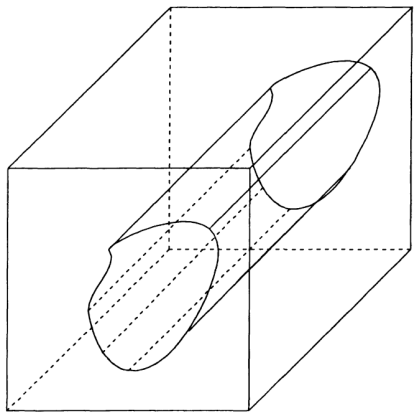


Figure 2: Example of cylinder for $k = 3$. [2]

- *Cylinder intersection* is $\cap_{i=1}^{k} S_i$ where S_i is cylinder in dimension i.

Lower bound for multiparty communication

- *Observation*: Player i's communication does not depend upon x_i, so it can be viewed as partitioning input according to cylinders in dimension i.
- At end of protocol, cube $\{0,1\}^{n^k}$ is partitioned using cylinder intersections.
- If the protocol communicates c bits, then the partition consists of at most 2^c monochromatic cylinder intersections. How?

Lower bound for multiparty communication

Lemma 12

If every partition of $M(f)$ into monochromatic cylinder intersections requires at least R cylinder intersections, then the k-party communication complexity is at least $\lceil \log_2 R \rceil$, where $M(f)$ is the k-dimensional table whose $(x_1, ..., x_k)$th entry is $f(x_1, ..., x_k)$.

Observation in previous slide and taking idea from two-party model proves this lemma. □

Discrepancy-based lower bound for Multiparty Communication

k-party Discrepancy of a function

$$Disc(f) = \frac{1}{2^{nk}} \max_{T} \left| \sum_{(a_1,\cdots,a_k)\in T} f(a_1,\cdots,a_k) \right|$$

where T ranges over all cylinder intersections

Discrepancy-based lower bound for Multiparty Communication

k-party Discrepancy of a function

$$Disc(f) = \frac{1}{2^{nk}} \max_{T} \left| \sum_{(a_1,\cdots,a_k)\in T} f(a_1,\cdots,a_k) \right|$$

where T ranges over all cylinder intersections

To upper bound the discrepancy, $\epsilon(f)$ is required. Define (k,n)-cube D to be subset of $\{0,1\}^{nk}$ consisting of 2^k points $\{a_1, a_1'\} \times \{a_2, a_2'\} \times \cdots \times \{a_k, a_k'\}$ where each $a_i, a_i' \in \{0,1\}^n$.

k-party $\epsilon(f)$

$\epsilon(f) = \mathbb{E}_D \left[\prod_{\mathbf{a}\in D} f(\mathbf{a})\right]$

Discrepancy-based lower bound for Multiparty Communication

Lemma 13

$Disc(f) \leq (\epsilon(f))^{1/2^k}$

Proof idea: Recall the proof for the lemma $Disc(f) \leq \epsilon(f)^{1/4}$ in the 2-party case.

$$\epsilon(f) = \mathbb{E}_D \left[\prod_{\mathbf{a} \in D} f(\mathbf{a}) \right]$$

For each a_i in $\mathbf{a}$, we follow the same steps as in 2-party case and arrive at

$$\epsilon(f) \geq (\mathbb{E}_{\mathbf{a} \setminus a_i} \mathbb{E}_{a_i} [f(a) g(a_i)])^2$$

Repeating the same for all 2^k values of a_i, we obtain the desired result.

- Introduction
- Lower bound methods
  - The fooling set method
  - The tiling method
  - The rank method
  - The discrepancy method
- Comparison between lower bounds
- Multiparty communication complexity
- Other Communication models
- Bibliography

Other Communication Models

- **Randomized Protocols**:
  All players have access to a shared random string r. $R(f)$ is defined as *expected* number of bits communicated by the players.
  Randomization can make a significant difference. Example: equality function has a randomized communication protocol with $O(\log n)$ complexity.
- **Non-deterministic protocols**:
  All players are provided with an additional third input z (non-deterministic guess) of some length m. Apart from this guess, protocol is deterministic.
  $f(x, y) = 1$ iff $\exists z$ that makes players output 1.
  Cost of protocol is $m +$ numbers of bits communicated.

Other Communication Models

- **Average case protocols**
- **Computing a non-Boolean function**
- **Asymmetric communication**

- Introduction
- Lower bound methods
  - The fooling set method
  - The tiling method
  - The rank method
  - The discrepancy method
- Comparison between lower bounds
- Multiparty communication complexity
- Other Communication models
- Bibliography

References

[1] Sanjeev Arora and Boaz Barak. *Computational Complexity: A Modern Approach*. 1st. USA: Cambridge University Press, 2009. ISBN: 0521424267.

[2] Eyal Kushilevitz and Noam Nisan. *Communication Complexity*. Cambridge University Press, 1996. DOI: 10.1017/CBO9780511574948.

[3] Alexander Sherstov. *Matrix Rank in Communication Complexity*. http://web.cs.ucla.edu/~sherstov/teaching/2012-winter/docs/lecture03.pdf. 2012.

Thank you