# CS 2429 - Foundations of Communication Complexity

**Lecturer: Toniann Pitassi**

## 1 Applications of Communication Complexity

There are many applications of communication complexity. In our survey article, "The Story of Set Disjointness" we give many applications via reductions to set disjointness (both 2-party as well as NOF model). Applications discussed in our survey article:

(1.) Streaming

(2.) Data Structures

(3.) Circuit Complexity

(4.) Proof Complexity

(5.) Game Theory

(6.) Quantum Computation

Below we discuss another application that is different than above. The essential difference is that it relies on the communication complexity of computing a certain *relation*, rather than a function.

### 1.1 Circuit Depth via Communication Complexity

In order to get circuit lower bounds, we need to extend our notion of 2-party communication complexity so that it can compute relations.

**Definition** A relation $R$ is a subset $R \subseteq X \times Y \times Z$

Given a relation $R$ the cc problem associated with $R$ follows:
Alice gets $x \in X$
Bob gets $y \in Y$
Alice and Bob must both compute (and output) some $z$ s.t. $(x, y, z) \in R$

A protocol for relations is the same as a protocol for functions, in each step it must specifiy which party sends a message and the value of that message.

Note that for a given relation there may be more than on $z$ satisfying the above property, Alice and Bob only need to give one such $z$. In general, lower bounds are harder to prove for relations as we need to show it is hard for Alice and Bob to compute *any* $z$.

**Definition** For any boolean function $f : \{0,1\}^n \rightarrow \{0,1\}$ and $X = f^{-1}(1)$, $Y = f^{-1}(0)$. We define $R_f \subseteq X \times Y \times \{1, 2, ...n\}$ to be the associated relation where,

- $R_f = \{(x, y, i) | x \in X, y \in Y, x_i \neq y_i\}$

$R_f$ is the set of all $(x, y, i)$ where $f(x) = 1$, $f(y) = 0$ and $x$ and $y$ differ on bit $i$. Similarly if $f$ is monotone then

- $M_f \subseteq X \times Y \times \{1, 2, ...n\}$ is the set of all $(x, y, i)$ such that $x \in X$, $y \in Y$ and $x_i = 1$, $y_i = 0$.

(Recall that for a monotone boolean function $f$, $f(x) = 1$ implies that for all $x'$ where $x'_i \geq x_i$ on every $i$, $x'$ is also a 1 of the function.)

Communication complexity lower bounds on $M_f$ give bounds on monotone circuit depth of $f$ and lower bounds on $R_f$ give circuit depth bounds for general circuits.

Let $d(f)$ and $d^{monotone}(f)$ denote the min depth of a circuit computing $f$ over $\wedge$, $\vee$, $\neg$, and the min depth of a monotone circuit computing $f$ over $\wedge$, $\vee$ respectively. In both cases the circuits must have bounded fan-in.

**Theorem 1** *(Karchmer and Widerson '80s)*

1. *For every boolean function $f : \{0,1\}^n \to \{0,1\}$, $cc(R_f) = d(f)$*
2. *For $f$ monotone, $cc(M_f) = d^{monotone}(f)$.*

For formulas it is known that $2^{d(f)} =$ formula-size$(f)$ so proving lower bounds on communication complexity of relations is also equivalent to proving formula size lower bounds.

It is a major open problem to get even super log-depth lower bounds for the general case. But for the monotone case the method above has been used to show that $NC^i_{monotone} \neq NC^{i+1}_{monotone}$ for all $i$ [see Theorem 2 and 3].

**Proof of Theorem 1 "⇒"**

Let $C$ be a circuit for $f$, depth$(C) = d$. We can assume that all the negations in the circuit are at the leaves. (If not, the negations can be pushed to the leaves without affecting depth in any circuit by repeated application of DeMorgan's laws.)

We want to use the circuit to obtain a protocol for $R_f$.

The protocol will involve Alice and Bob taking a particular path down the circuit with Alice, deciding the branch to take at $OR$ gates and Bob deciding at $AND$ gates. As long as the two parties maintain the invariant that at each subnode $v$ $C_v(x) = 1$ while $C_v(y) = 0$ then the leaf reached is a bit $i$ where $x_i \neq y_i$.

**The protocol follows:**

Starting from the top of the circuit, for each each node $v$ with children $v_L$, $v_R$

if the gate is an $OR$ Alice says 0 if $C_{v_L}(x) = 1$ and 1 otherwise.

if the gate is an $AND$ Bob says 0 $C_{v_L}(y) = 1$ and 1 otherwise.

At the end of the exchange, both Alice and bob recurse on $v_L$ if the message sent was 0 and $v_R$ if the message sent was 1.

Clearly at the top of the circuit, for any inputs $(x, y)$, $C(x) \neq C(y)$. Suppose at some point during the protocol Alice and Bob are at some inner node $v$ where $C_v(x) \neq C_v(y)$.

Case 1 $v$ is an or node.

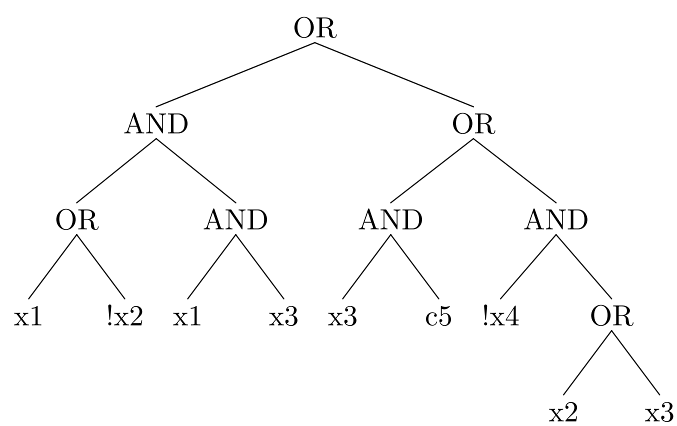
Then $C_v(y) = 0$ implies that both $C_{v_L}(y)$ and $C_{v_R}(y)$ are also 0. By choosing the subcircuit for which her input evaluates to 1, Alice ensures that the recursion continues on a subcircuit where the two inputs differ.

Case 2 $v$ is an and node.

Likewise, $C_v(x) = 1 \Rightarrow C_{v_L}(x) = C_{v_R}(x) = 1$ so by choosing the subcircuit for which his input evaluates to 0 Bob can also maintain the above invariant.

By induction, when the protocol reaches a leaf, both $A$ and $B$ know an $i$ at which their inputs differ. The total number of bits sent is bounded by the depth of the circuit. If $C$ was monotone the same protocol reaches a left where $x_i = 1$.

**Example**




Suppose Alice and Bob have inputs (01101) and (01010) respectively. Then on the circuit above the sequence of bits sent would be.

$Alice$ : 0 (go right)
$Bob$ : 1(go left)
$Alice$ : 0 (go left)

At which point they reach $x_3$ a bit on which they differ.

**Proof of Theorem 1 "$\Leftarrow$"**

Given a protocol for $R_f$ we can construct a circuit computing $f$ of bounded depth.

Consider a protocol tree $T$ for $R_f$. Convert $T$ into a circuit as follows:

1. For each node where the message is sent by Alice, replace the node with an $OR$ gate
2. For each node where the message is sent by Bob, replace the node with an $AND$ gate
3. At each leaf of the protocol tree, with associated monochromatic rectangle $A \times B$ and input bit $i$

   **Claim** Exactly one of the following hold

(a) $\forall \alpha \in A$, $\alpha_i = 1$ and $\forall \beta \in B$, $\beta_i = 0$

(b) $\forall \alpha \in A$, $\alpha_i = 0$ and $\forall \beta \in B$, $\beta_i = 1$

Assign the leaves in case (a) to be $z_i$ and and the leaves in case (b) to be $\bar{z}_i$.

Given the claim we can prove by induction that the circuit thus constructed calculates $f(z)$.

**Proof of Claim**
Let $\alpha \in A$, $\alpha_i = \sigma$. Then for every $\beta \in B$, $\beta_i = \bar{\sigma}$ which in turn implies that $\forall \alpha \in A$, $\alpha_i = \sigma$.

**Theorem 2** *(KW)*
*The monotone depth of st-connectivity is* $\Omega(\log^2 n)$.

Theorem 2 separates monotone $NC^1$ from monotone $NC^2$. A similar lower bound proved for clique separates $monotone - P$ from $monotone - NP$.

**Theorem 3** *Theorem(Raz, McKenzie)*
*For every $i$ there exists a monotone function in monotone-$NC^{i+1}$ but not in monotone-$NC^i$.*

# 2 Analyzing Communication Complexity Lower Bound Methods via Linear Programs

For the rest of this lecture, we will discuss the Jain and Klauck paper. Quantities of interest:

- Discripancy and Generalized discrepancy. (Generalized discrepancy: give a discrepancy bound for a function $g$ that is close to $f$.)
- Rectangle bound. (This is the same as corruption.) and smooth rectangle bound. (Smooth: give a rectangle bound for a function $g$ that is close to $f$.)
- (NEW) Partition bound
- Info theoretic techniques

They show that the smooth rectangle bound subsumes discrepancy bound, generalized discrepancy, and rectangle bound. So it subsumes all but the info theoretic techniques. IMPORTANT NOTE: In previous lectures we we defined discrepancy to be the inverse of the definition of discrepancy defined here. (This is just for notational convenience, so we can more easily compare the communication complexity with rectangle/discrepancy measures.)

So we have

$$exp(R(f)) \geq prt_\epsilon(f) \geq srec_\epsilon(f) \geq rec_\epsilon(f).$$

$$rec^z_\epsilon(f) \geq (1/2 - \epsilon)disc^\lambda(f) - 1/2$$

Where $prt$ is the partition bound, $srec$ is smooth rectangle bound, and $rec$ is the rectangle bound. The second inequality says that the rectangle bound dominates the discrepancy bound; thus the smooth rectangle bound also dominates the smooth-discrepancy bound.

## 2.1 LPs Primal and Dual Refresher

Recall that an LP is defined as follows. $A$ is an $m$-by-$n$ matrix, and $\vec{y}$, $\vec{b}$ are $n$-dimensional vectors.

$$\begin{array}{ll} min & \vec{b}\vec{y} \\ & A^T\vec{y} \geq c \\ & \vec{y} \geq 0 \end{array}$$

The dual is as follows, where $\vec{x}$ and $\vec{c}$ are $m$-dimensional vectors.

$$\begin{array}{ll} max & \vec{c}\vec{x} \\ & A\vec{x} \leq \vec{b} \\ & \vec{x} \geq 0 \end{array}$$

## 2.2 Rectangle Bound

The $\epsilon$-rectangle bound of $f$, denoted by $rec_\epsilon(f)$ is $max\{rec^z_\epsilon(f) : z = 0, 1\}$ where $rec^z_{\epsilon(f)}$ is the optimal value of the following linear program.

**PRIMAL**

$$\begin{array}{ll} min: & \sum_R w_R \\ & \forall (x,y) \in f^{-1}(z) : \sum_{R:(x,y)\in R} w_R \geq 1-\epsilon \\ & \forall (x,y) \in f^{-1} - f^{-1}(z) : \sum_{R:(x,y)\in R} w_R \leq \epsilon \\ & \forall R : w_R \geq 0 \end{array}$$

**DUAL**

$$\begin{array}{ll} max: & \sum_{(x,y)\in f^{-1}(z)} (1-\epsilon)\cdot \mu_{x,y} - \sum_{(x,y)\in f^{-1}-f^{-1}(z)} \epsilon \cdot \mu_{x,y} \\ & \forall R : \sum_{(x,y)\in f^{-1}(z)\cap R} \mu_{x,y} - \sum_{(x,y)\in (R\cap f^{-1})-f^{-1}(z)} \mu_{x,y} \leq 1 \\ & \forall (x,y) : \mu_{x,y} \geq 0 \end{array}$$

We will now show how a small probabilistic communication complexity protocol implies a small value of the above primal program. That is, $exp(R(F)) \geq rec_\epsilon(f)$. Suppose that we have a $c$-bit probabilistic protocol for $f$. For each choice $r$ of the public random bits, we have a partition of $M_f$ into rectangles. Let $P_{f,r}$ denote the partition of $M_f$ induced by the coin tosses $r$. Note that

$P_{f,r}$ partitions $M_f$ into at most $2^c$ rectangles. We will define $w_{z,R}$ to be the probability over all coin tosses of rectangle $R$ appearing, where $R$ is a $z$-rectangle. That is, $w_{z,R}$ will be equal to the number of copies of $R$ over all partitions $P_{f,r}$ divided by the number of coin tosses (where $R$ is a $z$-rectangle). Then $\sum_R w_{z,R}$ will be equal to the total number of $z$-rectangles over all partitions divided by the number of coin tosses, which is at most $2^c$. It is left to show that our definition of $w_R$ satisfies the constraints. First, by definition, $\forall z \forall R : w_{z,R} \geq 0$. Secondly, since the protocol has error at most $\epsilon$ on all inputs in $f^{-1}$, we have:

$$\forall (x,y) \in f^{-1} : \sum_{R:(x,y)\in R} w_{f(x),R} \geq 1 - \epsilon.$$

Finally, since for every $(x,y)$, and for every choice of random bits $r$, the protocol always outputs some $z$, we have:

$$\forall (x,y) : \sum_z \sum_{R:(x,y)\in R} w_{z,R} = 1.$$

Thus we have shown that $exp(R(f)) \geq prt_\epsilon(f)$.

### 2.2.1 Smooth (Generalized) Rectangle

The $\epsilon$-rectangle bound of $f$, denoted by $rec_\epsilon(f)$ is $max\{rec^z_\epsilon(f) : z = 0, 1\}$ where $rec^z_{\epsilon(f)}$ is the optimal value of the following linear program.

**PRIMAL**

$$\begin{aligned}
min: \quad & \sum_R w_R \\
& \forall (x,y) \in f^{-1}(z) : \sum_{R:(x,y)\in R} w_R \geq 1 - \epsilon \\
& \forall (x,y) \in f^{-1}(z) : \sum_{R:(x,y)\in R} w_R \leq 1 \\
& \forall (x,y) \in f^{-1} - f^{-1}(z) : \sum_{R:(x,y)\in R} w_R \leq \epsilon \\
& \forall R : w_R \geq 0
\end{aligned}$$

**DUAL**

$$\begin{aligned}
max: \quad & \sum_{(x,y)\in f^{-1}(z)} ((1-\epsilon)\mu_{x,y} - \phi_{x,y}) - \sum_{(x,y)\in f^{-1}-f^{-1}(z)} \epsilon \cdot \mu_{x,y} \\
& \forall R : \sum_{(x,y)\in f^{-1}(z)\cap R} (\mu_{x,y} - \phi_{x,y}) - \sum_{(x,y)\in (R\cap f^{-1})-f^{-1}(z)} \mu_{x,y} \leq 1 \\
& \forall (x,y) : \mu_{x,y} \geq 0, \phi_{x,y} \geq 0
\end{aligned}$$

It is not hard to see that any solution satisfying the primal of the above smooth rectangle LP also satisfies the primal of the rectangle LP, since there are more constraints above. This it is easy to see that $srec_\epsilon(f) \geq rec_\epsilon(f)$.

### 2.3 Discrepancy Bound

The (inverse of) the natural definition of discrepancy, denoted by $disc(f)$ is:

$$disc(f) = max\{disc^{\lambda}(f) : \lambda \text{ a distribution on } X \times Y\}$$

$$disc^{\lambda}(f) = min\{\frac{1}{|\sum_{(x,y)\in R}(-1)^{f(x,y)}\lambda_{x,y}|} : R \in \mathcal{R}\}$$

This is equivalent to the optimal value of the following linear program.

**PRIMAL**

$$
\begin{aligned}
min: \quad & \sum_R w_R + v_R \\
& \forall (x,y) \in f^{-1}(1) : \sum_{R:(x,y)\in R} w_R - v_R \geq 1 \\
& \forall (x,y) \in f^{-1}(0) : \sum_{R:(x,y)\in R} v_R - w_R \geq 1 \\
& \forall R : w_R, v_R \geq 0
\end{aligned}
$$

**DUAL**

$$
\begin{aligned}
max: \quad & \sum_{(x,y)\in f^{-1}\mu_{x,y}} \\
& \forall R : \sum_{(x,y)\in f^{-1}(1)\cap R} \mu_{x,y} - \sum_{(x,y)\in f^{-1}(0)\cap R} \mu_{x,y} \leq 1 \\
& \forall R : \sum_{(x,y)\in f^{-1}(0)\cap R} \mu_{x,y} - \sum_{(x,y)\in f^{-1}(1)\cap R} \mu_{x,y} \leq 1 \\
& \forall (x,y) : \mu_{x,y} \geq 0
\end{aligned}
$$

To see the equivalence, let $\lambda$ be the distribution giving the largest discrepancy. Then for all $R$, $\sum_{(x,y)\in R}(-1)^{f(x,y)}\lambda_{x,y} \leq 1/k$. Define $\mu_{x,y}$ to be $k\lambda_{x,y}$. Then the dual LP for $disc(f)$ will have value $k$, and it is easy to see that the constraints are satisfied.

We will now show the following Lemma, which shows that the rectangle bound dominates discrepancy (is greater than or equal to, more or less).

**Lemma 4** $rec_{\epsilon}(f) \geq (1/2 - \epsilon)disc^{\lambda}(f) - 1/2$.

**Proof** To prove this we will show that a solution to the *dual* LP for $disc^{\lambda}(f)$ of cost $k$ implies a solution to the dual LP for $rec_{\epsilon}(f)$ of cost at least $k$. The basic idea is that the rectangle bound looks at the worst case discrepancy over the 0-rectangles and also over the 1-rectangles. So if there exists a distribution that induces large 0-discrepancy, then the overall discrepancy is also large.

Let $k = disc^{\lambda}(f)$. Take the solution $\mu_{x,y}$ to the dual discrepancy LP to be the solution to the dual rectangle LP. (So set $\mu_{x,y}$ to be equal to $\mu_{x,y}$.) The discrepancy constraints in the dual LP clearly imply the constraints for the rectangle dual LP.

It is left to show that the value of this LP is at least $(1/2 - \epsilon)k - 1/2$. We have

$$\begin{aligned}
rec^{z}_{\epsilon}(f) &\geq \sum_{(x,y)\in f^{-1}(z)} (1-\epsilon)\mu_{x,y} - \sum_{(x,y)\in f^{-1}-f^{-1}(z)} \epsilon\mu_{x,y} \\
&= k(\sum_{(x,y)\in f^{-1}(z)} (1-\epsilon)\lambda_{x,y} - \sum_{(x,y)\in f^{-1}-f^{-1}(z)} \epsilon\lambda_{x,y}) \\
&= k(\sum_{(x,y)\in f^{-1}(z)} \lambda_{x,y} - \epsilon) \\
&\geq k(\frac{1}{2} - \frac{1}{2k} - \epsilon) \\
&= (\frac{1}{2} - \epsilon)k - \frac{1}{2}
\end{aligned}$$

The last inequality follows because by the definition of discrepancy, if we choose $R$ to be the entire matrix, then we have $(1 - 2\sum_{(x,y)\in f^{-1}(z)} \lambda_{x,y}) \geq 1/k$, and thus $\sum_{(x,y)\in f^{-1}(z)} \lambda_{x,y} \leq \frac{1}{2} - \frac{1}{2k}$.

### 2.3.1 Smooth (Generalized) Discrepancy

The generalized discrepancy of Boolean function, $disc(f)$ is given by the optimal value of the following linear program.

**PRIMAL**

$$\begin{aligned}
min: \quad & \sum_{R} w_R + v_R \\
& \forall (x,y) \in f^{-1}(1) : 1+\epsilon \geq \sum_{R:(x,y)\in R} w_R - v_R \geq 1 \\
& \forall (x,y) \in f^{-1}(0) : 1+\epsilon \geq \sum_{R:(x,y)\in R} v_R - w_R \geq 1 \\
& \forall R : w_R, v_R \geq 0
\end{aligned}$$

**DUAL**

$$\begin{aligned}
max: \quad & \sum_{(x,y)\in f^{-1} \mu_{x,y} - (1+\epsilon)\phi_{x,y}} \\
& \forall R : \sum_{(x,y)\in f^{-1}(1)\cap R} (\mu_{x,y} - \phi_{x,y}) - \sum_{(x,y)\in f^{-1}(0)\cap R} (\mu_{x,y} - \phi_{x,y}) \leq 1 \\
& \forall R : \sum_{(x,y)\in f^{-1}(0)\cap R} (\mu_{x,y} - \phi_{x,y}) - \sum_{(x,y)\in f^{-1}(1)\cap R} (\mu_{x,y} - \phi_{x,y}) \leq 1 \\
& \forall (x,y) : \mu_{x,y} \geq 0, \phi_{x,y} \geq 0
\end{aligned}$$

## 2.4 Partition Bound

**PRIMAL**

$$
\begin{aligned}
min: \quad & \sum_z \sum_R w_{z,R} \\
& \forall (x,y) \in f^{-1}: \sum_{R:(x,y)\in R} w_{f(x,y),R} \geq 1-\epsilon \\
& \forall (x,y): \sum_{R:(x,y)\in R} \sum_z w_{z,R} = 1 \\
& \forall z \forall R: w_{z,R} \geq 0
\end{aligned}
$$

**DUAL**

$$
\begin{aligned}
max: \quad & \sum_{(x,y)\in f^{-1}} (1-\epsilon)\mu_{x,y} + \sum_{x,y} \phi_{x,y} \\
& \forall z \forall R: \sum_{(x,y)\in f^{-1}(z)\cap R} \mu_{x,y} + \sum_{(x,y)\in R} \phi_{x,y} \leq 1 \\
& \forall (x,y) \mu_{x,y} \geq 0, \phi_{x,y} \in R
\end{aligned}
$$

We wil now show that $prt_\epsilon(f) \geq srec_\epsilon(f)$. To show this we wil show that a solution to the *primal* LP for $prt_\epsilon(f)$ of cost $c$ implies a solution to the primal LP for $srec_\epsilon(f)$ of cost at most $c$.

Fix $z' \in Z$. Let $\{w_{z,R} : z \in Z, R \in R\}$ be an optimal solution of the primal for $prt_\epsilon(f)$. We define $\forall R : w_R$ to be equal to $w_{z',R}$. Thus we have:

$$
\begin{aligned}
\forall (x,y) \in f^{-1}(z'): \sum_{R:(x,y)\in R} w_{z',R} \geq 1-\epsilon \quad &\rightarrow \quad \sum_{R:(x,y)\in R} w_R \geq 1-\epsilon \\
\forall (x,y) \in f^{-1} - f^{-1}(z'): \sum_{R:(x,y)\in R} w_{f(x,y),R} \geq 1-\epsilon \quad &\rightarrow \quad \sum_{R:(x,y)\in R} w_R \leq \epsilon \\
\forall (x,y): \sum_{R:(x,y)\in R} \sum_z w_{z,R} = 1 \quad &\rightarrow \quad \sum_{R:(x,y)\in R} w_R \leq 1
\end{aligned}
$$

Thus $w_R$ forms a feasible solution to the primal for $srec_\epsilon^z$ as desired.

### 2.4.1 Epsilon-Partition Bound

Laplante et al define a slight relaxation of the partition bound...