1. COMMUNICATION COMPLEXITY (CONTINUED)

In the last lecture, we’ve learned that for a given function \( f : \Sigma^n \times \Sigma^n \rightarrow \{0, 1\} \), the communication complexity of the function, \( cc(f) \), is the length of bits to be transmitted between two parties. A protocol is a binary tree whose leaves are labeled 0 or 1 and the length of the longest path from the root to the leaves defines \( cc(f) \). And \( \chi_f \) is the least number of monochromatic rectangles to partition \( M_f \) where \( M_f(x, y) = f(x, y) \). Each monochromatic rectangle associates with each leaf in the protocol. Hence, we obtain \( cc(f) \geq \log \chi_f \) since there are at most \( 2^{cc(f)} \) leaves. Today we will figure out the lower bound of \( \chi_f \) in two ways by examining the rank and discrepancy of the matrix \( M_f \).

2. THE MATRIX RANK BOUND

First, we will show the bound of communication complexity using the rank of \( M_f \) which is called the matrix rank bound. The rank of a matrix is the maximum number of linearly independent rows (or columns). For example, the communication matrix for the equality function \( EQ \),

\[
M_{EQ} = \begin{bmatrix}
1 & \cdots & 0 \\
\vdots & \ddots & \vdots \\
0 & \cdots & 1
\end{bmatrix},
\]

is a diagonal matrix. It is clear that each row is linearly independent to each other. Thus, the rank of the matrix is \( 2^n = N \). We are interested in the communication complexity of the following function

\[
IP(x, y) = \sum_{i=1}^{n} x_i y_i
\]

going over \( GF(2) \). We will first see the relationship between the \( cc(f) \) and the rank of \( M_f \) for an arbitrary function \( f : \Sigma^n \times \Sigma^n \rightarrow \Sigma \).

**Theorem 1.** Given a function \( f \), let \( R \) be the rank of a matrix, then \( cc(f) \geq \log (2R(M_f) - 1) \).

**Proof.** For a communication matrix \( M_f \), consider a partition of \( M_f \) into \( \chi_f \) monochromotatic rectangles. Let all 1-monochromatic rectangles\(^1\) be \( R_1, R_2, \ldots, R_\ell \) those partition all 1’s in \( M_f \). Thus, for every \( R_i \), there is a corresponding matrix \( M_i \) where \( M_i[x, y] = 1 \) if \( (x, y) \in R_i \), and 0 otherwise. For example, for a given \( M_f \) below and \( R_i = \{(3,1), (3,3), (4,1), (4,3)\} \), \( M_i \) is defined as below.

\[
M_f = \begin{bmatrix}
1 & 1 & 0 & 0 & 1 & 0 \\
1 & 1 & 0 & 1 & 0 & 1 \\
1 & 0 & 1 & 1 & 0 & 1 \\
1 & 0 & 1 & 0 & 1 & 1
\end{bmatrix}, M_i = \begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 0
\end{bmatrix}
\]

Since every element itself is a monochromatic rectangle, any matrices can always be partitioned into the number of monochromatic rectangles. Indeed, \( M_f \) can be expressed as

\(^1\)Monochromatic rectangles whose values are all 1.
$M_f = \sum_{i=1}^{\ell} M_i$ because every one should be in one of the rectangles. Hence, applying the subadditivity property of rank, $R(A + B) \leq R(A) + R(B)$, we obtain

$$R(M_f) = R\left(\sum_{i=1}^{\ell} M_i\right) \leq \sum_{i=1}^{\ell} R(M_i) = \ell.$$  

Similarly, we can compute the bound on rank by considering 0-monochromatic rectangles: Let $M'_f = 1_{N \times N} - M_f$. If the partition has $k$ 0-monochromatic rectangles, then

$$R(M'_f) \leq k.$$  

Now,

$$R(M_f) = R(1_{N \times N} - M'_f) \leq R(1_{N \times N}) + R(M'_f) = 1 + R(M'_f).$$  

Thus, $R(M'_f) \geq R(M_f) - 1$. Since $\chi_f = \ell + k$, the lower bound of $\chi_f$ in terms of rank of matrix is

$$\chi_f = \ell + k \geq R(M_f) + R(M'_f) \geq 2R(M_f) - 1.$$  

Therefore, we finally showed that,

$$cc(f) \geq \log(2R(M_f) - 1).$$

\[\square\]

**Corollary 1.** $cc(f) \geq \log R(M_f)$.

**Proof.** Since $cc(f) \geq \log \ell_1 \geq \log R(M_f)$. Or $cc(f) \geq \log (R(M_f) + R(M'_f)) \geq \log (R(M_f))$ because the rank of a matrix is non-negative. \[\square\]

Now, we will examine $R(M_{IP})$ to determine the lower bound of $cc(IP)$. Instead of using $M_{IP}$, a 0-1 matrix, we will use a matrix which maps 0 and 1 to 1 and -1, respectively\footnote{For a function $f : \Sigma^n \times \Sigma^n \to \{0, 1\}$, the new function $f'$ can be defined as $f' : \Sigma^n \times \Sigma^n \to \{1, -1\}$ and can be expressed as $f'(x, y) = (-1)^{f(x, y)}$.}. For example, for $n = 1$ or $n = 2$,

$$H_{2^1 \times 2^1} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \quad H_{2^2 \times 2^2} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}.$$  

This matrix is called Hadamard matrix which is a square matrix where its elements are $\pm 1$ and its rows are linearly independent. In addition, we can see that the block structure is repeating as follows,

$$\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \otimes \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 \end{bmatrix}.$$  

This operation is called a tensor product (In the context of matrix, it is often called Kronecker product.)

**Claim 1.** $R(H) = N$ where $N$ is the dimension of $H$. 

2For a function $f : \Sigma^n \times \Sigma^n \to \{0, 1\}$, the new function $f'$ can be defined as $f' : \Sigma^n \times \Sigma^n \to \{1, -1\}$ and can be expressed as $f'(x, y) = (-1)^{f(x, y)}$.  

2
Proof. We will prove the claim by showing that any two rows of $H$ are orthogonal by showing that the dot product of them is 0. To begin with, let $H_x$ and $H_z$ denote row $x$ and $z$ of $H$, respectively. The matrix $H$ looks like

$$H_{N \times N} = \begin{pmatrix} y^1 & y^2 & \cdots & y^i & \cdots & y^N \\ \downarrow & \downarrow & & \downarrow & & \downarrow \\ \vdots & & \ddots & & \vdots \\ x & \rightarrow & (1)^{IP(x,y^1)} & \cdots & (1)^{IP(x,y^i)} & \cdots & (1)^{IP(x,y^N)} \\ \vdots & & \ddots & & \vdots \\ z & \rightarrow & (1)^{IP(z,y^1)} & \cdots & (1)^{IP(z,y^i)} & \cdots & (1)^{IP(z,y^N)} \\ \vdots & & \ddots & & \vdots \end{pmatrix},$$

where $H[x,y^i] = (1)^{IP(x,y^i)}$. Then, $H_x, H_z \in \{-1, 1\}^N$ can be represented as

$$H_x = \left( (1)^{IP(x,y^1)}, (1)^{IP(x,y^2)}, \ldots, (1)^{IP(x,y^N)} \right),$$

$$H_z = \left( (1)^{IP(z,y^1)}, (1)^{IP(z,y^2)}, \ldots, (1)^{IP(z,y^N)} \right).$$

So, the dot product of $H_x$ and $H_z$ is expressed as

$$\sum_{j=1}^{N} (1)^{IP(x,y^j)} (1)^{IP(z,y^j)} = \sum_{j=1}^{N} (1)^{IP(x,y^j) + IP(z,y^j)} = \sum_{j=1}^{N} (1)^{IP(x \oplus z, y^j)}$$

because common bits between $x$ and $z$ counted twice, only different bits are considered. Since $x$ and $z$ are not the same, there are at least one 1-bit in the $x \oplus z$. Let $k$ be the number of different bits. Now, we rewrite the term as

$$\sum_{j=1}^{N} (1)^{IP(x \oplus z, y^j)} = |\{y| IP(x \oplus z, y) \text{ is even}\}| - |\{y| IP(x \oplus z, y) \text{ is odd}\}|$$

$$= |\{y| \text{even # of 1's in the } k \text{ bits}\}| - |\{y| \text{odd # of 1's in the } k \text{ bits}\}|$$

$$= 0.$$ 

Therefore, the dot product is 0 and two rows are orthogonal to each other. Finally, we can conclude that all rows are linearly independent and the rank of $H$ is equal to the number of rows, $R(H) = N$. \qed

Corollary 2. $cc(IP) \geq n$ because $cc(IP) \geq \log R(M_{IP}) = \log R(H) = \log N = n$.

We’ve showed the matrix rank bound for $\chi_f$. In addition, using the bound, we’ve proved that the function $IP$ over $GF(2)$ requires at least $n$ bits to communicate.

3. The Discrepancy Bound

Now, we will introduce the discrepancy of a matrix which will define new bound for $\chi_f$. 

"
Definition 1. The discrepancy of a rectangle $R$ is defined as

$$\text{Disc} (R) = \frac{1}{2^{2n}} \left| \sum_{x \in A, y \in B} M(x, y) \right|$$

where $R = A \times B$. The measure ranges from 0 to 1 and represents how unbalanced a rectangle is. If it consists of all -1’s or 1’s, then the discrepancy is high and the equal number of -1’s and 1’s yields 0. In addition, the discrepancy of a matrix is defined as

$$\text{Disc} (M) = \max_R \text{Disc} (R).$$

Hence, the discrepancy of a matrix is high if there is a large unbalanced rectangle.

For a function $f$ let $P$ be a protocol with has $\ell$ leaves. This partitions $M_f$ into $\ell$ monochromatic rectangles $R_i$, $1 \leq i \leq \ell$. Since all entries in $R_i$ is of same value, this rectangle is very unbalanced. It’s easy to see that the discrepancy of a monochromatic rectangle is proportional to its size, $\text{Disc} (R_i) = 2^{-2n} |R_i|$. Thus, a low matrix discrepancy ensures that all monochromatic rectangles are small. As a consequence, a low discrepancy means high number of monochromatic rectangles, thus, the communication complexity is high as well.

Claim 2. $\chi_f \geq \text{Disc} (M_f)^{-1}$

Proof. Let $\chi_f$ be $k$, so that we can partition $M_f$ into $k$ monochromatic rectangles. It means that there exists a monochromatic rectangle whose size is at least $\frac{2^{2n}}{k}$. Thus, the following holds.

$$\text{Disc} (R) \geq \frac{2^{2n}}{k} / 2^{2n} = \frac{1}{k} = \frac{1}{\chi_f}.$$ 

By the definition of discrepancy, $\chi_f \geq \text{Disc} (R)^{-1} \geq \text{Disc} (M_f)^{-1}$. \hfill $\square$

Corollary 3. $cc(f) \geq -\log (\text{Disc} (M_f))$.

Lastly, we will look at the upper bound of $\text{Disc} (R)$ for the function $IP$.

Theorem 2. For the function $IP$, let $R$ be a rectangle such that $R = S \times T$, then $\text{Disc} (R) \leq \frac{1}{N^2} \sqrt{|S| |T| |N|} = \frac{1}{N^2} \sqrt{|S| |T|}$. The product of size of two sets $S$ and $T$ is equal to the size of rectangle $R$, $|S| |T| = |R|$.

Proof. For the function $IP$, the discrepancy of a rectangle $R$ is $\frac{1}{2^2} \left| \sum_{x \in S, y \in T} H(x, y) \right|$. The term $\left| \sum_{x \in S, y \in T} H(x, y) \right|$ is the absolute value of sum of elements in the rectangle $R$. It is absolute value of sum of specified values in the matrix below.

$$H = \begin{bmatrix}
\vdots & \vdots & \vdots & \vdots \\
\cdots & -1^{(s_1, t_1)} & \cdots & -1^{(s_1, t_1)} \\
\vdots & \vdots & \vdots & \vdots \\
\cdots & -1^{(s_2, t_1)} & \cdots & -1^{(s_2, t_1)} \\
\vdots & \vdots & \vdots & \vdots \\
\cdots & -1^{(s_i, t_1)} & \cdots & -1^{(s_i, t_1)} \\
\vdots & \vdots & \vdots & \vdots \\
\cdots & -1^{(s_j, t)} & \cdots & -1^{(s_j, t)} \\
\vdots & \vdots & \vdots & \vdots \\
\cdots & \vdots & \vdots & \vdots \\
\end{bmatrix}$$
Let $H_x$ denote the row indexed by $x$. Given $T$, let $I_T$ be the characteristic vector of $T$ (e.g., if $T = \{3, 5, 7\}$, then $I_T = 00101010\cdots0$.) Then, we rewrite the expression,
\[
\left| \sum_{x \in S, y \in T} H(x, y) \right| = \left| \sum_{x \in S} \langle H_x, I_T \rangle \right|
\]
where $\langle H_x, I_T \rangle$ is the dot product of $H_x$ and $I_T$. Below we use the Cauchy-Schwartz inequality: For any two vectors $u$ and $v$, $|\langle u, v \rangle| \leq ||u||_2 \cdot ||v||_2$ where $||\cdot||_2$ is the $\ell^2$ norm of a vector.

\[
\left| \sum_{x \in S} \langle H_x, I_T \rangle \right| = \left| \left\langle \sum_{x \in S} H_x, I_T \right\rangle \right|
\]
by distributive law
\[
\leq \left| \sum_{x \in S} H_x \right| \cdot ||I_T||_2 \text{ by the Cauchy-Schwarz inequality}
\]
\[
= \left| \sum_{x \in S} H_x \right| \cdot \sqrt{|T|},
\]
because $I_T$ has $|T|$ number of 1’s and $N - |T|$ 0’s, $||I_T||_2 = \sqrt{1^2 + 1^2 + \cdots + 1^2} = \sqrt{|T|}$. Then,
\[
\left| \sum_{x \in S} H_x \right|_2 = \sqrt{\left\langle \sum_{x \in S} H_x, \sum_{x \in S} H_x \right\rangle}
\]
by definition
\[
= \sqrt{\sum_{x \in S} \sum_{x' \in S} \langle H_x, H_{x'} \rangle}
\]
by distributive law
\[
= \sqrt{\sum_{x \in S} \langle H_x, H_x \rangle}
\]
by linear independency
\[
\leq \sqrt{\sum_{x \in S} N}
\]
\[
= \sqrt{|S| N}.
\]
Since $S$ and $T$ is at most $N$, we finally get
\[
\text{Disc } (H) \leq \frac{\sqrt{|S| |T| N}}{N^2} \leq \frac{\sqrt{NNN}}{N^2} = N^{-1/2} = 2^{-n/2}.
\]

\textbf{Corollary 4.} $cc(\text{IP}) \geq n/2$ because $cc(\text{IP}) \geq -\log(\text{Disc } (M_{\text{IP}})) \geq \log(2^{n/2}) = \frac{n}{2}$.\hfill \Box