15 Communication Complexity (continued)

15.1 Randomized and Distributional Complexity

Let \( f : \{0,1\}^{2n} \rightarrow \{0,1\} \), and define
\[
C(f) = \min_{\mathcal{P}} \max_{(x,y)} |\mathcal{P}(x,y)|, \tag{15.1.1}
\]
where \( \mathcal{P} \) is over all protocols that compute \( f \), and \( |\mathcal{P}(x,y)| \) is the message string. Note that \( C(f) \leq n \).

Correction: the theorem from last time that states \( C(\beta) \geq \log \text{rk}(M_f) \) where \( M_f = (f(x,y))_{2^n \times 2^n} \) was incorrectly attributed to Yau last time: the correct attribution is Mehlhorn-Schmidt.

The Randomized Communication Complexity of \( f \) is denoted \( C_\varepsilon(f) \), and is defined by the same equation (15.1.1), except that \( \mathcal{P} \) ranges over protocols that compute \( f \) with some error allowed, of probability \( \leq \varepsilon \). More precisely, we require that \((\forall x, y)(\text{Pr}(\text{error}) \leq \varepsilon)\).

Distributional Complexity: The randomization over inputs
\[
D_{\varepsilon,\mu}(f) = \min \left\{ C(f^*) \left| \Pr_\mu(f^*(x,y) \neq f(x,y)) \leq \varepsilon \right. \right\} \tag{15.1.2}
\]

Lemma 15.1.1. \( \forall \mu, R_{\varepsilon}(F) \geq D_{\varepsilon,\mu}(F) \).
In fact, \( R_\varepsilon(f) = \max \mu \) \( D_{\varepsilon,\mu}(f) =: D_\varepsilon(f) \). (We won’t use this.)
\[
IP_\varepsilon(x, y) = \sum x_i y_i \pmod{2}.
\]

**Theorem 15.1.2.** \( C_\varepsilon(IP_X) = \Omega(n) \) (i.e. \( \geq c \cdot n \)).

Let’s switch notation: let \( f : \Omega \rightarrow \{\pm 1\} \), with \( S \subset \Omega \). The (normalized) discrepancy of \( f \) over \( S \) is
\[
\Delta(f, S) = \frac{\left| \sum_{x \in S} f(x) \right|}{|\Omega|}.
\]

If \( f \) is homogeneous on \( S \) then \( \Delta(f, S) = \frac{|S|}{|\Omega|} \).

The discrepancy of \( f \) is \( \Delta(f) = \max_{S \in \mathcal{F}} \Delta(f, S) \) where \( \mathcal{F} \) is a particular family of subsets of \( \Omega \).

Now, recall that our domain is \( \Omega = \{0, 1\}^n \times \{0, 1\}^n \). We wanted to prove the

**Theorem 15.1.3.**
\[
C_\varepsilon(f) \geq \log \left( \frac{1 - 2\varepsilon}{\Delta_{\Box}(f)} \right),
\]
where the \( \Box \) is over all rectangles (in the big \( 2^n \times 2^n \)-rectangle of inputs). (note the numerator was originally \( \frac{1}{2} - \varepsilon \) and was then changed.)

To bound \( C_\varepsilon \) from below, we estimate \( D_{\varepsilon,\mu} \) with respect to the uniform distribution \( \mu \). Let \( s := D_{\varepsilon,\mu} \).

Now, \( \Delta := \Delta_{\Box}(f) \), i.e., for every rectangle: say, label the rectangles \( R_j \), of sizes \( k_j \times \ell_j \); one has
\[
\left| \sum_{R_j} f(x, y) \right| \leq \Delta \cdot 2^{2n}.
\]

So \( P \) is a deterministic protocol with \( \leq \varepsilon \) fraction of error, and the message length is \( s \). If we have a cover by \( 2^s \) rectangles, homogeneous with respect to a fraction \( f^* \approx \varepsilon f \), let’s say each \( R_j \) has \( a_j \) 1’s and \( b_j \) -1’s, with \( a_j \geq b_j \): the number of errors is \( b_j \).
Now $0 \leq a_j - b_j \leq \Delta \cdot 2^{2n}$, and $a_j + b_j = k_j \ell_j$. So, adding these, $2b_j \geq k_j \ell_j - \Delta 2^{2n}$.

So

\[ 2\varepsilon 2^{2n} \geq 2 \cdot \text{total error} \geq 2^{2n} - 2^s \cdot \Delta \cdot 2^{2n}, \quad (15.1.5) \]

\[ 2\varepsilon \geq 1 - 2^s \Delta \quad (15.1.6) \]

\[ 2^s \Delta \geq 1 - 2\varepsilon \quad (15.1.7) \]

\[ 2^s \geq \frac{1 - 2\varepsilon}{\Delta} \quad (15.1.8) \]

\[ s \geq \log \frac{1 - 2\varepsilon}{\Delta}. \quad (15.1.9) \]

Now to complete the proof we need to learn about Hadamard matrices.

### 15.2 Hadamard Matrices

We have the following claim about the discrepancy of $IP_x$ over rectangles:

**Claim 15.2.1.** $\pm 1$-representation of $IP_-$ matrix is Hadamard.

**Definition 15.2.2.** A $N \times N$-matrix is Hadamard if

1. every entry is $\pm 1$

2. rows are orthogonal, i.e. $AA^T = N \cdot I = \begin{pmatrix} N & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & N \end{pmatrix}$

Exercises:

**Exercise 15.2.3.** $\text{rk}(A \otimes B) = \text{rk}(A) \cdot \text{rk}(b)$.

**Exercise 15.2.4.** If $k_1 = \ell_1$ and $k_2 = \ell_2$ and eigenvalues of $A$ are $\lambda_1, \ldots, \lambda_{k_1}$ and of $B$ are $\mu_1, \ldots, \mu_{k_2}$ (full lists counting multiplicities over $\mathbb{C}$), then the eigenvalues of $A \otimes B$ are $\lambda_i \mu_j$.

**Exercise 15.2.5.** If $A, B$ are Hadamard then $A \otimes B$ is Hadamard.

**Exercise 15.2.6.** $S_n := \bigotimes^n \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ is a $2^n \times 2^n$ Hadamard matrix. This is called the $2^n \times 2^n$ Sylvester matrix.
Exercise 15.2.7. Prove: if $\exists$ an $N \times N$ Hadamard matrix then $N = 2$ or $4 \mid N$.

Conjecture 15.2.8. This is also sufficient: if $4 \mid N$ then there exists an $n \times N$ Hadamard matrix.

Exercise 15.2.9. If $p \equiv 1 \pmod{4}$ is prime, then there exists a Hadamard matrix of size $(p-1) \times (p-1)$. Hint: use the quadratic character (Legendre symbol) modulo $p$.

One question is, what is the density of Hadamard numbers (numbers for which a Hadamard matrix of that size exists).

Bad fact: the density of the currently known Hadamard numbers is 0.

Here, $\text{density}(A) := \lim_{n \to \infty} \frac{|A \cap \{1, \ldots, n\}|}{n}$. But the conjectural (15.2.8) density is $1/4$.

Lemma 15.2.10. (J.H. Lindsey’s Lemma): If $H$ is an $N \times N$ Hadamard matrix and $R$ is a $k \times \ell$ rectangle in $H$, then

$$\left| \sum_{R} h_{ij} \right| \leq \sqrt{k\ell N}, \quad k, \ell \leq N. \quad (15.2.1)$$

Corollary 15.2.11.

$$\Delta \leq \frac{N^{3/2}}{N^2} = \frac{1}{\sqrt{N}} \quad (15.2.2)$$

Now, $C_\varepsilon(f) \geq \log_2 \frac{1-2\varepsilon}{2^{2n}} = \log_2(1-2\varepsilon) + \frac{n}{2} = \Omega(n)$, assuming that $M_f(\pm1)$ is Hadamard.

We have that $M_n = ((-1)^{|A \cap B|})_{2^n \times 2^n}$ for $A, B \subset \{1, \ldots, n\}$. Note that $|A \cap B|$ can be reduced modulo two here because it’s an exponent of $-1$.

Claim 15.2.12.

$$M_{n+1} = \begin{pmatrix} M_n & M_n \\ -M_n & M_n \end{pmatrix}. \quad (15.2.3)$$

Recall from Exercises 15.2.6 that $\otimes^n \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = S_n$ is called the $2^n \times 2^n$ Sylvester matrix.

Claim 15.2.13. $M_n$ is Hadamard.
Exercise 15.2.14. (a hint for Exercise 15.2.6) \[ \sum_{A} (-1)^{|A \cap B_1|} \cdot (-1)^{|A \cap B_2|} = \delta_{B_1,B_2}. \]

Now, let’s end with some magic. First note that if \( A \) is orthogonal and \( x \in \mathbb{R}^n \), then \( \|Ax\| = \|x\| \). Now, we have \((AB)^T = B^T A^T\), so \((AA^T)^T = AA^T\).

Let’s suppose that \( AA^T = I \). Does it follow that \( A^T A = I \)? In general it is not obvious that if \( AB = I \) then \( BA = I \). To do this we really only need to prove that the existence of a right inverse is equivalent to the existence of a left inverse. This is because, in a semigroup, \( ab = 1 \) and \( ca = 1 \) imply \( b = c \). Existence of a right inverse is the same as the rows being linearly independent, while the existence of a left inverse is the same as the columns being linearly independent. So if the matrix is square, having a right inverse is equivalent to having a left inverse (for finite-dimensional matrices). Example: multiplying by \( x \) or differentiating in the space of polynomials in \( x \).

Finally, we need to prove Lindsey’s lemma:

**Proof.** (Lindsey’s Lemma): We will need Cauchy-Schwarz (note that Schwarz has a “c” and no “t” so it’s a German Schwarz):

**Theorem 15.2.15.** (Cauchy-Schwarz): \[ |x \cdot y| \leq \|x\| \cdot \|y\|. \]

We know that \( \|Ax\|^2 = (Ax)^T(Ax) = x^T A^T A x = x^T x = \|x\|^2 \).

Now we want to know the sum of the entries that fall in a rectangle \( R \), i.e. \( \sum_{R} h_{i,j} = a^T H b \), where \( a \) has a 1 in the entries corresponding to the rows used by \( R \) and \( b \) has a 1 in the entries corresponding to the columns used by \( R \) (we put \( a \) and \( b \) as column vectors). So \( |a^T H b| \leq \|a^T\| \cdot \|H b\| = \sqrt{k} \|H b\| \).

Now \( HH^T = N \cdot I \), and \( \frac{1}{\sqrt{N}} H \) is orthogonal. So \( \| \left( \frac{1}{\sqrt{N}} H \right) b \| = \| b \| \) and \( \| H b \| = \sqrt{N} \| b \| = \sqrt{N} \). This is a magical proof: note that 99% of the magic is in the Cauchy-Schwarz.

This completes the proof of Theorem 15.1.3.

### 15.3 Indian Head Poker

Let’s move on to something different: recall Indian Head Poker: three people each put a card on their respective foreheads so that they can see the other two cards but not their own. Then they bet on whose card will win. So we have a function \( f(x, y, z) \), with \( C(f) \leq n \), which has to do with the cards
(e.g. is someone’s card higher than the other, etc.). Let’s find an explicit function \( f \) such that \( C(f) = \Omega(n) \). Finding explicit functions is usually what people are most interested in (random functions cannot be computed).

Suppose \( f : \{0, 1\}^{3n} \to \{0, 1\} \). We want to find a function that’s difficult to compute: one is the Generalized Inner Product (GIP): \( GIP(x, y, z) = \sum x_i y_i z_i \mod 2 \).

What other examples are there? For two players one has

\[
\text{Exercise 15.3.1.} \quad C_\epsilon \left( \left( \frac{x + y}{p} \right) \right) = \Omega(n),
\]

where the \((-)\) here is the Legendre symbol.

\[ \text{Theorem 15.3.2.} \quad C_\epsilon \left( \left( \frac{x + y + z}{p} \right) \right) = \Omega(n). \]

This has to do with the quadratic character. One also has \( C_\epsilon(GIP) = \Omega(n) \).

For \( k \) players,

\[
C(GIP_k) = \Omega \left( \frac{n}{4^k} \right),
\]

(15.3.2)

and

\[
C(QCH) = \Omega \left( \frac{n}{2^k} \right).
\]

(15.3.3)

Note that for both of these, they are only difficult to communicate if \( k \ll \log(n) \). We don’t know any functions that are difficult to compute if \( k \sim \log(n) \).

\[ \text{Question 15.3.3.} \ (\text{Open question}): \text{Find an explicit } f \text{ with } C_k(f) > (\log n)^2 \text{ with } k > \log n \text{ players.} \]

Note: the proof of \( C(GIP_k) \) involves repeated Cauchy-Schwarz. The proof of \( C(QCH) \) is an inductive proof using Cauchy-Schwarz whose base case uses Weil’s character estimates.