Electronic Colloquium on Computational Complexity, Report No. 111 (2024)

An XOR Lemma for Deterministic Communication Complexity

Siddharth Iyer*
siyer@cs.washington.edu

Anup Rao*
anuprao@cs.washington.edu

July 1, 2024

Abstract

We prove a lower bound on the communication complexity of computing the n-fold xor of an arbitrary function f, in terms of the communication complexity and rank of f. We prove that $D(f^{\oplus n}) \geq n \cdot \left(\frac{\Omega(D(f))}{\log \mathsf{rk}(f)} - \log \mathsf{rk}(f)\right)$, where here $D(f), D(f^{\oplus n})$ represent the deterministic communication complexity, and $\mathsf{rk}(f)$ is the rank of f. Our methods involve a new way to use information theory to reason about deterministic communication complexity.

1. Introduction

How is the complexity of computing a Boolean function f on 1 input related to the complexity of computing f on n inputs? In this work, we give new lower bounds for the deterministic communication complexity of computing f on n inputs, making the first progress on this question in many years. We refer the reader to the textbooks [KN97, RY20] for the broader context surrounding these problems and the model of communication complexity.

Given a function $f : \mathcal{X} \times \mathcal{Y} \to \{0, 1\}$, define the functions

$$f^n(x_1, \ldots, x_n, y_1, \ldots, y_n) = f(x_1, y_1), f(x_2, y_2), \ldots, f(x_n, y_n),$$
$$f^{\oplus n}(x_1, \ldots, x_n, y_1, \ldots, y_n) = f(x_1, y_1) \oplus f(x_2, y_2) \oplus \ldots \oplus f(x_n, y_n).$$

So, f^n computes f on n distinct inputs, and $f^{\oplus n}$ computes the parity of the outputs of f. Because every protocol computing f^n is also a protocol for computing $f^{\oplus n}$, the complexity of computing $f^{\oplus n}$ can only be smaller. An important example to keep in mind is when x, y are bits and $f(x, y) = x \oplus y$. Then the communication complexity of f and $f^{\oplus n}$ are both 2, so there is no increase in the complexity of the xor for such functions.

The communication complexity of a function f is related to the number of *monochromatic rectangles* needed to cover the inputs to f. A monochromatic rectangle is a pair $A \subseteq \mathcal{X}, B \subseteq \mathcal{Y}$ such that f is constant when restricted to $A \times B$. Let $D(f)$ denote the communication complexity of f, and let $C(f)$ denote the minimum number of monochromatic rectangles needed to cover the inputs of f. It is a standard fact that $D(f) \geq \log C(f)$. Prior to our work, the best known result concerning the complexity of computing these functions was proved by Feder, Kushilevitz, Naor and Nisan [FKNN95], who showed that when $\sqrt{D(f)} > \log \log(|\mathcal{X}| \cdot |\mathcal{Y}|)$, $D(f^n)$ grows with n:

*Supported by NSF award 2131899.

ISSN 1433-8092

Theorem 1 ([FKNN95]). $D(f^n) \geq \log C(f^n) \geq n \cdot (\sqrt{D(f)} - \log\log(|\mathcal{X}| \cdot |\mathcal{Y}|))$.

Another important parameter of f is its rank. The function f can be viewed as a Boolean matrix M whose (x, y)'th entry is $(-1)^{f(x,y)}$. We write $\mathsf{rk}(f)$ to denote the rank of this matrix. Because M has ± 1 entries, it can have at most $2^{\mathsf{rk}(f)}$ distinct rows and at most $2^{\mathsf{rk}(f)}$ distinct columns. This observation leads to the following corollary of Theorem 1:

Corollary 2. $D(f^n) \geq n \cdot (\sqrt{D(f)} - \log \mathsf{rk}(f) - 1)$.

There have been a number of results concerning the randomized communication complexity of f^n and $f^{\oplus n}$ in recent years. These results rely on definitions from information complexity and simulations of protocols that have small information complexity. See [SK87, Raz92, Raz95, CSWY01, BJKS02, JRS03, BBCR10, HJMR10, BR11, Bra15, Kol16, She18, JPY12, BRWY13b, BRWY13a, Yu22, GKR16, RR15, IR24]. However, communication complexity is a model where the deterministic and randomized complexity can be quite far from each other. For example, the randomized communication complexity of the equality function is a constant, but there is no deterministic protocol that beats the performance of the trivial protocol.

In fact, a number of connections between the model of communication complexity and other models of computation are only meaningful when using deterministic protocols. A good example is the connection between circuit depth and communication complexity observed by Karchmer and Wigderson [KW90]. The randomized communication complexity of every Karchmer-Wigderson game is small, because the game can efficiently be solved by hashing. So, lower bounds on circuit depth can only be obtained by studying deterministic communication complexity. Karchmer, Raz and Wigderson [KRW95] conjectured that the communication complexity of this problem increases when the function is composed with itself. Recently, there have been attempts toward this conjecture and on understanding Karchmer-Wigderson games using ideas from information theory [GMWW17, MW19]. If the conjecture is true, this would imply that there is no way to simulate every polynomial time algorithm in logarithmic time with a parallel algorithm. Achieving such tantalizing results motivates us to study the questions about deterministic communication complexity we consider in this paper.

Before the present paper, techniques from information theory had not led to results about the deterministic communication complexity of f^n or $f^{\oplus n}$. That is because known methods to simulate protocols with small information lead to simulations that introduce errors, even if the protocols being simulated do not make errors. In the present paper, we use information theory to obtain results about deterministic communication complexity without introducing any errors. That is the key technical contribution of our work. Our proofs are short, but they circumvent a barrier to applying information theory in the setting of deterministic communication protocols.

Lovász and Saks [LS88] conjectured that there is a constant c such that $D(f) \leq (\log \mathsf{rk}(f))^c$. This is called the *log-rank* conjecture. To date, the best known upper bound is $D(f) \leq \sqrt{\mathsf{rk}(f)}$ [Lov14, ST23], and it is known that there are f with $D(f) \geq (\log \mathsf{rk}(f))^{2-o(1)}$ [GPW18]. Recall that $D(f) \geq \log \mathsf{rk}(f)$. Our main result gives stronger lower bounds when $D(f) \gg (\log \mathsf{rk}(f))^2$:

Theorem 3. $D(f^{\oplus n}) \geq \log C(f^{\oplus n}) \geq n \cdot \left(\frac{\Omega(D(f))}{\log \mathsf{rk}(f)} - \log \mathsf{rk}(f)\right)$.

In comparison to Theorem 1, our result gives lower bounds even for computing the xor $f^{\oplus n}$. The key new step of our proof is the following theorem, whose proof uses the sub-additivity of entropy in an essential way:

Theorem 4. *If $f^{\oplus n}$ has a monochromatic rectangle of size 2^k, then f has a monochromatic rectangle of size $2^{k/n-2}$.*

The above theorem allows us to use a monochromatic rectangle of large density in $f^{\oplus n}$ to find a monochromatic rectangle of even larger density in f. Combined with some reasoning about the rank of the function, we are able to use Theorem 4 to obtain a deterministic protocol that proves Theorem 3. In the rest of this paper, we give the details of the proofs of these two theorems.

2. Preliminaries and Notation

For a variable $X = X_1, \ldots, X_n$, we write $X_{<i}$ to denote $X_1, \ldots, X_{i-1}$. We define $X_{>i}$ similarly. All logarithms are taken base 2. We recall some basic definitions regarding entropy of random variables. Let A be a random variable distributed according to $p(a)$. The entropy of A is defined as

$$H(A) := \mathop{\mathbb{E}}_{p(a)}\left[\log \frac{1}{p(a)}\right].$$

Proposition 5. *For any random variable A with finite support, we have $H(A) \leq \log|\mathsf{supp}(A)|$, with equality if A is distributed according to the uniform distribution.*

If A and B are two jointly distributed random variables distributed according to $p(ab)$ then the entropy of A conditioned on B is defined as

$$H(A|B) := \mathop{\mathbb{E}}_{p(a,b)}\left[\log \frac{1}{p(a|b)}\right].$$

The entropy of jointly distributed random variables satisfy the chain rule:

$$H(A, B) = H(A) + H(B|A).$$

Additionally, it is known that the conditional entropy of a random variable cannot exceed its entropy.

Lemma 6. *For any two jointly distributed random variables, A, B, we have $H(A|B) \leq H(A)$, with equality if A and B are independent.*

We need the following basic fact about rank:

Proposition 7. *For any two matrices A_1 and A_2, we have $\mathsf{rk}(A_1 + A_2) \leq \mathsf{rk}(A_1) + \mathsf{rk}(A_2)$.*

We need the following lemma that shows that a protocol with a small number of leaves can be computed by a protocol with small communication (see [RY20], Theorem 1.7).

Lemma 8. *If π is a deterministic protocol with ℓ leaves, there exists a deterministic protocol computing $\pi(x, y)$ with communication at most $\lceil 2\log_{3/2}\ell\rceil$.*

3. Proof of Theorem 4

Let R be a monochromatic rectangle for $f^{\oplus n}$ of size 2^k, and let $(X,Y) \in R$ be uniformly random. Because R is a rectangle, X and Y are independent. Using the chain rule, we get

$$\begin{aligned}
k = H(XY) = H(X) + H(Y) & & \text{(because } X, Y \text{ are independent)}\\
&= \sum_{i=1}^{n} H(X_i|X_{<i}) + H(Y_i|Y_{>i}) & \text{(by the chain rule)}\\
&= \sum_{i=1}^{n} H(X_i|X_{<i}Y_{>i}) + H(Y_i|X_{<i}Y_{>i}X_i) & \text{(because } X, Y \text{ are independent)}\\
&= \sum_{i=1}^{n} H(X_iY_i|X_{<i}Y_{>i}). & \text{(by the chain rule)}
\end{aligned}$$

This implies there exist $i, x_{<i}, y_{>i}$ such that

$$H(X_iY_i|x_{<i}y_{>i}) \geq k/n.$$

Define the random variables $U = f(x_1, Y_1) \oplus \ldots \oplus f(x_{i-1}, Y_{i-1})$ and $V = f(X_{i+1}, y_{i+1}) \oplus \ldots \oplus f(X_n, y_n)$. By the chain rule, and since U, V are bits, we get

$$\begin{aligned}
H(X_iY_i|x_{<i}y_{>i}UV) + 2 &\geq H(X_iY_i|x_{<i}y_{>i}UV) + H(UV|x_{<i}y_{>i})\\
&= H(X_iY_iUV|x_{<i}y_{>i})\\
&\geq H(X_iY_i|x_{<i}y_{>i})\\
&\geq k/n,
\end{aligned}$$

so there is some fixed value of u, v such that

$$H(X_iY_i|x_{<i}y_{>i}uv) \geq k/n - 2.$$

The desired rectangle is the support of (X_i, Y_i) conditioned on this fixed value of $(x_{<i}, y_{>i}, u, v)$, which we call T. Because (X, Y) is distributed uniformly in R, the distribution of (X_i, Y_i) conditioned on $(x_{<i}, y_{>i}, u, v)$ is a product distribution, and so T is a rectangle. By Proposition 5, $|T| \geq 2^{k/n-2}$. Because each input $(x_i, y_i) \in T$ corresponds to some input $(x, y) \in R$ with $f^{\oplus n}(x, y)$ fixed, and we have fixed $x_{<i}, y_{>i}$ and the xor of the function value in the first $i-1$ as well as the last $n-i$ coordinates, $f(x_i, y_i)$ is determined within T, and T is a monochromatic rectangle of f.

4. Proof of Theorem 3

The proof uses Theorem 4 and standard ideas along the lines of [NW95] to obtain a protocol for f. We shall prove that f has a protocol tree whose number of leaves is bounded by

$$2^{O((\log C(f^{\oplus n})^{1/n} + \log \mathsf{rk}(f)) \cdot \log \mathsf{rk}(f))} \tag{1}$$

By applying Lemma 8 to this protocol, we obtain a protocol with communication

$$O\Big((\log C(f^{\oplus n})^{1/n} + \log \mathsf{rk}(f)) \log \mathsf{rk}(f)\Big) \geq D(f),$$

which proves that

$$\log C(f^{\oplus n})^{1/n} \geq \frac{\Omega(D(f))}{\log \mathsf{rk}(f)} - \log \mathsf{rk}(f),$$

yielding the theorem.

We prove the bound by induction on $|\mathcal{X}| \cdot |\mathcal{Y}|$ and $\mathsf{rk}(f)$. If $\mathsf{rk}(f) < 5$, or $|\mathcal{X}| \cdot |\mathcal{Y}| \leq 1$, we obtain a protocol with a constant number of leaves. Otherwise, by averaging, $f^{\oplus n}$ has a monochromatic rectangle of size

$$\frac{|\mathcal{X}|^n \cdot |\mathcal{Y}|^n}{C(f^{\oplus n})}.$$

Theorem 4 then implies that f contains a monochromatic rectangle R of size at least

$$\frac{|\mathcal{X}| \cdot |\mathcal{Y}|}{4 \cdot C(f^{\oplus n})^{1/n}}.$$

We can use R to partition the matrix corresponding to f as follows

$$\begin{bmatrix} R & A \\ B & Z \end{bmatrix}.$$

Since R has rank 1, we have

$$\begin{aligned}
\mathsf{rk}(f) &\geq \mathsf{rk}\left(\begin{bmatrix} 0 & A \\ B & Z \end{bmatrix}\right) - 1 && \text{(Proposition 7)} \\
&\geq \mathsf{rk}\left(\begin{bmatrix} 0 & A \end{bmatrix}\right) + \mathsf{rk}\left(\begin{bmatrix} 0 \\ B \end{bmatrix}\right) - 1 && \text{(by Gaussian elimination)} \\
&\geq \mathsf{rk}\left(\begin{bmatrix} R & A \end{bmatrix}\right) + \mathsf{rk}\left(\begin{bmatrix} R \\ B \end{bmatrix}\right) - 3. && \text{(Proposition 7)}
\end{aligned}$$

So, we must have either

$$\mathsf{rk}(\begin{bmatrix} R & A \end{bmatrix}) \leq (\mathsf{rk}(f) + 3)/2, \tag{2}$$

or

$$\mathsf{rk}\Big(\begin{bmatrix} R \\ B \end{bmatrix}\Big) \leq (\mathsf{rk}(f) + 3)/2.$$

If Equation (2) holds, Alice sends a bit to Bob indicating whether her input is consistent with R. Otherwise, Bob sends a bit indicating whether his input is consistent with R. Without loss of generality, assume that Equation (2) holds.

Let f_0 and f_1 denote the sub-functions of f obtained by restricting to $\begin{bmatrix} R & A \end{bmatrix}$ and $\begin{bmatrix} B & Z \end{bmatrix}$ respectively. Since every rectangle cover of $f^{\oplus n}$ yields a rectangle cover of $f_0^{\oplus n}$ and a rectangle cover of $f_1^{\oplus n}$, we have

$$\max\{C(f_0^{\oplus n}), C(f_1^{\oplus n})\} \leq C(f^{\oplus n}).$$

If Alice's input is consistent with R, we may repeat the argument with the function f_0 which satisfies $\mathsf{rk}(f_0) \leq (\mathsf{rk}(f) + 3)/2 \leq 4\mathsf{rk}(f)/5$, so long as $\mathsf{rk}(f) \geq 5$. Otherwise, if Alice's input is inconsistent with R, we repeat the argument with the function f_1 which has at most

$$|\mathcal{X}| \cdot |\mathcal{Y}| \cdot \Big(1 - \frac{1}{4 \cdot C(f^{\oplus n})^{1/n}}\Big)$$

inputs.

The number of recursive steps where the rank reduces by a factor of 4/5 is at most $O(\log \mathsf{rk}(f))$. Moreover, since the matrix corresponding to f has at most $2^{\mathsf{rk}(f)}$ distinct rows and columns, the number of steps where the input space shrinks by a factor of $(1 - \frac{1}{4 \cdot C(f^{\oplus n})^{1/n}})$ is at most $8 \cdot \mathsf{rk}(f) \cdot C(f^{\oplus n})^{1/n}$. That is because after so many steps the number of inputs is at most

$$2^{2\cdot \mathsf{rk}(f)} \cdot \left(1 - \frac{1}{4 \cdot C(f^{\oplus n})^{1/n}}\right)^{8\mathsf{rk}(f)\cdot C(f^{\oplus n})^{1/n}} \leq 2^{2\mathsf{rk}(f)} \cdot e^{-2\mathsf{rk}(f)} < 1.$$

The number of leaves in the protocol we have designed is at most

$$\binom{8 \cdot \mathsf{rk}(f) \cdot C(f^{\oplus n})^{1/n} + O(\log \mathsf{rk}(f))}{O(\log \mathsf{rk}(f))} \leq 2^{O((\log C(f^{\oplus n})^{1/n} + \log \mathsf{rk}(f)) \log \mathsf{rk}(f))},$$

since $C(f^{\oplus n}) \geq 1$. This proves Equation (1).

References

[BBCR10] Boaz Barak, Mark Braverman, Xi Chen, and Anup Rao. How to compress interactive communication. In *Proceedings of the Forty-Second ACM Symposium on Theory of Computing*, STOC '10, page 67–76, New York, NY, USA, 2010. Association for Computing Machinery.

[BJKS02] Ziv Bar-Yossef, T. S. Jayram, Ravi Kumar, and D. Sivakumar. An information statistics approach to data stream and communication complexity. In *43rd Symposium on Foundations of Computer Science (FOCS 2002), 16-19 November 2002, Vancouver, BC, Canada, Proceedings*, pages 209–218. IEEE Computer Society, 2002.

[BR11] Mark Braverman and Anup Rao. Information equals amortized communication. In *2013 IEEE 54th Annual Symposium on Foundations of Computer Science*, pages 748–757, Los Alamitos, CA, USA, oct 2011. IEEE Computer Society.

[Bra15] Mark Braverman. Interactive information complexity. *SIAM Journal on Computing*, 44(6):1698–1739, 2015.

[BRWY13a] Mark Braverman, Anup Rao, Omri Weinstein, and Amir Yehudayoff. Direct product via round-preserving compression. In Fedor V. Fomin, Rusins Freivalds, Marta Z. Kwiatkowska, and David Peleg, editors, *Automata, Languages, and Programming - 40th International Colloquium, ICALP 2013, Riga, Latvia, July 8-12, 2013, Proceedings, Part I*, volume 7965 of *Lecture Notes in Computer Science*, pages 232–243. Springer, 2013.

[BRWY13b] Mark Braverman, Anup Rao, Omri Weinstein, and Amir Yehudayoff. Direct products in communication complexity. In *54th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2013, 26-29 October, 2013, Berkeley, CA, USA*, pages 746–755. IEEE Computer Society, 2013.

[CSWY01] Amit Chakrabarti, Yaoyun Shi, Anthony Wirth, and Andrew Chi-Chih Yao. Informational complexity and the direct sum problem for simultaneous message complexity. In *42nd Annual Symposium on Foundations of Computer Science, FOCS 2001, 14-17 October 2001, Las Vegas, Nevada, USA*, pages 270–278. IEEE Computer Society, 2001.

[FKNN95] Tomás Feder, Eyal Kushilevitz, Moni Naor, and Noam Nisan. Amortized communication complexity. *SIAM Journal on Computing*, 24(4):736–750, 1995.

[GKR16] Anat Ganor, Gillat Kol, and Ran Raz. Exponential separation of information and communication for boolean functions. *J. ACM*, 63(5), nov 2016.

[GMWW17] Dmitry Gavinsky, Or Meir, Omri Weinstein, and Avi Wigderson. Toward better formula lower bounds: The composition of a function and a universal relation. *SIAM J. Comput.*, 46(1):114–131, 2017.

[GPW18] Mika Göös, Toniann Pitassi, and Thomas Watson. Deterministic communication vs. partition number. *SIAM Journal on Computing*, 47(6):2435–2450, 2018.

[HJMR10] Prahladh Harsha, Rahul Jain, David McAllester, and Jaikumar Radhakrishnan. The communication complexity of correlation. *IEEE Transactions on Information Theory*, 56(1):438–449, 2010.

[IR24] Siddharth Iyer and Anup Rao. Xor lemmas for communication via marginal information. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*, STOC 2024, page 652–658, New York, NY, USA, 2024. Association for Computing Machinery.

[JPY12] Rahul Jain, Attila Pereszlényi, and Penghui Yao. A direct product theorem for the two-party bounded-round public-coin communication complexity. In *53rd Annual IEEE Symposium on Foundations of Computer Science, FOCS 2012, New Brunswick, NJ, USA, October 20-23, 2012*, pages 167–176. IEEE Computer Society, 2012.

[JRS03] Rahul Jain, Jaikumar Radhakrishnan, and Pranab Sen. A direct sum theorem in communication complexity via message compression. In Jos C. M. Baeten, Jan Karel Lenstra, Joachim Parrow, and Gerhard J. Woeginger, editors, *Automata, Languages and Programming*, pages 300–315, Berlin, Heidelberg, 2003. Springer Berlin Heidelberg.

[KN97] Eyal Kushilevitz and Noam Nisan. *Communication complexity*. Cambridge University Press, Cambridge, 1997.

[Kol16] Gillat Kol. Interactive compression for product distributions. In *Proceedings of the Forty-Eighth Annual ACM Symposium on Theory of Computing*, STOC '16, page 987–998, New York, NY, USA, 2016. Association for Computing Machinery.

[KRW95] Mauricio Karchmer, Ran Raz, and Avi Wigderson. Super-logarithmic depth lower bounds via the direct sum in communication complexity. *Comput. Complex.*, 5(3/4):191–204, 1995.

[KW90] Mauricio Karchmer and Avi Wigderson. Monotone circuits for connectivity require super-logarithmic depth. *SIAM Journal on Discrete Mathematics*, 3(2):255–265, 1990.

[Lov14] Shachar Lovett. Communication is bounded by root of rank. In *Proceedings of the Forty-Sixth Annual ACM Symposium on Theory of Computing*, STOC '14, page 842–846, New York, NY, USA, 2014. Association for Computing Machinery.

[LS88] László Miklós Lovász and Michael E. Saks. Lattices, mobius functions and communications complexity. *[Proceedings 1988] 29th Annual Symposium on Foundations of Computer Science*, pages 81–90, 1988.

[MW19] Or Meir and Avi Wigderson. Prediction from partial information and hindsight, with application to circuit lower bounds. *Comput. Complex.*, 28(2):145–183, 2019.

[NW95] Noam Nisan and Avi Wigderson. On rank vs. communication complexity. *Comb.*, 15(4):557–565, 1995.

[Raz92] Alexander A. Razborov. On the distributional complexity of disjointness. *Theor. Comput. Sci.*, 106(2):385–390, 1992.

[Raz95] Ran Raz. A parallel repetition theorem. In *Proceedings of the Twenty-Seventh Annual ACM Symposium on Theory of Computing*, STOC '95, page 447–456, New York, NY, USA, 1995. Association for Computing Machinery.

[RR15] Sivaramakrishnan Natarajan Ramamoorthy and Anup Rao. How to compress asymmetric communication. In *Proceedings of the 30th Conference on Computational Complexity*, CCC '15, page 102–123, Dagstuhl, DEU, 2015. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

[RY20] Anup Rao and Amir Yehudayoff. *Communication Complexity: and Applications.* Cambridge University Press, 2020.

[She18] Alexander A. Sherstov. Compressing interactive communication under product distributions. *SIAM Journal on Computing*, 47(2):367–419, 2018.

[SK87] Georg Schnitger and Bala Kalyanasundaram. The probabilistic communication complexity of set intersection. In *Proceedings of the Second Annual Conference on Structure in Complexity Theory, Cornell University, Ithaca, New York, USA, June 16-19, 1987*, pages 41–47. IEEE Computer Society, 1987.

[ST23] Benny Sudakov and István Tomon. Matrix discrepancy and the log-rank conjecture. https://arxiv.org/abs/2311.18524, 2023.

[Yu22] Huacheng Yu. Strong XOR lemma for communication with bounded rounds : (extended abstract). In *63rd IEEE Annual Symposium on Foundations of Computer Science, FOCS 2022, Denver, CO, USA, October 31 - November 3, 2022*, pages 1186–1192. IEEE, 2022.

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il