The Computational Complexity Column

by

V. Arvind

Institute of Mathematical Sciences, CIT Campus, Taramani
Chennai 600113, India
arvind@imsc.res.in
http://www.imsc.res.in/~arvind

Communication complexity is a fundamental sub-area of complexity theory that studies the amount of communication required between several parties to jointly perform a computational task. In the basic model, Alice and Bob need to jointly compute a boolean function $f(x, y)$. Alice holds the input part x and Bob holds y. The communication cost is the number of bits exchanged between them during computation. The *log-rank conjecture*, formulated twenty-five years ago, states that there is a deterministic protocol for computing f whose communication cost is polynomial in the logarithm of the rank of the associated communication matrix M_f.

There has been exciting progress on this long-standing open problem in the last year. In this timely survey article, Shachar Lovett explains the recent results with intuitive proof sketches, and points out various future directions for progress.

Recent advances on the log-rank conjecture in communication complexity

Shachar Lovett*

Abstract

The log-rank conjecture is one of the fundamental open problems in communication complexity. It speculates that the deterministic communication complexity of any two-party function is equal to the log of the rank of its associated matrix, up to polynomial factors. Despite much research, we still know very little about this conjecture. Recently, there has been renewed interest in this conjecture and its relations to other fundamental problems in complexity theory. This survey describes some of the recent progress, and hints at potential directions for future research.

1 Introduction

Communication complexity studies the amount of communication needed in order to evaluate a function, whose output depends on information distributed amongst two or more parties. Since its first introduction by Yao [35], communication complexity was extensively studied, to a large extent because of its applications in diverse fields, such as circuit complexity, VLSI design, proof complexity, streaming algorithms, data structures and more. Still, there are many fundamental problems about the communication complexity of functions which are wide open. We refer the reader to the book of Kushilevitz and Nisan [8] for more details on communication complexity and its applications, and to the book of Lee and Shraibman [16] for an exposition of more recent lower bound techniques in communication complexity.

In this survey, we focus on the communication complexity between two parties. Let $f : X \times Y \to \{0, 1\}$ be a boolean function, where one party holds an inputs $x \in X$, the other party holds an input $y \in Y$, and their

*CSE department, UC San-Diego. e-mail: slovett@cse.ucsd.edu.

goal is to evaluate $f(x, y)$ while minimizing their communication. For most of this survey, we will focus on deterministic protocols, which is the simplest communication model. The *deterministic communication complexity* of f is the minimal number of bits communicated by an optimal deterministic protocol computing f, and is denoted by $\mathrm{CC}^{\mathrm{det}}(f)$.

There is a simple lower bound on the deterministic communication complexity of functions, first observed by Mehlhorn and Schmidt [21], based on the rank of their associated matrix. Let M_f be the $X \times Y$ matrix with $M_{x,y} = f(x, y)$. A deterministic protocol computing f in which the players send c bits of communication, corresponds to a partition of the matrix M_f to 2^c rectangles (a rectangle is a set $A \times B$ with $A \subset X, B \subset Y$) such that the value of M_f is constant on each rectangle. Such rectangles are called *monochromatic*. As the rank (as a real matrix) of a monochromatic rectangle is at most one, we get that $\mathrm{rank}(M_f) \leq 2^c$. Equivalently, if we shorthand $\mathrm{rank}(f) = \mathrm{rank}(M_f)$ then

$$\mathrm{CC}^{\mathrm{det}}(f) \geq \log \mathrm{rank}(f).$$

The *log-rank conjecture* proposed by Lovász and Saks [15] speculates that this simple bound is tight for all boolean functions, up to polynomial factors.

Conjecture 1.1 (The log-rank conjecture [15])**.** *There exists a universal constant $C > 0$ such that for any boolean function f,*

$$\mathrm{CC}^{\mathrm{det}}(f) \leq C(\log \mathrm{rank}(f))^C.$$

Validity of the log-rank conjecture is one of the fundamental open problems in communication complexity. It is true in all known examples, but still very little progress has been made towards resolving it. In the special case where M_f is the adjacency matrix of a graph G, an essentially equivalent conjecture given by van Nuffelen [34] and Fajtlowicz [4] replaces the communication complexity by the (weaker notion) of log of the chromatic number of the graph; equivalently, that $\chi(G) \leq \exp(\log^{O(1)} \mathrm{rank}(G))$.

A simple upper bound is that $\mathrm{CC}^{\mathrm{det}}(f) \leq \mathrm{rank}(f)$, which is exponentially worse than what is conjectured by the log-rank conjecture. It follows from the simple observation that if $\mathrm{rank}(f) = r$, then there could be at most 2^r distinct rows in M_f. Hence, one can assume without loss of generality that $|X| \leq 2^r$, and consider a protocol in which the first player simply sends its input x. In the special case of graphs, Kotlov and Lovász [7] proved that if a graph has rank r, then its chromatic number is at most $2^{r/2}$. This was later improved to $(4/3)^r$ by Kotlov [9].

In terms of lower bounds, a sequence of works [1, 23, 25, 28] culminating in an example due to Kushilevitz (unpublished, cf. [23]) shows that there

exist functions for which $\mathrm{CC}^{\det}(f) \geq (\log \mathrm{rank}(f))^{\log_3 6}$. Hence,the constant C in Conjecture 1.1, if it exists, must satisfy $C \geq \log_3 6 \approx 1.63$.

Recently, there was renewed interest in the log-rank conjecture and its relations to several other problems in complexity theory. Ben-Sasson, Ron-Zewi and the author [2] studied the relation of the log-rank conjecture to the approximate duality conjecture of [3], and showed that if one assumes a number-theoretic conjecture (the polynomial Freiman-Ruzsa conjecture) then the trivial upper bound can be reduced by a logarithmic factor.

Theorem 1.2 ([2]). *Assuming the polynomial Freiman-Ruzsa conjecture over $\mathbb{F}_2^n$, for any boolean function f,*

$$\mathrm{CC}^{\det}(f) \leq O(\mathrm{rank}(f)/\log \mathrm{rank}(f)).$$

Gavinsky and the author [5] studied the relation between deterministic and randomized protocols for low rank matrices, and showed that in order to prove the log-rank conjecture, it suffices to prove that any low rank matrix has an efficient randomized protocol. In fact, they show that even weaker notions of protocols are sufficient, like low information cost protocols or efficient zero-communication protocols. We will show here the following result.

Theorem 1.3 ([5]). *If a boolean function f has a randomized protocol of complexity c, then it also has a deterministic protocol of complexity $O(c \cdot \log^2(\mathrm{rank}(f)))$.*

Finally, the author [14] proved a new (unconditional) upper bound, based on discrepancy of low rank matrices, which improves the previous upper bound by nearly a quadratic factor.

Theorem 1.4 ([14]). *For any boolean function f,*

$$\mathrm{CC}^{\det}(f) \leq O\left(\sqrt{\mathrm{rank}(f)} \cdot \log \mathrm{rank}(f)\right).$$

The goal of this survey is to explain these recent works, discuss their relations to other fundamental problems in complexity theory, and speculate on what directions seem the most likely to yield further advances for the log-rank conjecture. This is by no means a comprehensive survey. In particular, a related line of research which will not be discussed here is the study of the log-rank conjecture restricted to special families of functions. For example, the case of XOR functions (functions of the form $f(x, y) = F(x \oplus y)$) and related problems has received considerable attention recently [10,18–20,30,31,36–38].

Paper organization. In Section 2 we present a result of Nisan and Wigderson which allows to reduce the problem of constructing deterministic protocols to the simpler problem of exhibiting a large monochromatic rectangle. As this result is used repeatedly, we include its proof for completeness. In Section 3 we discuss the approximate duality conjecture in additive combinatorics, its relations to the log-rank conjecture and to constructions of two-source extractors. In Section 4 we show that low-rank functions with efficient randomized protocols also have efficient deterministic protocols. In Section 5 we apply bounds on the discrepancy of low-rank functions to deduce better upper bounds on deterministic protocols. In Section 6 we discuss several directions for further research, including relations to the problem of matrix rigidity.

2 From monochromatic rectangles to protocols

The log-rank conjecture speculates that if M_f has a low rank, then it can be partitioned into a small number of monochromatic rectangles. In particular, it must have a large monochromatic rectangle. A beautiful reduction of Nisan and Wigderson [23] shows that if one can prove that any low rank boolean matrix has a large monochromatic rectangle, then it can be bootstrapped to design a protocol with nearly the same efficiency. As this reduction would be useful for us, we review it below. We recall that a monochromatic rectangle is a subset $R = A \times B \subset X \times Y$ such that $f(x,y)$ is constant for all $(x,y) \in R$.

Theorem 2.1 ([23]). *Assume that for any function $f : X \times Y \to \{0,1\}$ with $\mathrm{rank}(f) = r$, there exists a monochromatic rectangle of size $|R| \geq 2^{-c(r)}|X \times Y|$. Then, any boolean function of rank r is computable by a deterministic protocol of complexity $O(\log^2 r + \sum_{i=0}^{\log r} c(r/2^i))$.*

Before giving the proof, we note that if $c(r) = \mathrm{poly}\log(r)$ then Theorem 2.1 implies a protocol with deterministic communication complexity $\mathrm{poly}\log(r)$, hence proving the log-rank conjecture. On the other end of the spectrum, if $c(r) = r^\alpha$ for some $\alpha < 1$ then Theorem 2.1 implies a protocol with deterministic communication complexity $O(r^\alpha)$.

Proof. Let f be a function with $\mathrm{rank}(M_f) = r$, and let R be the assumed monochromatic rectangle of size $2^{-c(r)} \cdot |X \times Y|$. Consider the partition of the matrix M_f as

$$M_f = \begin{pmatrix} R & S \\ P & Q \end{pmatrix}$$

As R is monochromatic, $\text{rank}(R) \leq 1$. Hence, $\text{rank}(S) + \text{rank}(P) \leq r + 1$. Assume, without loss of generality, that $\text{rank}(S) \leq r/2 + 1$ (otherwise, exchange the roles of the rows player and columns player). The row player sends one bit, indicating whether the input x is in the top part or in the bottom part of the matrix. If it is in the top part then the rank decreases to $\text{rank}(R\ S) \leq \text{rank}(R) + \text{rank}(S) \leq r/2 + 2$. If it is in the bottom part, the rank might not decrease, but the size of the matrix reduces to at most $(1 - 2^{-c(r)})|X \times Y|$. Iterating this process defines a protocol tree. We next bound the number of leaves of the protocol. By standard techniques, any protocol tree can be balanced so that the communication complexity is logarithmic in the number of leaves (cf. [8, Chapter 2, Lemma 2.8]).

Consider the protocol which stops once the rank drops to approximately $r/2$. The protocol tree in this case has at most $O(2^{c(r)} \cdot \log(|XY|))$ leaves, and hence can be simulated by a protocol sending only $O(c(r) + \log\log(|XY|))$ bits. Note that since we can assume f has no repeated rows or columns, $|XY| \leq 2^{2r}$ and hence $\log\log(|XY|) \leq \log(r) + 1$. Next, consider the phase where the protocol continues until the rank drops to $r/4$. Again, this protocol can be simulated by $O(c(r/2)+\log(r))$ bits of communication. Summing over $r/2^i$ for $i = 0, \ldots, \log(r)$ gives the bound. □

3 Approximate duality and the log-rank conjecture

Nisan and Wigderson [23] proved another interesting fact: any low rank boolean matrix contains a large rectangle which is slightly biased. The bias of f over a rectangle R is defined as

$$\text{bias}(f|R) = \left|\mathbb{E}_{(x,y)\in R}\left[(-1)^{f(x,y)}\right]\right| = \left|\Pr_{(x,y)\in R}[f(x,y) = 0] - \Pr_{(x,y)\in R}[f(x,y) = 1]\right|.$$

We also define $\text{bias}(f)$ to be the bias of f over the full space $X \times Y$. We will later see a generalization of this fact, called discrepancy, which is measured against the worst case distribution of inputs.

Theorem 3.1 ([23]). *Let $f : X \times Y \to \{0, 1\}$ with $\text{rank}(f) = r$. Then there exists a rectangle R of size $|R| \geq |X \times Y|/O(r^{3/2})$ such that $\text{bias}(f|R) \geq 1/O(r^{3/2})$.*

Let us restrict f to the rectangle R so that we may assume for simplicity $\text{bias}(f) \geq \varepsilon = 1/O(r^{3/2})$. Thus, we may ask whether it is easier to study the structure of low rank matrices, if we further assume that they are somewhat

biased. Recall that Theorem 2.1 requires us to find a large monochromatic rectangle. This raises the following problem.

Problem 3.2. *Let f be a boolean function such that* $\mathrm{rank}(f) = r$ *and* $\mathrm{bias}(f) \geq \varepsilon$. *What is the largest monochromatic rectangle that M_f must contain?*

The previous discussion shows that this problem is essentially equivalent to the log-rank conjecture, as long as the bias is inverse polynomially related to the rank. The main idea of Ben-Sasson et al. [2] is to consider a related problem, where instead of considering the matrices over the reals, we consider them over the binary finite field $\mathbb{F}_2$. In the following, we denote by $\mathrm{rank}_{\mathbb{F}_2}(M_f)$ the rank of a matrix over $\mathbb{F}_2$; note that the rank over $\mathbb{F}_2$ is always at most the rank over the reals, e.g. $\mathrm{rank}_{\mathbb{F}_2}(M_f) \leq \mathrm{rank}(M_f)$.

Approximate duality. We now introduce a seemingly unrelated problem. Let $A, B \subset \mathbb{F}_2^r$ be subsets. The *approximate duality* measure of A, B is

$$\varepsilon = \left|\mathbb{E}_{a\in A, b\in B}[(-1)^{\langle a,b\rangle}]\right| = \left|\Pr_{a\in A, b\in B}[\langle a,b\rangle = 0] - \Pr_{a\in A, b\in B}[\langle a,b\rangle = 1]\right|.$$

We say the sets are ε-approximate dual if their approximate duality measure is at least ε. Note that $\varepsilon = 1$ corresponds to sets which are orthogonal (possibly after applying an affine shift to one of the sets). The *approximate duality conjecture* of Ben-Sasson and Ron-Zewi [3] speculates that any large sets which are approximate dual, must contain large subsets which are dual.

Conjecture 3.3 (Approximate duality conjecture [3])**.** *Let $A, B \subset \mathbb{F}_2^r$ be sets which are ε-approximate dual. Then there exist subsets $A' \subset A, B' \subset B$ and a value $c \in \mathbb{F}_2$ such that*

$$\langle a,b\rangle = c \qquad \forall a \in A', b \in B',$$

where

$$\frac{|A|}{|A'|}, \frac{|B|}{|B'|} \leq 2^{O\left(\sqrt{r\log(1/\varepsilon)}\right)}.$$

The bound in Conjecture 3.3, if true, is the best possible, as the following example shows. Let $A = B$ be the set of all vectors in $\mathbb{F}_2^r$ of hamming weight $\sqrt{r}/10$. Then the probability that a uniformly chosen $a \in A, b \in B$ intersect is at most $1/100$, and hence A, B are ε-approximate dual for $\varepsilon \geq 0.98$. On the other hand, the largest subsets $A' \subset A, B' \subset B$ which are orthogonal come from choosing $A' = A \cap (\{0,1\}^{r/2} \times 0^{r/2})$ to be the set of vectors supported

on the first half of the coordinates, and $B' = B \cap (0^{r/2} \times \{0,1\}^{r/2})$ to be the vectors supported on the last half of the coordinates. One can then verify that $|A|/|A'| = |B|/|B'| = \exp(\Omega(\sqrt{r}))$. The bound for general $\varepsilon > 0$ can be similarly obtained, by considering $A = B$ to be the vectors in $\mathbb{F}_2^r$ of hamming weight $O(\sqrt{r \log(1/\varepsilon)})$.

Approximate duality and the log-rank conjecture. Let us now relate the approximate duality conjecture with the log-rank conjecture. By Theorem 3.1, if $\mathrm{rank}(M_f) = r$ (where the rank is over the reals) we may assume (by potentially restricting f to a large rectangle) that $\mathrm{bias}(f) \geq \varepsilon = 1/O(r^{3/2})$. Moreover, $\mathrm{rank}_{\mathbb{F}_2}(f) \leq \mathrm{rank}(f) = r$. Equivalently put, there are vectors $a_x, b_y \in \mathbb{F}_2^r$ such that

$$\langle a_x, b_y \rangle = f(x,y).$$

Let us define $A = \{a_x : x \in X\}, B = \{b_y : y \in Y\}$. Then by definition, since $\mathrm{bias}(f) \geq \varepsilon$, the sets A, B are ε-approximate dual. Then, by the approximate duality conjecture, there are large subsets $A' \subset A, B' \subset B$ such that $\langle a, b \rangle$ is constant for all $a \in A', b \in B'$. That is, the rectangle $A' \times B'$ is monochromatic! Working out the parameters, the approximate duality conjecture implies that M_f contains a monochromatic rectangle R of size $|R| \geq \exp(-O(\sqrt{r \log(r)}))|X \times Y|$. As this holds for any matrix of rank r, Theorem 2.1 implies that f has a deterministic protocol of complexity at most $O(\sqrt{r \log(r)})$. Thus, we obtain the following corollary.

Corollary 3.4. *If Conjecture 3.3 is true, then any boolean function f with $\mathrm{rank}(f) = r$ has a deterministic protocol of complexity $O(\sqrt{r \log(r)})$.*

Of course, we do not know if Conjecture 3.3 is true or not. Ben-Sasson and Ron-Zewi proved the following weak version of it, which has no direct implication for the log-rank conjecture.

Theorem 3.5 ([3]). *For any $\alpha > 0$ there exist $\varepsilon > 0$ such that the following holds. Let $A, B \subset \mathbb{F}_2^r$ be sets which are $(1-\varepsilon)$-approximate dual. Then there exist subsets $A' \subset A, B' \subset B$ and a value $c \in \mathbb{F}_2$ such that*

$$\langle a, b \rangle = c \qquad \forall a \in A', b \in B',$$

where

$$\frac{|A|}{|A'|}, \frac{|B|}{|B'|} \leq 2^{\alpha r}.$$

Ben-Sasson, Ron-Zewi and the author [2] proved a slightly stronger version, assuming a number-theoretic conjecture known as the polynomial Freiman-Ruzsa conjecture. This conjecture can be defined over arbitrary Abelian groups, but we only need it for the additive group $\mathbb{F}_2^n$.

Conjecture 3.6 (The polynomial Freiman-Ruzsa conjecture over $\mathbb{F}_2^n$)**.** *Let $A \subset \mathbb{F}_2^n$ be a set, and let $A + A = \{a_1 + a_2 : a_1, a_2 \in A\}$ be its sumset. If $|A+A| \leq K|A|$ then there exists an affine subspace $V \subset \mathbb{F}_2^n$ of size $|V| \leq |A|$ such that*

$$|A \cap V| \geq K^{-O(1)}|A|.$$

The polynomial Freiman-Ruzsa conjecture is one of the fundamental open problems in additive combinatorics, see e.g. [6] for a discussion of the conjecture. A quasi-polynomial analog of it was proved by Sanders [29], see also [13] for an exposition. If one assumes Conjecture 3.6 to hold, Ben-Sasson et al [2] proved an improved bound on the approximate duality conjecture.

Theorem 3.7 ([2])**.** *Assume that the polynomial Freiman-Ruzsa conjecture over $\mathbb{F}_2^n$ (Conjecture 3.6) is true. Let $A, B \subset \mathbb{F}_2^r$ be sets which are ε-approximate dual for $\varepsilon \geq 2^{-\sqrt{r}}$. Then there exist subsets $A' \subset A, B' \subset B$ and a value $c \in \mathbb{F}_2$ such that*

$$\langle a, b \rangle = c \qquad \forall a \in A', b \in B',$$

where

$$\frac{|A|}{|A'|}, \frac{|B|}{|B'|} \leq 2^{O(r/\log(r))}.$$

Theorem 1.2 follows as an immediate corollary from the combination of Theorem 3.7 with Theorem 2.1. We restate it below for the convenience of the reader.

Theorem 1.2 (restated) *Assuming the polynomial Freiman-Ruzsa conjecture over $\mathbb{F}_2^n$, for any boolean function f,*

$$\mathrm{CC}^{\det}(f) \leq O(\mathrm{rank}(f)/\log \mathrm{rank}(f)).$$

Approximate duality and two-source extractors. The original application of [3] for the approximate duality conjecture was for the construction of pseudo-random graphs, specifically construction of two-source extractors from certain constructions of two-source dispersers. In the following, we focus for simplicity on the case of dispersers and extractors which output a single bit, and we somewhat abuse the standard notations in this field. Let $G = (U, V, E)$ be a bi-partite graph. The graph G is a k-Ramsey graph (also called a disperser), if it contains no bi-partite clique or independent set of size $k \times k$. Equivalently, for any subsets $A \subset U, B \subset V$ of size $|A| = |B| = k$, if we denote by $E(A, B)$ the set of induced edges between A and B, then

$$1 \leq |E(A, B)| \leq |A||B| - 1.$$

The graph is called a (k, ε) two-source extractor if in fact the number of edges between A, B is close to what might be expected in a random graph, that is

$$(1/2 - \varepsilon)|A||B| \leq |E(A, B)| \leq (1/2 + \varepsilon)|A||B|.$$

Ben-Sasson and Ron-Zewi [3] showed that certain constructions of Ramsey graphs are inherently also two-source extractors for weaker parameters. Consider the following construction of a bi-partite graph $G = (U, V, E)$: $U, V \subset \mathbb{F}_2^n$, and for $u \in U, v \in V$ we have $(u, v) \in E$ if $\langle u, v \rangle = 1$. Assume that G is not a (k, ε) two-source extractor. That is, there are subsets $A \subset U, B \subset V$ of size $|A| = |B| = k$ such that (say) $|E(A, B)| \geq (1/2 + \varepsilon)|A||B|$. This means that the approximate duality measure between A, B is at least 2ε, which by the approximate duality conjecture (Conjecture 3.3) implies that we can find large subsets $A' \subset A, B' \subset B$ such that (say) $|E(A', B')| = 0$. Then, we conclude that the graph G is not a k'-Ramsey graph for $k' = \min(|A'|, |B'|)$. Otherwise put, any bi-partite graph, constructed in this way, which is k'-Ramsey, must also be a (k, ε) two-source extractor, where k is somewhat larger than k'. For further details we refer the reader to the original paper [3].

4 From randomized to deterministic protocols

The log-rank conjecture speculates that low rank boolean functions have efficient deterministic protocols. We already saw in Theorem 2.1 that a sufficient condition is that any low rank boolean matrix contains a large monochromatic rectangle. Here, we describe another reduction, due to Gavinsky and the author [5]. We will show that it is also sufficient to construct a randomized protocol computing the function.

A randomized protocol computing a function $f(x, y)$ is a protocol, in which both parties are allowed to use randomized strategies, such that for every input x, y, the protocol computes the correct value $f(x, y)$ with probability at least $2/3$. Note that a randomized protocol is a distribution over deterministic protocols. The complexity of a randomized protocol is the maximal number of bits that may be sent by the protocol. We recall Theorem 1.3 for the convenience of the reader.

Theorem 1.3 (restated) *If a boolean function has a randomized protocol of complexity c, then it also has a deterministic protocol of complexity $O(c \cdot \log^2(\mathrm{rank}(f)))$.*

Proof. Let $p(x, y)$ denote the probability that the protocol computes f correctly on inputs x, y, where by assumption $p(x, y) \geq 2/3$. We can increase the success probability by repeating the protocol a few times, and

computing the majority of the values obtained. Specifically, if we repeat the protocol $O(\log 1/\varepsilon)$ times, we obtain a randomized protocol which uses $c' = O(c \log(1/\varepsilon))$ bits and computes $f(x,y)$ correctly with probability $1-\varepsilon$. A randomized protocol is a distribution over deterministic protocols; hence, if we consider the uniform distribution over inputs, we get by an averaging argument that there exists a deterministic protocol $\pi(x,y)$ of complexity c' such that

$$\left|\{(x,y) \in X \times Y : \pi(x,y) = f(x,y)\}\right| \geq (1-\varepsilon)\,|X \times Y|.$$

A deterministic protocol of complexity c' corresponds to a partition to $N = 2^{c'}$ many rectangles. We next argue that there exists a large rectangle on which f is nearly fixed. Let $R_1, \ldots, R_N$ denote the rectangles corresponding to the protocol π. Denote by $\mu(R) = |R|/|X \times Y|$ the fractional size of a rectangle, and by $\alpha(R) = |\{(x,y) \in R : \pi(x,y) \neq f(x,y)\}|/|R|$ the fraction of elements in R on which the protocol π makes a mistake. By assumption, we have

$$\sum_{i=1}^{N} \mu(R_i) = 1; \qquad \sum_{i=1}^{N} \mu(R_i)\alpha(R_i) \leq \varepsilon.$$

One can verify that these imply that there must be a rectangle $R = R_i$ such that

$$\mu(R) \geq 1/2N; \qquad \alpha(R) \leq 2\varepsilon.$$

As π is fixed on R, we can assume without loss of generality that

$$|\{(x,y) \in R : f(x,y) = 1\}| \geq (1-2\varepsilon)|R|.$$

Let $r = \mathrm{rank}(f)$. We next show that by setting $\varepsilon = 1/8r$, there exists a large sub-rectangle $R' \subset R$ on which f is monochromatic.

Claim 4.1. *Let f be a boolean function of rank r, and assume there exists a rectangle R on which $f(x,y) = 1$ for at least $1 - 1/4r$ of the elements in R. Then, there exists a sub-rectangle $R' \subset R$ of size $|R'| \geq |R|/8$ such that $f(x,y) = 1$ for all $(x,y) \in R'$.*

Proof. Let $R = A \times B$. Let $A' \subset A$ be the set of rows for which at most $1/2r$ fraction of the elements are -1,

$$A' = \left\{x \in A : |\{y \in B : f(x,y) = -1\}| \leq |B|/2r\right\}.$$

By Markov inequality, $|A'| \geq |A|/2$. Let $x_1, \ldots, x_r \in A'$ be indices so that their rows span f restricted to $A' \times B$. Let

$$B' = \{y \in B : f(x_1,y) = \ldots = f(x_r,y) = 1\}.$$

Since each of the rows $x_1, \ldots, x_r$ contain at most $1/2r$ fraction of elements which are -1 we have $|B'| \geq |B|/2$. Now, this implies that all rows in $A' \times B'$ are either the all 1 or all -1. Choosing the largest half gives the required rectangle. This gives a monochromatic rectangle $R' \subset R$ of size $|R'| \geq |R|/8$. □

To conclude, we would like to apply Theorem 2.1 in order to show the existence of a deterministic protocol. The reader can verify, that although the conditions of Theorem 2.1 require one to show that any low rank function has a large monochromatic rectangle, in fact for the proof to go through, it suffices to assume that this holds only for functions which are restrictions of f to rectangles. The same argument as above shows that for any rectangle $R \subset X \times Y$, there exists a sub-rectangle $R' \subset R$ of size $|R'| \geq 2^{-O(c \log(r))}|R|$ on which f is monochromatic. Note that, as the bound c does not improve as the rank decreases, we incur an additional multiplicative factor of $\log(r)$ in the communication complexity. We deduce that there exists a deterministic protocol computing f of complexity $O(c \log^2(r))$, as claimed. □

5 Discrepancy of matrices and the log-rank conjecture

Let $f : X \times Y \to \{-1, 1\}$ be a boolean function. For a distribution μ on $X \times Y$, the *discrepancy* of f with respect to μ is the maximal correlation that f has with rectangles,

$$\mathrm{disc}(f; \mu) = \max_R \left| \sum_{(x,y) \in R} f(x, y)\mu(x, y) \right|$$

where R ranges over all rectangles. The discrepancy of f is its discrepancy for the worse case distribution,

$$\mathrm{disc}(f) = \min_\mu \mathrm{disc}(f; \mu).$$

Discrepancy is a well-studied property in the context of communication complexity lower bounds, see e.g. the survey [12] for details. On the other hand, it is known that low-rank boolean matrices have noticeable discrepancy [11,17]: if f has rank r then

$$\mathrm{disc}(f) \geq \frac{1}{8\sqrt{r}}. \tag{1}$$

A result of the author [14] shows that discrepancy can be used to prove upper bounds as well. We restate Theorem 1.4 for the convenience of the reader.

Theorem 1.4 (restated) *For any boolean function f,*

$$\mathrm{CC}^{\det}(f) \leq O\left(\sqrt{\mathrm{rank}(f)} \cdot \log \mathrm{rank}(f)\right).$$

The following lemma is the main technical tool. It shows that a function with high discrepancy contains a large rectangle which is almost monochromatic. In fact, this is true with respect to any distribution over the inputs. We make the following definitions: given a distribution μ over $X \times Y$, let $\mu(R) = \sum_{(x,y)\in R} \mu(x,y)$ denote the probability of an input landing in R, and $\mathbb{E}_\mu[f] = \sum_{(x,y)\in X\times Y} \mu(x,y)f(x,y)$ the average of f with respect to μ. For a rectangle R such that $\mu(R) > 0$, let $\mu|R$ the distribution μ conditioned on being in R, that is, $(\mu|R)(x,y) = 1_{(x,y)\in R} \cdot \mu(x,y)/\mu(R)$.

Lemma 5.1. *Let $f : X \times Y \to \{-1,1\}$ be a function with $\mathrm{disc}(f) = \delta$. Then for any $\varepsilon > 0$ and any distribution μ over $X \times Y$, there exists a rectangle R with*

$$\mu(R) \geq 2^{-O(\delta^{-1}\cdot \log(1/\varepsilon))}$$

such that $\left|\mathbb{E}_{\mu|R}[f]\right| \geq 1 - \varepsilon$.

Proof of Theorem 1.4, assuming Lemma 5.1. Let f be any boolean function of rank r. Apply Lemma 5.1 with μ the uniform distribution over $X \times Y$, $\delta \geq 1/8\sqrt{r}$ and $\varepsilon = 1/4r$, to deduce the existence of a rectangle $R \subset X \times Y$ of size $|R| \geq 2^{-O(\sqrt{r}\log(r))}|X \times Y|$ such that $f(x,y) = v$ for $1 - 1/4r$ fraction of elements in R. Apply Claim 4.1 to deduce that there exists a sub-rectangle $R' \subset R$ of size $|R'| \geq |R|/8$ on which f is monochromatic. By Theorem 2.1, this implies that any function of rank r has a deterministic protocol of complexity $O(\sqrt{r}\log(r))$. □

We now turn to prove Lemma 5.1. The proof of Lemma 5.1 which we give below is a simplification of the original proof of [14], which was presented to us by Salil Vadhan [32].

Proof of Lemma 5.1. Let us assume without loss of generality that $\mathbb{E}_\mu[f] \geq 0$, otherwise apply the lemma to $-f$. Let σ be any distribution over $X \times Y$ such that $\mathbb{E}_\sigma[f] = 0$. By assumption, there exists a rectangle R_1 such that

$$\left|\sum_{(x,y)\in R_1} \sigma(x,y)f(x,y)\right| \geq \delta.$$

Let $R_1 = A \times B$ and define $A' = X \setminus A, B' = Y \setminus B$. Consider the four rectangles

$$R_1 = A \times B, R_2 = A' \times B, R_3 = A \times B', R_4 = A' \times B'.$$

As $\sum_{(x,y)\in X\times Y} \sigma(x,y)f(x,y) = \mathbb{E}_\sigma[f] = 0$, there must exist a rectangle $R \in \{R_1, R_2, R_3, R_4\}$ such that

$$\sum_{(x,y)\in R} \sigma(x,y)f(x,y) \geq \delta/3.$$

This holds for any distribution σ for which $\mathbb{E}_\sigma[f] = 0$. Hence, we can apply von Neumann's Minimax Theorem [22] and deduce that there exists a distribution ρ over rectangles, such that for any distribution σ for which $\mathbb{E}_\sigma[f] = 0$, we have

$$\mathbb{E}_{R\sim\rho}\left[\sum_{(x,y)\in R} \sigma(x,y)f(x,y)\right] \geq \delta/3.$$

Equivalently,

$$\sum_{(x,y)\in X\times Y} \Pr_{R\sim\rho}[(x,y) \in R] \cdot \sigma(x,y)f(x,y) \geq \delta/3.$$

Fix $(x_1, y_1) \in f^{-1}(1)$ and $(x_2, y_2) \in f^{-1}(-1)$. Let σ be the distribution given by $\sigma(x_1, y_1) = \sigma(x_2, y_2) = 1/2$. As $\mathbb{E}_\sigma[f] = 0$ we have

$$\Pr_{R\sim\rho}[(x_1, y_1) \in R] - \Pr_{R\sim\rho}[(x_2, y_2) \in R] \geq (2/3)\delta.$$

Let p be the *minimal* probability that $(x_1, y_1) \in R$ over all $(x_1, y_1) \in f^{-1}(1)$, where R is sampled according to ρ; and let q be the *maximal* probability that $(x_2, y_2) \in R$ over all $(x_2, y_2) \in f^{-1}(-1)$. We established that

$$p - q \geq (2/3)\delta.$$

Fix $t \geq 1$ and let $R_1, \ldots, R_t \sim \rho$ be chosen independently, and let $R^* = R_1 \cap \ldots \cap R_t$ be their intersection. We will show that for an appropriate choice of t, the rectangle R^* satisfies the requirements of the lemma with positive probability (and hence such a rectangle exists). We will use the fact that for any $x \in X, y \in Y$,

$$\Pr[(x,y) \in R^*] = \Pr_{R\sim\rho}[(x,y) \in R]^t.$$

Consider the random variable

$$T = \mu(R^*) - (1/\varepsilon) \cdot \mu(R^* \cap f^{-1}(-1)).$$

By linearity of expectation, we have

$$\begin{aligned}\mathbb{E}[T] &= \sum_{(x,y)\in f^{-1}(1)} \mu(x,y) \Pr[(x,y) \in R^*] - \sum_{(x,y)\in f^{-1}(-1)} \mu(x,y)((1/\varepsilon)-1) \Pr[(x,y) \in R^*] \\ &\geq \mu(f^{-1}(1)) \cdot p^t - \mu(f^{-1}(-1)) \cdot q^t/\varepsilon \\ &\geq 1/2 \cdot (p^t - q^t/\varepsilon),\end{aligned}$$

where we used our initial assumption that $\mathbb{E}_\mu[f] = \mu(f^{-1}(1)) - \mu(f^{-1}(-1)) \geq 0$. Setting $t = O(p/\delta \cdot \log(1/\varepsilon))$ gives

$$q^t/p^t \leq (1 - (2/3)\delta/p)^t \leq \varepsilon/2.$$

For this choice of t, we have

$$\mathbb{E}[T] \geq p^t/4 = 2^{-O(\delta^{-1} \cdot \log(1/\varepsilon))}.$$

Let R^* be a rectangle which achieves this average, that is

$$\mu(R^*) - (1/\varepsilon) \cdot \mu(R^* \cap f^{-1}(-1)) \geq 2^{-O(\delta^{-1} \cdot \log(1/\varepsilon))}.$$

In particular, we learn that both $\mu(R^*) \geq 2^{-O(\delta^{-1} \cdot \log(1/\varepsilon))}$ (which satisfies the first requirement) and furthermore that $\mu(R^* \cap f^{-1}(-1)) \leq \varepsilon \cdot \mu(R^*)$, which implies that $\mathbb{E}_{\mu|R^*}[f] \geq 1 - \varepsilon$ (which satisfies the second requirement). □

6 Further research

There are several directions for further research. We describe a few concrete ones below.

6.1 Randomized protocols vs approximate rank

The *approximate rank* of a boolean function $f(x,y)$ is the minimal rank of an $X \times Y$ real matrix M such that

$$2/3 \leq M(x,y)f(x,y) \leq 1.$$

Similar to the log rank lower bound for the deterministic communication complexity, the log of the approximate rank is a lower bound on the randomized communication complexity of a function. The log-rank conjecture for randomized protocols speculates that it is also an upper bound, up to polynomial factors. As a first step, one can attempt to generalize Theorem 1.4 to approximate rank and randomized protocols.

Problem 6.1. *Let f be a boolean function with approximate rank r. Show that f has a randomized protocol of complexity $\sqrt{r} \cdot \operatorname{poly}\log(r)$.*

6.2 Quantum protocols for low-rank matrices

The work of [5] shows that if low-rank functions have certain types of efficient protocols (randomized protocols, low information cost protocols, or zero-communication protocols), then up to a poly-logarithmic factor in the rank, they also have efficient deterministic protocols. One type of protocol which they were not able to analyze is quantum protocols. This is interesting on its own right, but also because to the best of our current knowledge, it may be that quantum protocols are only polynomially better than randomized protocols, for any complete boolean function (exponential separations are known for partial functions, see e.g. [26,27]). Thus, understanding quantum protocols, even just for low-rank functions, seems like an important step towards a better understanding of quantum protocols in general.

Problem 6.2. *Let f be a boolean function which can be computed by a quantum protocol of complexity c. Show that f can also be computed by a deterministic protocol of complexity $c \cdot \operatorname{poly}\log(\operatorname{rank}(f))$.*

6.3 The structure of low-rank sparse matrices, and matrix rigidity

The proof of Theorem 1.4 applies to boolean matrices. We conjecture in [14] that it can be generalized to show that any low rank sparse matrix contains a large zero rectangle.

Conjecture 6.3. *Let M be an $n \times n$ real matrix with $\operatorname{rank}(M) = r$ and such that $M_{i,j} \neq 0$ for at most εn^2 entries. Then, there exist $A, B \subset [n]$ such that*

$$M_{a,b} = 0 \qquad \forall a \in A, b \in B$$

such that $|A|, |B| \geq n \cdot \exp(-O(\sqrt{\varepsilon r}))$.

The reader can observe the similarities of Conjecture 6.3 to the approximate duality conjecture (Conjecture 3.3) which we discussed. Note that here we consider the case where nearly all the elements are zero, while in the approximate duality conjecture we only assumed a small bias. Nevertheless, the same construction shows that the bounds in Conjecture 6.3, if true, are the best possible.

A matrix M is called (r, s)-rigid, if its rank cannot be made smaller than r by changing at most s entries in M. The problem of explicitly constructing

rigid matrices was introduced by Valiant [33] in the context of arithmetic circuits lower bounds, and was also studied by Razborov [24] in the context of separation of the analogs of PH and PSPACE in communication complexity. Despite much research, the best results to date are achieved by the so-called "untouched minor" argument, which gives explicit matrices which are (r, s)-rigid with $s = \Omega\left(\frac{n^2}{r}\log\left(\frac{n}{r}\right)\right)$. See e.g. the survey of Lokam [12] for details. We will prove the following corollary of Conjecture 6.3, which improves previous bounds by a logarithmic factor.

Corollary 6.4. *Assuming Conjecture 6.3, there exists an explicit $n \times n$ real matrix which is (r, s)-rigid for $s = \Omega\left(\frac{n^2}{r}\log^2\left(\frac{n}{r}\right)\right)$.*

Proof. Let M be an $n \times n$ matrix of rank r, such that all $r \times r$ minors of M have full rank. For example, such a matrix may be constructed as $M = NN^t$ where N is an $n \times r$ matrix such that any r rows of N are linearly independent. Assume that M is not (r, s)-rigid. Then, we can decompose

$$M = L + S, \quad \text{rank}(L) < r, \quad S \text{ is s-sparse.}$$

Let $s = \varepsilon n^2$. The matrix S is both s-sparse and low rank, as $\text{rank}(S) \leq \text{rank}(M) + \text{rank}(L) < 2r$. Hence, by Conjecture 6.3, there exist $A, B \subset [n]$ of size $|A|, |B| \geq n \cdot \exp(-O(\sqrt{\varepsilon r}))$ such that $S_{a,b} = 0$ for all $a \in A, b \in B$. Hence, $M_{a,b} = L_{a,b}$. If $|A|, |B| \geq r$, we must have that $\text{rank}(L) \geq \text{rank}(M) = r$. So, $n \cdot \exp(-O(\sqrt{\varepsilon r})) < r$ and the corollary follows by rearranging the terms. □

Acknowledgements I thank V. Arvind for inviting me to write this survey. I thank Dmitry Gavinsky, Noga Ron-Zewi and Adi Shraibman for helpful comments on earlier versions of this manuscript.

References

[1] Noga Alon and Paul D Seymour. A counterexample to the rank-coloring conjecture. *Journal of Graph Theory*, 13(4):523–525, 1989.

[2] Eli Ben-Sasson, Shachar Lovett, and Noga Ron-Zewi. An Additive Combinatorics Approach Relating Rank to Communication Complexity. *Proceedings of the 53rd Annual Symposium on Foundations of Computer Science*, pages 177–186, 2012.

[3] Eli Ben-Sasson and Noga Zewi. From affine to two-source extractors via approximate duality. In *Proceedings of the 43rd annual ACM symposium on Theory of computing*, pages 177–186. ACM, 2011.

[4] Siemion Fajtlowicz. On conjectures of graffiti. *Discrete mathematics*, 72(1):113–118, 1988.

[5] D. Gavinsky and S. Lovett. En Route to the log-rank Conjecture: New Reductions and Equivalent Formulations. *Electronic Colloquium on Computational Complexity (ECCC)*, 20(80), 2013.

[6] Ben Green. Finite field models in additive combinatorics. *arXiv preprint math/0409420*, 2004.

[7] Andrew Kotlov and László Lovász. The rank and size of graphs. *Journal of Graph Theory*, 23(2):185–189, 1996.

[8] E. Kushilevitz and N. Nisan. Communication Complexity. *Cambridge University Press*, 1997.

[9] A. Kotlov. Rank and Chromatic Number of a Graph. *Journal of Graph Theory 26(1)*, pages 1–8, 1997.

[10] Ming Lam Leung, Yang Li, and Shengyu Zhang. Tight bounds on communication complexity of symmetric xor functions in one-way and smp models. In *Theory and Applications of Models of Computation*, pages 403–408. Springer, 2011.

[11] N. Linial, S. Mendelson, G. Schechtman, and A. Shraibman. Complexity Measures of Sign Matrices. *Combinatorica 27(4)*, pages 439–463, 2007.

[12] Satyanarayana V. Lokam. Complexity lower bounds using linear algebra. *Found. Trends Theor. Comput. Sci.*, 4(1–2):1–155, January 2009.

[13] Shachar Lovett. An exposition of sanders quasi-polynomial freiman-ruzsa theorem. In *Electronic Colloquium on Computational Complexity (ECCC)*, volume 19, page 29, 2012.

[14] Shachar Lovett. Communication is bounded by root of rank. *arXiv preprint arXiv:1306.1877*, 2013.

[15] L. Lovász and M. Saks. Lattices, Möbius Functions and Communication Complexity. *Annual Symposium on Foundations of Computer Science*, pages 81–90, 1988.

[16] Troy Lee and Adi Shraibman. *Lower bounds in communication complexity*. Now Publishers Inc, 2009.

[17] N. Linial and A. Shraibman. Learning Complexity vs. Communication Complexity. *Combinatorics, Probability & Computing 18(1-2)*, pages 227–245, 2009.

[18] Troy Lee and Shengyu Zhang. Composition theorems in communication complexity. In *Automata, Languages and Programming*, pages 475–489. Springer, 2010.

[19] Yang Liu and Shengyu Zhang. Quantum and randomized communication complexity of xor functions in the smp model. In *Electronic Colloquium on Computational Complexity (ECCC)*, volume 20, page 10, 2013.

[20] Ashley Montanaro and Tobias Osborne. On the communication complexity of xor functions. *arXiv preprint arXiv:0909.3392*, 2009.

[21] Kurt Mehlhorn and Erik M Schmidt. Las vegas is better than determinism in vlsi and distributed computing. In *Proceedings of the fourteenth annual ACM symposium on Theory of computing*, pages 330–337. ACM, 1982.

[22] J v Neumann. Zur theorie der gesellschaftsspiele. *Mathematische Annalen*, 100(1):295–320, 1928.

[23] N. Nisan and A. Wigderson. On Rank vs. Communication Complexity. *Proceedings of the 35rd Annual Symposium on Foundations of Computer Science*, pages 831–836, 1994.

[24] Alexander Razborov. On rigid matrices (in russian). *Technical report, Steklov Mathematical Institute*, 1989.

[25] Alexander A Razborov. The gap between the chromatic number of a graph and the rank of its adjacency matrix is superlinear. *Discrete mathematics*, 108(1):393–396, 1992.

[26] Ran Raz. Exponential separation of quantum and classical communication complexity. In *Proceedings of the thirty-first annual ACM symposium on Theory of computing*, pages 358–367. ACM, 1999.

[27] Oded Regev and Bo'az Klartag. Quantum one-way communication can be exponentially stronger than classical communication. In *Proceedings of the 43rd annual ACM symposium on Theory of computing*, pages 31–40. ACM, 2011.

[28] Ran Raz and Boris Spieker. On the âĂIJlog rankâĂİ-conjecture in communication complexity. *Combinatorica*, 15(4):567–588, 1995.

[29] Tom Sanders. On the bogolyubov-ruzsa lemma. *arXiv preprint arXiv:1011.0107*, 2010.

[30] Amir Shpilka, Avishay Tal, and Ben lee Volk. On the structure of boolean functions with small spectral norm. *CoRR*, abs/1304.0371, 2013.

[31] Xiaoming Sun, Chengu Wang, et al. Randomized communication complexity for linear algebra problems over finite fields. In *Symposium on Theoretical Aspects of Computer Science*, volume 14, pages 477–488, 2012.

[32] Salil Vadhan, 2013. personal communication.

[33] LeslieG. Valiant. Graph-theoretic arguments in low-level complexity. In Jozef Gruska, editor, *Mathematical Foundations of Computer Science 1977*, volume 53 of *Lecture Notes in Computer Science*, pages 162–176. Springer Berlin Heidelberg, 1977.

[34] Cyriel van Nuffelen. A bound for the chromatic number of a graph. *American Mathematical Monthly*, pages 265–266, 1976.

[35] Andrew Chi-Chih Yao. Some complexity questions related to distributive computing (preliminary report). In *Proceedings of the eleventh annual ACM symposium on Theory of computing*, pages 209–213. ACM, 1979.

[36] Shengyu Zhang. Efficient quantum protocols for xor functions. *arXiv preprint arXiv:1307.6738*, 2013.

[37] Zhiqiang Zhang and Yaoyun Shi. Communication complexities of symmetric xor functions. *Quantum information and computation*, 9(3&4):0255–0263, 2009.

[38] Zhiqiang Zhang and Yaoyun Shi. On the parity complexity measures of boolean functions. *Theoretical Computer Science*, 411(26):2612–2618, 2010.