Lower Bounds in Communication Complexity

Lower Bounds in Communication Complexity

Troy Lee

Columbia University
New York, NY 10027
USA
troyjlee@gmail.com

Adi Shraibman

Weizmann Institute
Rehovot 76100
Israel
adi.shribman@gmail.com

now
the essence of knowledge

Boston – Delft

Foundations and Trends® in
Theoretical Computer Science

Published, sold and distributed by:
now Publishers Inc.
PO Box 1024
Hanover, MA 02339
USA
Tel. +1-781-985-4510
www.nowpublishers.com
sales@nowpublishers.com

Outside North America:
now Publishers Inc.
PO Box 179
2600 AD Delft
The Netherlands
Tel. +31-6-51115274

The preferred citation for this publication is T. Lee and A. Shraibman, Lower Bounds in Communication Complexity, Foundation and Trends® in Theoretical Computer Science, vol 3, no 4, pp 263–399, 2007

ISBN: 978-1-60198-258-2

Foundations and Trends® in
Theoretical Computer Science
Volume 3 Issue 4, 2007

Editorial Scope

Foundations and Trends® in Theoretical Computer Science will publish survey and tutorial articles in the following topics:

- Algorithmic game theory
- Computational algebra
- Computational aspects of combinatorics and graph theory
- Computational aspects of communication
- Computational biology
- Computational complexity
- Computational geometry
- Computational learning
- Computational Models and Complexity
- Computational Number Theory
- Cryptography and information security
- Data structures
- Database theory
- Design and analysis of algorithms
- Distributed computing
- Information retrieval
- Operations Research
- Parallel algorithms
- Quantum Computation
- Randomness in Computation

Information for Librarians

Foundations and Trends® in Theoretical Computer Science, 2007, Volume 3, 4 issues. ISSN paper version 1551-305X. ISSN online version 1551-3068. Also available as a combined paper and online subscription.

Foundations and Trends$^{\circledR}$ in
Theoretical Computer Science
Vol. 3, No. 4 (2007) 263–399
© 2009 T. Lee and A. Shraibman
DOI: 10.1561/0400000040

now
the essence of knowledge

Lower Bounds in Communication Complexity

Troy Lee[1] **and Adi Shraibman**[2]

[1] *Columbia University, New York, NY 10027, USA, troyjlee@gmail.com*
[2] *Department of Mathematics, Weizmann Institute, Rehovot, 76100, Israel, adi.shribman@gmail.com*

Abstract

The communication complexity of a function $f(x,y)$ measures the number of bits that two players, one who knows x and the other who knows y, must exchange to determine the value $f(x,y)$. Communication complexity is a fundamental measure of complexity of functions. Lower bounds on this measure lead to lower bounds on many other measures of computational complexity. This monograph surveys lower bounds in the field of communication complexity. Our focus is on lower bounds that work by first representing the communication complexity measure in Euclidean space. That is to say, the first step in these lower bound techniques is to find a geometric complexity measure, such as rank or trace norm, that serves as a lower bound to the underlying communication complexity measure. Lower bounds on this geometric complexity measure are then found using algebraic and geometric tools.

Contents

1

Introduction

Communication complexity studies how much communication is needed in order to evaluate a function whose output depends on information distributed amongst two or more parties. Yao [101] introduced an elegant mathematical framework for the study of communication complexity, applicable in numerous situations, from an e-mail conversation between two people, to processors communicating on a chip. Indeed, the applicability of communication complexity to other areas, including circuit and formula complexity, VLSI design, proof complexity, and streaming algorithms, is one reason why it has attracted so much study. See the excellent book of Kushilevitz and Nisan [56] for more details on these applications and communication complexity in general.

Another reason why communication complexity is a popular model for study is simply that it is an interesting mathematical model. Moreover, it has that rare combination in complexity theory of a model for which we can actually hope to show tight lower bounds, yet these bounds often require the development of nontrivial techniques and sometimes are only obtained after several years of sustained effort.

In the basic setting of communication complexity, two players Alice and Bob wish to compute a function $f\colon X \times Y \to \{T,F\}$ where X,Y are arbitrary finite sets. Alice holds an input $x \in X$, Bob $y \in Y$, and they wish to evaluate $f(x,y)$ while minimizing the number of bits communicated. We let Alice and Bob have arbitrary computational power as we are really interested in how much information must be exchanged in order to compute the function, not issues of running time or space complexity.

Formally, a communication protocol is a binary tree where each internal node v is labeled either by a function $a_v\colon X \to \{0,1\}$ or a function $b_v\colon Y \to \{0,1\}$. Intuitively each node corresponds to a turn of either Alice or Bob to speak. The function a_v indicates, for every possible input x, how Alice will speak if the communication arrives at that node, and similarly for b_v. The leaves are labeled by an element from $\{T,F\}$. On input x,y the computation traces a path through the tree as indicated by the functions a_v, b_v. The computation proceeds to the left child of a node v if $a_v(x) = 0$ and the right child if $a_v(x) = 1$, and similarly when the node is labeled by b_v. The protocol correctly computes f if for every input x,y, the computation arrives at a leaf ℓ labeled by $f(x,y)$.

The cost of a protocol is the height of the protocol tree. The deterministic communication complexity of a function f, denoted $D(f)$, is the minimum cost of a protocol correctly computing f. Notice that, as we have defined things, the transcript of the communication defines the output, thus both parties "know" the answer at the end of the protocol. One could alternatively define a correct protocol where only one party needs to know the answer at the end, but this would only make a difference of one bit in the communication complexity.

If we let $n = \min\{\lceil \log|X| \rceil, \lceil \log|Y| \rceil\}$ then clearly $D(f) \leq n+1$ as either Alice or Bob can simply send their entire input to the other, who can then compute the function and send the answer back. We refer to this as the trivial protocol. Thus the communication complexity of f will be a natural number between 1 and $n+1$, and our goal is to determine this number. This can be done by showing a lower bound on how much communication is needed, and giving a protocol of matching complexity.

The main focus of this survey is on showing lower bounds on the communication complexity of explicit functions. We treat different variants of communication complexity, including randomized, quantum, and multiparty models. Many tools have been developed for this purpose from a diverse set of fields including linear algebra, Fourier analysis, and information theory. As is often the case in complexity theory, demonstrating a lower bound is usually the more difficult task.

One of the most important lower bound techniques in communication complexity is based on matrix rank. In fact, it is not too much of an exaggeration to say that a large part of communication complexity is the study of different variants of matrix rank. To explain the rank bound, we must first introduce the *communication matrix*, a very useful and common way of representing a function $f\colon X \times Y \to \{T,F\}$. We will consider both a *Boolean* and a *sign* version of the communication matrix, the difference being in the particular integer representation of $\{T,F\}$. A Boolean matrix has all entries from $\{0,1\}$, whereas a sign matrix has entries from $\{-1,+1\}$. The Boolean communication matrix for f, denoted B_f, is a $|X|$-by-$|Y|$ matrix where $B_f[x,y] = 1$ if $f(x,y) = T$ and $B_f[x,y] = 0$ if $f(x,y) = F$. The sign communication matrix for f, denoted A_f, is a $\{-1,+1\}$-valued matrix where $A_f[x,y] = -1$ if $f(x,y) = T$ and $A_f[x,y] = +1$ if $f(x,y) = F$. Depending on the particular situation, it can be more convenient to reason about one representation or the other, and we will use both versions throughout this survey. Fortunately, this choice is usually simply a matter of convenience and not of great consequence — it can be seen that they are related as $B_f = (J - A_f)/2$, where J is the all-ones matrix. Thus the matrix rank of the two versions, for example, will differ by at most one.

Throughout this survey we identify a function $f\colon X \times Y \to \{T,F\}$ with its corresponding (sign or Boolean) communication matrix. The representation of a function as a matrix immediately puts tools from linear algebra at our disposal. Indeed, Mehlhorn and Schmidt [69] showed how matrix rank can be used to lower bound deterministic communication complexity. This lower bound follows quite simply from the properties of a deterministic protocol, but we delay a proof until Section 2.

Theorem 1.1 (Mehlhorn and Schmidt [69]). For every sign matrix A,

$$\log \operatorname{rank}(A) \leq D(A).$$

The rank bound has nearly everything one could hope for in a lower bound technique. From a complexity point of view it can be efficiently computed, i.e., computed in time polynomial in the size of the matrix. Furthermore, it frees us from thinking about communication protocols and lets us just consider the properties of A as a linear operator between Euclidean spaces, with all the attendant tools of linear algebra to help in doing this. Finally, it is even conjectured that one can always show polynomially tight bounds via the rank method. This log-rank conjecture is one of the greatest open problems in communication complexity.

Conjecture 1 (Lovász and Saks [67]). There is a constant c such that for every sign matrix A

$$D(A) \leq (\log \operatorname{rank}(A))^c + 2.$$

The additive term is needed because a rank-one sign matrix can require two bits of communication. Thus far the largest known separation between log rank and deterministic communication, due to Nisan and Wigderson [73], shows that in Conjecture 1 the constant c must be at least 1.63.

The problems begin, however, when we start to study other models of communication complexity such as randomized, quantum, or multiparty variants. Here one can still give a lower bound in terms of an appropriate variation of rank, but the bounds now can become very difficult to evaluate. In the case of multiparty complexity, for example, the communication matrix becomes a communication tensor, and one must study tensor rank. Unlike matrix rank, the problem of computing tensor rank is NP-hard [41], and even basic questions like the largest possible rank of an n-by-n-by-n real tensor remain open.

For randomized or quantum variants of communication complexity, as shown by Krause [52] and Buhrman and de Wolf [24], respectively, the relevant rank bound turns out to be *approximate rank.*

Definition 1.1. Let A be a sign matrix. The approximate rank of A with approximation factor α, denoted $\operatorname{rank}^\alpha(A)$, is

$$\operatorname{rank}^\alpha(A) = \min_{B : 1 \leq A[i,j]B[i,j] \leq \alpha} \operatorname{rank}(B).$$

As we shall see in Sections 4 and 5, the logarithm of approximate rank is a lower bound on randomized and quantum communication complexity, where the approximation factor α relates to the success probability of the protocol. In analogy with the log-rank conjecture, it is also reasonable to conjecture here that this bound is polynomially tight.

Approximate rank, however, can be quite difficult to compute. While we do not know if it is NP-hard, similar rank minimization problems subject to linear constraints are NP-hard, see for example Section 7.3 of [98]. Part of this difficulty stems from the fact that approximate rank is an optimization problem over a nonconvex function.

This brings us to the main theme of our survey. We focus on lower bound techniques which are real-valued functions and ideally possess some "nice" properties, such as being convex. The development and application of these techniques follow a three-step approach which we now describe. This approach can be applied in much the same way for different models, be they randomized, quantum, or multiparty.

Say that we are interested in a complexity measure CC, a mapping from functions to the natural numbers, which could represent any one of the above models.

(1) Embed the problem in $\mathbb{R}^{m\times n}$. That is, find a function $\mathcal{G}\colon\mathbb{R}^{m\times n}\to\mathbb{R}$ such that

$$\mathcal{G}(A) \leq \mathrm{CC}(A),$$

for every sign matrix A. As is the case with rank and approximate rank, often $\mathcal{G}$ will itself be naturally phrased as a minimization problem.

(2) Find an equivalent formulation of $\mathcal{G}$ in terms of a maximization problem. This will of course not always be possible, as in the case of approximate rank. This can be done, however, for rank and for a broad class of optimization problems over convex functions.

(3) Prove lower bounds on $\mathcal{G}$ by exhibiting an element of the feasible set for which the objective function is large. We call such an element a *witness* as it witnesses that $\mathcal{G}$ is at least as large as a certain value.

We will delay most of the technical details of this approach to the main body of the survey, in particular to Section 6 where we discuss the use of duality to perform the key Step 2 to go from a "min" formulation to a "max" formulation. Here we limit ourselves to more general comments, providing some intuition as to why and in what circumstances this approach is useful.

Step 1 We are all familiar with the idea that it can be easier to find the extrema of a smooth real-valued function than a discrete-valued function. For example, for smooth functions the powerful tools of calculus are available. To illustrate, think of integer programming versus linear programming. The latter problem can be solved in polynomial time, while even simple instances of integer programming are known to be NP-hard.

The intuition behind the first step is the same. The complexity of a protocol is a discrete-valued function, so in determining communication complexity we are faced with an optimization problem over a discrete-valued function. By working instead with a real-valued lower bound $\mathcal{G}$ we will have more tools at our disposal to evaluate $\mathcal{G}$. Moreover, if $\mathcal{G}$ is "nice" — for example being an optimization problem over a convex function — then the set of tools available to us is particularly rich. For instance, we can use duality to enact Step 2.

We do potentially pay a price in performing Step 1 and working with a "nicer" function $\mathcal{G}$. It could be the case that $\mathcal{G}(A)$ is much smaller than $\mathrm{CC}(A)$ for some sign matrices A. Just as in approximation algorithms, we seek a bound that is not only easier to compute but also approximates $\mathrm{CC}(A)$ well. We will say that a representation $\mathcal{G}(A)$ is *faithful* if there is some constant k such that $\mathrm{CC}(A) \leq \mathcal{G}(A)^k$ for all sign matrices A.

Step 2 A communication complexity measure $\mathrm{CC}(A)$ is naturally phrased as a minimization problem — looking for a protocol of minimum cost. Often times, as with the case of approximate rank, our lower bound $\mathcal{G}$ is also naturally phrased as a minimization problem.

The difficulty, of course, is that to lower bound a minimization problem one has to deal with the universal quantifier $\forall$ — we have to show that every possible protocol requires a certain amount of communication.

When our complexity measure $\mathcal{G}$ is of a nice form, however, such as a minimization problem of a convex function, we can hope to find an *equivalent* formulation of $\mathcal{G}$ in terms of a maximization problem. A maximization problem is much easier to lower bound since we simply have to demonstrate a particular feasible instance for which the target function is large. In some sense this can be thought of as an "algorithmic approach" to lower bounds. In Section 6 we will show how this can be done for a large class of complexity measures known as *approximate norms*.

This is an instance of a more general phenomena: showing a statement about *existence* is often easier than proving a statement about *nonexistence*. The former can be certified by a witness, which we do not always expect for the latter. Take the example of graph planarity, i.e., the question of whether a graph can be drawn in the plane in such a way that its edges intersect only at their endpoints. While it can be tricky to find such a drawing, at least we know what form the answer will take. To show that a graph is nonplanar, however, seems like a much more daunting task unless one has heard of Kuratowski's Theorem or Wagner's Theorem. These theorems reduce the problem of nonexistence to that of existence: for example, Wagner's Theorem states that a graph is nonplanar if and only if it contains K_5, the complete graph on five vertices, or $K_{3,3}$ the complete three-by-three bipartite graph, as a minor. Not surprisingly, theorems of this flavor are key in efficient algorithmic solutions to planarity and nonplanarity testing.

Step 3 Now that we have our complexity measure $\mathcal{G}$ phrased in terms of a maximization problem, we are in much better shape. Any element from the feasible set can be used to show a lower bound, albeit not necessarily a good one. As a simple example, going back to the rank lower bound, we observe that a natural way to prove a lower bound on rank is to find a large set of columns (or rows) that are independent.

Finding a good witness to prove a lower bound for a certain complexity measure $\mathcal{G}$ can still be a very difficult task. This is the subject we take up in Section 7. There are still only a few situations where we know how to choose a good witness, but this topic has recently seen a lot of exciting progress and more is certainly still waiting to be discovered.

Approximate norms The main example of the three-step approach we study in this survey is for *approximate norms.* We now give a more technical description of this case; the reader can skip this section at first reading, or simply take it as an "impression" of what is to come.

Let Φ be any norm on $\mathbb{R}^{m\times n}$, and let $\alpha \geq 1$ be a real number. The *α-approximate norm* of an $m \times n$ sign matrix A is

$$\Phi^\alpha(A) = \min_{B:1\leq A[i,j]B[i,j]\leq\alpha} \Phi(B).$$

The limit as $\alpha \to \infty$ motivates the definition

$$\Phi^\infty(A) = \min_{B:1\leq A[i,j]B[i,j]} \Phi(B).$$

In Step 1 of the framework described above we will usually take $\mathcal{G}(A) = \Phi^\alpha(A)$ for an appropriate norm Φ. We will see that the familiar matrix trace norm is very useful for showing communication complexity lower bounds, and develop some more exotic norms as well. We discuss this step in each of the model specific chapters, showing which norms can be used to give lower bounds on deterministic (Section 2), nondeterministic (Section 3), randomized (Section 4), quantum (Section 5), and multiparty (Section 8) models.

The nice thing about taking $\mathcal{G}$ to be an approximate norm is that we can implement Step 2 of this framework in a general way. As described in Section 6, duality can be applied to yield an *equivalent* formulation for any approximate norm Φ^α in terms of a maximization. Namely, for a sign matrix A

$$\Phi^\alpha(A) = \max_W \frac{(1+\alpha)\langle A, W\rangle + (1-\alpha)\|W\|_1}{2\Phi^*(W)} \tag{1.1}$$

Here Φ^* is the dual norm:

$$\Phi^*(W) = \max_X \frac{\langle W, X\rangle}{\Phi(X)}$$

We have progressed to Step 3. We need to find a witness matrix W that makes the bound from Equation (1.1) large. As any matrix W at all gives a lower bound, we can start with an educated guess and modify it according to the difficulties that arise. This is similar to the case discussed earlier of trying to prove that a graph is planar — one can simply start drawing and see how it goes. The first choice of a witness that comes to mind is the target matrix A itself. This gives the lower bound

$$\Phi^{\alpha}(A) \geq \frac{(1+\alpha)\langle A, A\rangle + (1-\alpha)\|A\|_1}{2\Phi^*(A)} = \frac{mn}{\Phi^*(A)}. \tag{1.2}$$

This is actually not such a bad guess; for many interesting norms this lower bound is tight with high probability for a random matrix. But it is not always a good witness, and there can be a very large gap between the two sides of the inequality (Equation (1.2)). One reason that the matrix A might be a bad witness, for example, is that it contains a large submatrix S for which $\Phi^*(S)$ is relatively large.

A way to fix this deficiency is to take instead of A any matrix $P \circ A$, where P is a real matrix with nonnegative entries that sum up to 1. Here $\circ$ denotes the entrywise product. This yields a better lower bound

$$\Phi^{\alpha}(A) \geq \max_{\substack{P: P \geq 0 \\ \|P\|_1 = 1}} \frac{1}{\Phi^*(P \circ A)}. \tag{1.3}$$

Now, by a clever choice of P, we can for example give more weight to a good submatrix of A and less or zero weight to submatrices that attain large values on the dual norm. Although this new lower bound is indeed better, it is still possible to exhibit an exponential gap between the two sides of Equation (1.3). This is nicely explained by the following characterization given in Section 7.

Theorem 1.2. For every sign matrix A

$$\Phi^{\infty}(A) = \max_{\substack{P: P \geq 0 \\ \|P\|_1 = 1}} \frac{1}{\Phi^*(P \circ A)}.$$

The best value a witness matrix W which has the same sign as A in each entry can provide, therefore, is equal to $\Phi^{\infty}(A)$. It can be expected that there are matrices A for which $\Phi^\infty(A)$ is significantly smaller than $\Phi^\alpha(A)$ for say $\alpha = 2$.[1] This is indeed the case for some interesting communication complexity problems such as the SET INTERSECTION problem where $f(x,y) = \bigvee_i (x_i \wedge y_i)$, which will be a running example throughout the survey.

When $\Phi^\infty(A)$ is not a good lower bound on $\Phi^\alpha(A)$ for bounded α, there are only a few situations where we know how to choose a good witness. One case is where A is the sign matrix of a so-called *block composed function*, that is, a function of the form $(f \bullet g^n)(x,y) = f(g(x^1,y^1),\ldots,g(x^n,y^n))$ where $x = (x^1,\ldots,x^n)$ and $y = (y^1,\ldots,y^n)$. This case has recently seen exciting progress [92, 90, 94]. These works showed a lower bound on the complexity of a block composed function in terms of the approximate degree of f, subject to the inner function g satisfying some technical conditions. The strength of this approach is that the approximate degree of $f\colon \{0,1\}^n \to \{-1,+1\}$ is often easier to understand than its communication complexity. In particular, in the case where f is symmetric, i.e., only depends on the Hamming weight of the input, the approximate polynomial degree has been completely characterized [75]. These results are described in detail in Section 7.2.

Historical context The "three-step approach" to proving communication complexity lower bounds has already been used in the first papers studying communication complexity. In 1983, Yao [102] gave an equivalent "max" formulation of randomized communication complexity using von Neumann's minimax theorem. He showed that the 1/3-error randomized communication complexity is equal to the maximum over all probability distributions P, of the minimum cost of a deterministic protocol which errs with probability at most 1/3 with respect to P. Thus one can show lower bounds on randomized communication complexity by exhibiting a probability distribution which is hard for deterministic protocols. This principle is the starting point for many lower bound results on randomized complexity.

A second notable result using the "three-step approach" is a characterization by Karchmer, Kushilevitz, and Nisan [48] of nondeterministic

[1] Notice that $\Phi^\alpha(A)$ is a decreasing function of α.

communication complexity. Using results from approximation theory, they show that a certain linear program characterizes nondeterministic communication complexity, up to small factors. By then looking at the dual of this program, they obtain a "max" quantity which can always show near optimal lower bounds on nondeterministic communication complexity.

The study of quantum communication complexity has greatly contributed to our understanding of the role of convexity in communication complexity lower bounds, and these more recent developments occupy a large portion of this survey. The above two examples are remarkable in that they implement the "three-step approach" with an exact (near) representation of the communication model. For quantum communication complexity, however, we do not yet have such a characterization which is convenient for showing lower bounds. The search for good representations to approximate quantum communication complexity led in particular to the development of approximate norms [49, 83, 64]. Klauck (Lemma 3.1) introduced what we refer to in this survey as the μ^α-approximate norm, also known as the *generalized discrepancy method.* While implicit in Klauck and Razborov, the use of Steps 2 and 3 of the three-step approach becomes explicit in later works [64, 90, 94].

What is not covered In the 30 years since its inception, communication complexity has become a vital area of theoretical computer science, and there are many topics which we will not have the opportunity to address in this survey. We mention some of these here.

Much work has been done on protocols of a restricted form, for example one-way communication complexity where information only flows from Alice to Bob, or simultaneous message passing where Alice and Bob send a message to a referee who then outputs the function value. A nice introduction to some of these results can be found in [56]. In this survey we focus only on general protocols.

For the most part, we stick to lower bound methods that fit into the general framework described earlier. As we shall see, these methods do encompass many techniques proposed in the literature, but not all. In particular, a very nice approach which we do not discuss are lower bounds based on information theory. These methods, for example, can give an elegant proof of the optimal $\Omega(n)$ lower bound on the SET INTERSECTION problem. We refer the reader to [14] for more details.

We also restrict ourselves to the case where Alice and Bob want to compute a Boolean function. The study of the communication complexity of relations is very interesting and has nice connections to circuit depth and formula size lower bounds. More details on this topic can be found in Kushilevitz and Nisan [56].

Finally, there are some models of communication complexity which we do not discuss. Perhaps the most notable of these is the model of unbounded-error communication complexity. This is a randomized model where Alice and Bob only have to succeed on every input with probability strictly greater than 1/2. We refer the reader to [37, 91, 84] for interesting recent developments on this model.

References

[1] S. Aaronson and A. Ambainis, "Quantum search of spatial regions," *Theory of Computing*, vol. 1, pp. 47–79, 2005.

[2] S. Aaronson and A. Wigderson, "Algebrization: A new barrier in complexity theory," *ACM Transactions on Computing Theory*, vol. 1, 2009.

[3] A. Aho, J. Ullman, and M. Yannakakis, "On notions of information transfer in VLSI circuits," in *Proceedings of the 15th ACM Symposium on the Theory of Computing*, pp. 133–139, ACM, 1983.

[4] E. Allender, "A note on the power of threshold circuits," in *Proceedings of the 30th IEEE Symposium on Foundations of Computer Science*, pp. 580–584, 1989.

[5] N. Alon, "Problems and results in extremal combinatorics, Part i," *Discrete Mathematics*, vol. 273, pp. 31–53, 2003.

[6] N. Alon, "Perturbed identity matrices have high rank: Proof and applications," *Combinatorics, Probability, and Computing*, vol. 18, pp. 3–15, 2009.

[7] N. Alon, Y. Kohayakawa, C. Mauduit, C. Moreira, and V. Rödl, "Measures of pseudorandomness for finite sequences: Minimal values," *Combinatorics, Probability, and Computing*, vol. 15, pp. 1–29, 2006.

[8] N. Alon, Y. Matias, and M. Szegedy, "The space complexity of approximating the frequency moments," *Journal of Computer and System Sciences*, vol. 58, no. 1, pp. 137–147, 1999.

[9] N. Alon and A. Naor, "Approximating the cut-norm via Grothendieck's inequality," *SIAM Journal on Computing*, vol. 35, pp. 787–803, 2006.

[10] N. Alon and P. Seymour, "A counterexample to the rank-coloring conjecture," *Journal of Graph Theory*, vol. 13, no. 4, pp. 523–525, 1989.

[11] L. Babai, P. Frankl, and J. Simon, "Complexity classes in communication complexity theory," in *Proceedings of the 27th IEEE Symposium on Foundations of Computer Science*, IEEE, 1986.

[12] L. Babai, A. Gál, P. Kimmel, and S. Lokam, "Simultaneous messages vs. communication," *SIAM Journal on Computing*, vol. 33, no. 1, pp. 137–166, 2003.

[13] L. Babai, N. Nisan, and M. Szegedy, "Multiparty protocols and Logspace-hard pseudorandom sequences," in *Proceedings of the 21st ACM Symposium on the Theory of Computing*, pp. 1–11, ACM, 1989.

[14] Z. Bar-Yossef, T. Jayram, R. Kumar, and D. Sivakumar, "Information statistics approach to data stream and communication complexity," *Journal of Computer and System Sciences*, vol. 68, no. 4, pp. 702–732, 2004.

[15] R. Beals, H. Buhrman, R. Cleve, M. Mosca, and R. Wolf, "Quantum Lower Bounds by Polynomials," *Journal of the ACM*, vol. 48, no. 4, pp. 778–797, 2001.

[16] P. Beame, M. David, T. Pitassi, and P. Woelfel, "Separating deterministic from randomized NOF multiparty communication complexity," in *Proceedings of the 34th International Colloquium On Automata, Languages and Programming*, Lecture Notes in Computer Science. Springer-Verlag, 2007.

[17] P. Beame and D. Huynh-Ngoc, "Multiparty communication complexity of AC^0," Technical Report TR-08-082, ECCC, 2008.

[18] P. Beame, T. Pitassi, and N. Segerlind, "Lower bounds for Lovász-Schrijver systems and beyond follow from multiparty communication complexity," *SIAM Journal on Computing*, vol. 37, no. 3, pp. 845–869, 2006.

[19] P. Beame, T. Pitassi, N. Segerlind, and A. Wigderson, "A strong direct product lemma for corruption and the NOF complexity of disjointness," *Computational Complexity*, vol. 15, no. 4, pp. 391–432, 2006.

[20] R. Beigel and J. Tarui, "On ACC," *Computational Complexity*, vol. 4, pp. 350–366, 1994.

[21] S. Ben-David, N. Eiron, and H. Simon, "Limitations of learning via embeddings in Euclidean half spaces," *Journal of Machine Learning Research*, vol. 3, pp. 441–461, 2002.

[22] H. Buhrman, Personal communication, December 2007.

[23] H. Buhrman, R. Cleve, and A. Wigderson, "Quantum vs. classical communication and computation," in *Proceedings of the 30th ACM Symposium on the Theory of Computing*, pp. 63–68, ACM, 1998.

[24] H. Buhrman and R. de Wolf, "Communication complexity lower bounds by polynomials," in *Proceedings of the 16th IEEE Conference on Computational Complexity*, pp. 120–130, 2001.

[25] H. Buhrman and R. de Wolf, "Complexity Measures and Decision Tree Complexity: A Survey," *Theoretical Computer Science*, vol. 288, pp. 21–43, 2002.

[26] H. Buhrman, N. Vereshchagin, and R. de Wolf, "On computation and communication with small bias," in *Proceedings of the 22nd IEEE Conference on Computational Complexity*, pp. 24–32, IEEE, 2007.

[27] A. Chakrabarti, S. Khot, and X. Sun, "Near-optimal lower bounds on the multi-party communication complexity of set-disjointness," in *Proceedings of the 18th IEEE Conference on Computational Complexity*, IEEE, 2003.

[28] A. Chattopadhyay, "Discrepancy and the power of bottom fan-in depth-three circuits," in *Proceedings of the 48th IEEE Symposium on Foundations of Computer Science*, pp. 449–458, IEEE, 2007.

[29] A. Chattopadhyay, PhD thesis, McGill University, 2008.

[30] A. Chattopadhyay and A. Ada, "Multiparty communication complexity of disjointness," Technical Report TR-08-002, ECCC, 2008.

[31] F. Chung, "Quasi-random classes of hypergraphs," *Random Structures and Algorithms*, vol. 1, pp. 363–382, 1990.

[32] J. Clauser, M. Horne, A. Shimony, and R. Holt, "Proposed experiment to test local hidden-variable theories," *Physical Review Letters*, vol. 23, no. 15, pp. 880–884, 1969.

[33] M. David, T. Pitassi, and E. Viola, "Improved Separations between Nondeterministic and Randomized Multiparty Communication," in *APPROX-RANDOM*, volume 5171 of *Lecture Notes in Computer Science*, pp. 371–384, Springer, 2008.

[34] J. Degorre, M. Kaplan, S. Laplante, and J. Roland, "The communication complexity of non-signaling distributions," Technical Report 0804.4859, arXiv, 2008.

[35] S. Fajtlowicz, "On conjectures of graffiti," *Discrete Mathematics*, vol. 72, pp. 113–118, 1988.

[36] J. Ford and A. Gál, "Hadamard tensors and lower bounds on multiparty communication complexity," in *Proceedings of the 32th International Colloquium On Automata, Languages and Programming*, pp. 1163–1175, 2005.

[37] J. Forster, "A linear lower bound on the unbounded error probabilistic communication complexity," *Journal of Computer and System Sciences*, vol. 65, pp. 612–625, 2002.

[38] A. Frieze and R. Kannan, "Quick approximation to matrices and applications," *Combinatorica*, vol. 19, pp. 175–220, 1999.

[39] M. Goemans and D. Williamson, "Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming," *Journal of the ACM*, vol. 42, pp. 1115–1145, 1995.

[40] V. Grolmusz, "The BNS lower bound for multi-party protocols is nearly optimal," *Information and Computation*, vol. 112, no. 1, pp. 51–54, 1994.

[41] J. Håstad, "Tensorrank is NP-complete," *Journal of Algorithms*, vol. 11, pp. 644–654, 1990.

[42] J. Håstad and M. Goldmann, "On the power of small-depth threshold circuits," *Computational Complexity*, vol. 1, pp. 113–129, 1991.

[43] G. J. O. Jameson, *Summing and Nuclear Norms in Banach Space Theory*. Cambridge University Press, 1987.

[44] T. Jayram, R. Kumar, and D. Sivakumar, "Two applications of information complexity," in *Proceedings of the 35th ACM Symposium on the Theory of Computing*, pp. 673–682, ACM, 2003.

[45] T. Jiang and B. Ravikumar, "Minimal NFA problems are hard," *SIAM Journal on Computing*, vol. 22, pp. 1117–1141, 1993.

[46] W. Johnson and J. Lindenstrauss, "Basic concepts in the geometry of Banach spaces," in *Handbook of the Geometry of Banach Spaces, Vol. I*, pp. 1–84, Amsterdam: North-Holland, 2001.

[47] B. Kalyanasundaram and G. Schnitger, "The probabilistic communication complexity of set intersection," in *Proceedings of the 2nd Annual Conference on Structure in Complexity Theory*, pp. 41–49, 1987.

[48] M. Karchmer, E. Kushilevitz, and N. Nisan, "Fractional Covers and Communication Complexity," *SIAM Journal on Discrete Mathematics*, vol. 8, no. 1, pp. 76–92, 1995.

[49] H. Klauck, "Lower bounds for quantum communication complexity," in *Proceedings of the 42nd IEEE Symposium on Foundations of Computer Science*, IEEE, 2001.

[50] H. Klauck, "Rectangle size bounds and threshold covers in communication complexity," in *Proceedings of the 18th IEEE Conference on Computational Complexity*, IEEE, 2003.

[51] A. Klivans and A. Sherstov, "A lower bound for agnostically learning disjunctions," in *Proceedings of the 20th Conference on Learning Theory*, 2007.

[52] M. Krause, "Geometric arguments yield better bounds for threshold circuits and distributed computing," *Theoretical Computer Science*, vol. 156, pp. 99–117, 1996.

[53] I. Kremer, "Quantum communication," Technical Report, Hebrew University of Jerusalem, 1995.

[54] J. Krivine, "Constantes de Grothendieck et fonctions de type positif sur les sphères," *Advances in Mathematics*, vol. 31, pp. 16–30, 1979.

[55] E. Kushilevitz, N. Linial, and R. Ostrovsky, "The linear array conjecture of communication complexity is false," in *Proceedings of the 28th ACM Symposium on the Theory of Computing*, ACM, 1996.

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

[57] T. Lee, G. Schechtman, and A. Shraibman, "Lower bounds on quantum multiparty communication complexity," in *Proceedings of the 24th IEEE Conference on Computational Complexity*, IEEE, 2009.

[58] T. Lee and A. Shraibman, "An approximation algorithm for approximation rank," in *Proceedings of the 24th IEEE Conference on Computational Complexity*, IEEE, 2008. arXiv:0809.2093 [cs.CC].

[59] T. Lee and A. Shraibman, "Disjointness is hard in the multiparty number-on-the-forehead model," *Computational Complexity*, vol. 18, no. 2, pp. 309–336, 2009.

[60] T. Lee, A. Shraibman, and R. Špalek, "A direct product theorem for discrepancy," in *Proceedings of the 23rd IEEE Conference on Computational Complexity*, pp. 71–80, IEEE, 2008.

[61] T. Lee, A. Shraibman, and S. Zhang, "Composition theorems in communication complexity," Manuscript, 2009.

[62] N. Linial, S. Mendelson, G. Schechtman, and A. Shraibman, "Complexity measures of sign matrices," *Combinatorica*, vol. 27, no. 4, pp. 439–463, 2007.

[63] N. Linial and A. Shraibman, "Learning complexity versus communication complexity," *Combinatorics, Probability, and Computing*, vol. 18, pp. 227–245, 2009.

[64] N. Linial and A. Shraibman, "Lower bounds in communication complexity based on factorization norms," *Random Structures and Algorithms*, vol. 34, pp. 368–394, 2009.

[65] L. Lovász, "On the ratio of optimal integral and fractional covers," *Discrete Mathematics*, vol. 13, pp. 383–390, 1975.

[66] L. Lovász, "Communication complexity: A survey," in *Paths, flows, and VLSI-layout*, (B. Korte, L. Lovász, H. Prömel, and A. Schrijver, eds.), pp. 235–265, Springer-Verlag, 1990.

[67] L. Lovász and M. Saks, "Möbius functions and communication complexity," in *Proceedings of the 29th IEEE Symposium on Foundations of Computer Science*, pp. 81–90, IEEE, 1988.

[68] C. Lund and M. Yannakakis, "On the hardness of approximating minimization problems," *Journal of the ACM*, vol. 41, no. 5, pp. 960–981, 1994.

[69] K. Mehlhorn and E. Schmidt, "Las Vegas is better than determinism in VLSI and distributed computing," in *Proceedings of the 14th ACM Symposium on the Theory of Computing*, pp. 330–337, ACM, 1982.

[70] I. Newman, "Private versus common random bits in communication complexity," *Information Processing Letters*, vol. 39, pp. 67–71, 1991.

[71] M. Nielsen and I. Chuang, *Quantum Computation and Quantum Information.* Cambridge University Press, 2000.

[72] N. Nisan and M. Szegedy, "On the degree of Boolean functions as real polynomials," *Computational Complexity*, vol. 4, pp. 301–313, 1994.

[73] N. Nisan and A. Wigderson, "A note on rank vs. communication complexity," *Combinatorica*, vol. 15, no. 4, pp. 557–566, 1995.

[74] J. Orlin, "Contentment in graph theory: Covering graphs with cliques," *Indigationes Mathematicae*, vol. 80, pp. 406–424, 1977.

[75] R. Paturi, "On the degree of polynomials that approximate symmetric Boolean functions (preliminary version)," in *Proceedings of the 24th ACM Symposium on the Theory of Computing*, pp. 468–474, ACM, 1992.

[76] R. Paturi and J. Simon, "Probabilistic communication complexity," *Journal of Computer and System Sciences*, vol. 33, no. 1, pp. 106–123, 1986.

[77] R. Raz, "Fourier Analysis for Probabilistic Communication Complexity," *Computational Complexity*, vol. 5, no. 3/4, pp. 205–221, 1995.

[78] R. Raz, "Exponential separation of quantum and classical communication complexity," in *Proceedings of the 31st ACM Symposium on the Theory of Computing*, pp. 358–367, ACM, 1999.

[79] R. Raz, "The BNS-Chung criterion for multi-party communication complexity," *Computational Complexity*, vol. 9, no. 2, pp. 113–122, 2000.

[80] R. Raz and B. Spieker, "On the log rank conjecture in communication complexity," *Combinatorica*, vol. 15, no. 4, pp. 567–588, 1995.

[81] A. Razborov, "On submodular complexity measures," in *Boolean Function Complexity*, (M. Paterson, ed.), pp. 76–83, Cambridge University Press, 1992.

[82] A. Razborov, "On the distributional complexity of disjointness," *Theoretical Computer Science*, vol. 106, pp. 385–390, 1992.

[83] A. Razborov, "Quantum communication complexity of symmetric predicates," *Izvestiya: Mathematics*, vol. 67, no. 1, pp. 145–159, 2003.

[84] A. Razborov and A. Sherstov, “The sign rank of AC^0,” in *Proceedings of the 49th IEEE Symposium on Foundations of Computer Science*, pp. 57–66, IEEE, 2008.

[85] J. Reeds, “A new lower bound on the real Grothendieck constant,” Available at http://www.dtc.umn.edu/ reedsj, 1991.

[86] I. Schur, “Bemerkungen zur Theorie beschränkten Bilinearformen mit unendlich vielen Veränderlichen,” *Journal für die Reine und Angewandte Mathematik*, vol. 140, pp. 1–28, 1911.

[87] R. Shaltiel, “Towards proving strong direct product theorems,” *Computational Complexity*, vol. 12, no. 1–2, pp. 1–22, 2003.

[88] A. Sherstov, “Communication lower bounds using dual polynomials,” *Bulletin of the EATCS*, vol. 95, pp. 59–93, 2008.

[89] A. Sherstov, “Halfspace matrices,” *Computational Complexity*, vol. 17, no. 2, pp. 149–178, 2008.

[90] A. Sherstov, “The pattern matrix method for lower bounds on quantum communication,” in *Proceedings of the 40th ACM Symposium on the Theory of Computing*, pp. 85–94, ACM, 2008.

[91] A. Sherstov, “The unbounded-error communication complexity of symmetric functions,” in *Proceedings of the 49th IEEE Symposium on Foundations of Computer Science*, IEEE, 2008.

[92] A. Sherstov, “Separating AC^0 from depth-2 majority circuits,” *SIAM Journal on Computing*, vol. 38, no. 6, pp. 2113–2129, 2009.

[93] Y. Shi and Z. Zhang, “Communication complexities of XOR functions,” *Quantum information and computation*, 2009.

[94] Y. Shi and Y. Zhu, “Quantum communication complexity of block-composed functions,” *Quantum Information and Computation*, vol. 9, no. 5, 6, pp. 444–460, 2009.

[95] B. Tsirelson, “Quantum analogues of the Bell inequalities: The case of two spatially separated domains,” *Journal of Soviet Mathematics*, vol. 36, pp. 557–570, 1987.

[96] F. Unger, “Noise in classical and quantum computation and non-locality,” PhD thesis, University of Amsterdam, 2008.

[97] C. van Nuffelen, “A bound for the chromatic number of a graph,” *American Mathematical Monthly*, vol. 83, pp. 265–266, 1976.

[98] L. Vandenberghe and S. Boyd, “Semidefinite Programming,” *SIAM Review*, vol. 38, pp. 49–95, 1996.

[99] S. Vavasis, “On the complexity of nonnegative matrix factorization,” Technical Report arXiv:0708.4149 [cs.NA], ArXiv, 2007.

[100] M. Yannakakis, “Expressing combinatorial optimization problems by linear programs,” *Journal of Computer and System Sciences*, vol. 43, no. 3, pp. 441–466, 1991.

[101] A. Yao, “Some complexity questions related to distributive computing,” in *Proceedings of the 11th ACM Symposium on the Theory of Computing*, pp. 209–213, ACM, 1979.

[102] A. Yao, “Lower bounds by probabilistic arguments,” in *Proceedings of the 24th IEEE Symposium on Foundations of Computer Science*, pp. 420–428, 1983.

[103] A. Yao, "On ACC and threshold circuits," in *Proceedings of the 31st IEEE Symposium on Foundations of Computer Science*, pp. 619–627, IEEE, 1990.

[104] A. Yao, "Quantum circuit complexity," in *Proceedings of the 34th IEEE Symposium on Foundations of Computer Science*, pp. 352–360, IEEE, 1993.