Project Proposal: Lower bounds on Concrete Complexity

Anup Rao
University of Washington
anuprao@cs.washington.edu

Project Description

A key stumbling block in our attempt to understand the algorithmic complexity of many problems is our inability to prove lower bounds. It is quite possible that there are *linear* time algorithms for problems that we consider hard, such as boolean satisfiability, even though we do not even know of *subexponential* time algorithms for them[1]. Closing this gap in our knowledge is of fundamental importance. This proposal is directed at finding techniques to prove better lower bounds on complexity.

Given this failure to prove lower bounds on algorithmic complexity, it is natural to try and prove lower bounds on models of computation that are easier to reason about. This has been an enduring line of research in complexity theory. Sometimes, as in the study of *bounded depth circuits*, the models are weaker than general algorithms. Sometimes, as in the study of *communication complexity*, the models are much stronger than algorithms. And sometimes, as in the study of *data structures*, the models are of incomparable computational strength. We hope that developing a theory that is able to prove lower bounds on these diverse, yet theoretically tractable models will equip us with techniques that are relevant to all computational models. Proving new lower bounds on algorithms is a mountain that we eventually need to scale; my goal is to identify the next steps on this difficult journey.

I propose to investigate questions about communication complexity, data structures, certain restricted families of boolean circuits, and the extension complexity of linear programs. I make the case that there is a rich interplay of ideas relevant to proving lower bounds on these models, even though the models themselves seem very different. Moreover, I am uniquely placed to make the connections and advance this research program—I already have a track record of success with answering some fundamental questions in these domains. My focus is on simple, concrete models of computation. This is intentional—my main goal is to discover widely applicable techniques, and I believe that methods developed for simple models are more likely to be widely applicable. Some of the models, like linear programs and data structures, are extremely relevant to practice. Others, like communication complexity, are more theoretically motivated. I believe that all of them must be studied in any comprehensive theory of computation.

Communication Complexity

Communication is at the heart of most computational processes, and so the study of *communication complexity* (introduced by Yao [99]) has had many applications in computer science. It is fair to say that no other single model has had a bigger impact on the project of proving lower bounds in computational complexity. The model is general enough that it captures something essential about all computational processes, yet simple and natural enough that beautiful ideas from a wide range of mathematical disciplines can be used to prove powerful lower bounds. Lower bounds on communication

[1] We do know that there are functions that require exponential time to compute—this follows from the *time hierarchy* theorem. The challenge is to develop methods that can prove useful lower bounds for functions that we actually want to compute.

have been translated to lower bounds on boolean circuits, proof complexity, data structures, linear programs, distributed systems and streaming algorithms, and the list is still growing. The ideas used to prove lower bounds draw from results in probability theory, information theory, analysis, geometry and algebra.

Over the past few years, I have made a serious effort to advance the state of the art in communication complexity. My collaborators and I have found the best known results to several fundamental longstanding open problems, and new connections to other models of computation. I have been an organizer for semester long programs and workshops on the subject. Amir Yehudayoff and I have written a book about communication complexity and its applications[2]. We have worked hard to ensure that proofs appearing in the book are significantly simpler and easier to understand than the proofs that appear in the literature. I elaborate on these efforts in the Broad Impacts section.

Communication complexity measures the number of bits two or more parties need to send each other in order to compute some joint function of their inputs. Next, I discuss the relevant history for a couple of fundamental directions of research in communication complexity.

Direct Sums and Information Complexity

The *direct sum* question is a natural question one can ask about any computational model. Given a function $f(x)$, we define the function $f^k(x_1, \ldots, x_k) = (f(x_1), f(x_2), \ldots, f(x_k))$. Then the direct sum question is: How does the complexity of computing f^k relate to the complexity of computing f? If f has complexity c, what is the best lower bound we can prove on the complexity of f^k?

The goal is to prove results about direct sums that hold for *all* choices of f—such a result identifies something fundamental about the model of computation. In most reasonable models of computation, the complexity of computing f^k is at most $c \cdot k$. For models like boolean circuits, where the measure of complexity is the size of the smallest circuit computing f, we know of examples where the complexity of computing f^k can be significantly smaller[3] than $c \cdot k$. In fact, if one could show that the circuit complexity of computing f^k is $\omega(c \cdot \log^2 k)$, this would be a major breakthrough—it would imply that there is no linear time algorithm for matrix multiplication.

For deterministic 2-party communication complexity, if f depends on n bits and has complexity c, Feder, Kushilevitz, Naor and Nisan showed that at least $\Omega(k(\sqrt{c} - \log n))$ communication is required for f^k [36]. No stronger lower bound is known. Alon and Orlitsky [5] showed that if the parties are trying to compute a relation rather than a function, then there is an example where computing k copies requires only $O(c+k)$ bits of communication. When the 2 parties are allowed to use randomized protocols, the best known lower bound was proved by the PI along with Barak, Braverman and Chen. We showed that f^k requires $\Omega(c\sqrt{k}/\log(ck))$ bits of communication [12]. Subsequently, with Braverman, Weinstein and Yehudayoff, we built on these ideas and new observations of Jain, Pereszlényi and Yao [51] to show the stronger result that if c bits of communication are required to compute f with probability of success 2/3, then any protocol that uses communication $\ll c\sqrt{k}/\log(ck)$ can compute f^k with probability at most $2^{-\Omega(k)}$ [23]. When the measure of complexity is average-case, and the inputs are promised to be independent of each other, we prove stronger bounds $(\Omega(ck/\log^2(ck)))$. The results I was involved in proving are the best known lower bounds for direct sums in randomized communication complexity today.

The lower bounds on direct sums in randomized communication complexity were obtained by studying a new definition—the notion of the information complexity of a protocol. The evolution of this definition is illustrative of the broad impact good definitions can have. As far as I know, the story begins with attempts to prove lower bounds on the randomized communication complexity of disjointness, an important problem in communication complexity. Kalyanasundaram and Schnitger were the first to prove that disjointness requires linear randomized communication [53]. Their proof involved

[2] A draft of the book can be seen here: `https://homes.cs.washington.edu/~anuprao/pubs/book.pdf`.

[3] The examples rely on fast algorithms for matrix multiplication. One can show that there is are many $n \times n$ matrices A for which computing $f(x) = Ax$ requires a circuit of size $\Omega(n^2/\log n)$ when x is an n-bit column vector. Yet for any such f, f^n can be computed with a circuit of size $\ll n^3/\log n$ using the best known algorithms for matrix multiplication.

measuring information in communication protocols by reasoning about Kolmogorov complexity. The proof was subsequently simplified by Razborov [79], who used Shannon's entropy function, and then by Bar-Yossef, Jayram, Kumar and Sivakumar [11], who used Shannon's mutual information. Similar concepts were used by Raz [77] (subsequently simplified by Holenstein [48]) in a technical tour de force to prove the parallel repetition theorem, a key result in the study of hardness of approximation and probabilistically checkable proofs. I first became familiar with these concepts when I used them to prove stronger bounds for the parallel repetition theorem [74]. Meanwhile, Chakrabarti, Shi, Worth and Yao had made attempts to answer the direct sum question for randomized communication using ideas from information theory [25]. In [12], we gave a new definition for the information content of communication protocols that was inspired by the definitions used to prove the parallel repetition theorem. We combined this definition with an approach following the intuition of [25] to prove our lower bounds for the direct sum problem. So, the ideas for understanding the flow of information originated in the study of disjointness in communication complexity, made their way to the world of approximation algorithms and PCP's, and eventually came back to answer questions about direct sums in randomized communication complexity. Following our work on direct sums, Braverman and Garg took the ideas back in the other direction once again, giving new bounds for the parallel repetition theorem [19]. Thus, there has been a healthy exchange of ideas between the study of fundamental concepts in communication complexity and ideas useful to prove results about the hardness of approximation.

The result of this evolution of ideas was the definition of the *internal information* of a communication protocol. If X, Y are the inputs to a communication protocol, and M is the variable containing the public randomness and messages of the protocol, the internal information is $I = I(M; X|Y) + I(M; Y|X)$. Here $I(\cdot;\cdot)$ is Shannon's notion of mutual information. Each of the terms in this expression measures the information learnt by one of the parties during the execution of the protocol. Today it is clear that this is the right definition for the information of a 2-party protocol—it allows one to prove results like the direct sum theorem and reason about combinatorial problems like disjointness. Braverman and I [22] showed that under this definition, information complexity is the same as amortized communication complexity, giving a very natural equivalent definition of the same quantity.

If a protocol has communication c and internal information I, how many bits of communication are required to simulate the protocol? This is a fundamental question in its own right. Braverman and I showed that proving optimal direct sum theorems is equivalent to finding ways to compress protocols according to their internal information [22]. Ever since Shannon's seminal paper [83], we have known how to encode a message so that its length is essentially the same as its entropy. Can we get a similar savings in an interactive setting with the notion of information defined above? In [12], we showed that the compression can be carried out with $O(\sqrt{cI}\log c)$ bits of communication. For the case of single messages, the Braverman and I showed that near optimal compression of $I + o(I)$ is possible [22]. Subsequently, Braverman gave a different simulation with $2^{O(I)}$ bits of communication [18]. My student Ramamoorthy and I showed that if the information learnt by one of the parties is I_a, and the information learnt by the second party is I_b, then the simulation can be carried out with $O(I_a + I_b^{1/4} \cdot c^{3/4} \log c)$ bits of communication, a result that is useful when understanding protocols where the information revealed is asymmetric.

All of these results are about efficient ways to compress communication protocols. Ganor, Kol and Raz [41, 42] were the first to place non-trivial limits on how strong such compression schemes could be. They showed that there is a function that can be computed with information I, but cannot be computed with communication $2^{o(I)}$, proving that Braverman's simulation is tight, and showing that one cannot hope for the strongest possible direct sum theorem in randomized communication complexity. Their proof implies that for every $\epsilon > 0$, there is a function f with communication complexity c so that f^k can be computed with ϵck bits of communication. My student Sinha and I simplified the proof of this separation between information and communication [75]. It remains open to understand whether or not the other compression schemes, most notably that of [12], are tight.

Disjointness and the Number-on-Forehead Model

If there is a most important communication problem, it is the *set disjointness* problem. Each of k parties is given a subset of $\{1, 2, \ldots, n\}$ as input, and they wish to determine whether or not the intersection of all of their sets is empty. As we discussed above, attempts to prove lower bounds on the communication complexity of disjointness have led to the discovery of fundamentally new techniques like information complexity. Moreover, many lower bounds in other models, like data structures, extension complexity, streaming algorithms and distributed computing are proved by reduction to the lower bounds on the communication complexity of set disjointness.

When the number of parties $k = 2$, the key idea for all known lower bounds on the randomized communication complexity of disjointness [53, 79, 11, 21] is to count the information: one can identify a distribution on inputs and prove that any protocol with small communication must be devoting only a small fraction of its communication towards a typical element of $\{1, 2, \ldots, n\}$. These arguments give the optimal lower bounds for randomized 2-party protocols computing disjointness.

The *number-on-forehead* model is the most important communication model [26] where optimal lower bounds still elude us. Here there are k parties, and each party has an n-bit input written on their *forehead*—each party can see the $k - 1$ inputs written on the other parties' foreheads, but not their own. This is an extremely powerful model. For example, k parties can compute any degree $k - 1$ polynomial of their inputs very cheaply in this model—the variables of each monomial in such a polynomial are completely visible to one of the parties, so the parties can exress the polynomial as a sum of k polynomials that are computable by the parties. Then communication is needed only to compute the sum of the polynomials, which can be done cheaply.

Lower bounds on multiparty communication complexity are important because several computational models such as circuits, branching programs, and propositional proofs can be used to obtain efficient communication protocols. Strong enough communication complexity lower bounds for the computation of any explicit function can therefore be used to prove lower bounds on these models [26, 10, 30, 9, 78, 96]. In particular, lower bounds on the communication complexity of disjointness have found applications to proof systems [14], circuit lower bounds [46, 81, 95, 50], lower bounds on communication for problems related to combinatorial auctions [31, 64, 63, 32, 45, 66], lower bounds in distributed computing [33], and oracle separations for complexity classes [1] (see the survey [29]). Proving lower bounds for the number-on-forehead model is much more challenging, because the parties share a lot of common information. Although we can prove optimal lower bounds in the number-in-hand model, here the best known lower bounds are of the type $n/2^{\Theta(k)}$. Counting arguments show that there are functions that require communication $\Omega(n)$, but we do not know of specific functions that such large communication. The information theory based approaches used in the past seem to run into serious technical obstacles because of the strong correlations between the inputs.

Grolmusz described a clever protocol for disjointness in the number-on-forehead model, where the parties need to communicate only $O(\log^2 n + k^2 n/2^k)$ bits [44, 8].

A wide range of ideas have been employed to prove increasingly stronger lower bounds on the communication complexity of disjointness in the number-on-forehead model. When k is large, Tesson [92] and Beame, Pitassi, Segerlind and Wigderson [15] proved that the deterministic communication complexity is $\Omega(\log(n)/k)$. Sherstov [85, 86] introduced the *pattern matrix method* for proving lower bounds in the case $k = 2$. The method was used to separate certain circuit classes by relating their complexity to analytic properties of boolean functions, like their approximate degree. This technique was generalized to $k > 2$ by Chattopadhyay [27], Lee and Shraibman [59], and Chattopadhyay and Ada [28]. These last two papers proved lower bounds of the type $\Omega\big(n^{1/(k+1)}/2^{2^{O(k)}}\big)$ on the randomized communication complexity of disjointness. Beame and Huynh-Ngoc [13] extended these methods further to prove that the randomized communication complexity is at least $2^{\Omega\left(\sqrt{\log(n)/k}\right)}2^{-k}$. Sherstov [87, 88] proved the best known lower bounds on randomized communication complexity of $\Omega(\sqrt{n}/(k2^k))$. These results use powerful techniques such as Fourier analysis, Gowers norms, directional derivatives, and bounds on the approximate degree.

In a somewhat anticlimactic paper, Yehudayoff and I [76] proved that the deterministic communication complexity of disjointness is at least $\Omega(n/4^k)$, which almost matches Grolmusz's upper bound. In addition, we recovered Sherstov's bounds of $\Omega(\sqrt{n}/(k2^k))$ with a simpler (though still sophisticated) proof. Surprisingly, our deterministic lower bound does not use most of the complicated methods developed by past work—it uses a discrepancy based argument with a counterintuitive twist. It remains open to prove the optimal lower bound of $n/2^{\Theta(k)}$ on the randomized communication complexity.

Research Directions

Goal: New approaches to protocol compression. As far as we know, there could be a method to compress protocols giving communication at most $I \log C$. Such strong results were already proved by [12] for external information. Recently Sherstov [89] gave better bounds for product distributions, and Braverman and Kol [20] gave improved bounds for external information. These ideas seem to fall short of proving better bounds for internal information, but they give new ways to think about these problems.

Goal: Counterexamples to optimal compression. It is possible that known compression results are already optimal. In some ways this would be an even more interesting outcome to this line of research. The ideas developed in [41, 42, 75] suggest a way to generalize the counterexample to prove the stronger lower bounds required to show that [12] is tight. The example is too technical to discuss here.

Goal: Compression for multiparty protocols. All of our work on compression and direct sums had to do with two party communication complexity. One could ask similar questions for the multiparty scenario. While our results that have to do with information measured from the point of view of an external observer do carry over to the multiparty setting, the direct sum results do not carry over. The problem seems to be a matter of definition. The naive extension of the definition of information from the viewpoint of the parties, namely $I(XYZ; M|X) + I(XYZ; M|Y) + I(XYZ; M|Z)$, where X, Y, Z are the inputs and M are the messages in the protocol, does not make sense—one can show that every boolean function can be computed with constant information when the inputs X, Y, Z are distributed according to the number on the forehead model. It is an open question to find the right definitions for this scenario. I feel that this question is closely related to the goal of proving better lower bounds on the number-on-forehead model, which I discuss below.

Goal: Randomized lower bounds for disjointness in the number-on-forehead model. Can we prove lower bounds on the randomized communication to match our deterministic lower bounds [76]? The linear lower bounds for $k = 2$ are based on methods from information theory, and it is not clear why similar methods will not eventually give lower bounds in the number-on-forehead setting.

Let me outline a new approach towards obtaining such lower bounds that I am pursuing. Here I discuss the approach at a high level, omitting many details. A useful fact is that when $k = 2$, the set of inputs x, y that lead to a given sequence of messages forms a *combinatorial rectangle*: the set is of the form $A \times B$ for some sets A and B. Rectangles have the extremely useful feature that if x, y are independent, then they remain independent even if we condition on the event that they belong to the rectangle. This fact is used crucially in all information based methods to proving lower bound. Now when $k > 2$, the relevant set of inputs has less structure—it is a *cylinder intersection*. This is a set whose indicator function can be expressed as a product of functions $\chi_1 \cdot \chi_2 \cdot \ldots \chi_k$, where χ_i does not depend on the i'th input. Cylinder intersections do not share the crucial feature of combinatorial rectangles discussed above. For example, if $k = 3$, and x, y, z are mutually independent, then conditioning on their being in a cylinder intersection can make them dependent. Thus, we cannot rely on this property to give lower bounds for cylinder intersections.

My approach is to rely on a different feature shared by both cylinder intersections and rectangles. Suppose $x_0, x_1, y_0, y_1, z_0, z_1 \subseteq \{1, 2, \ldots, n\}$ are 6 sets. Then call the set of tuples

$$P = \{x_0, x_1\} \times \{y_0, y_1\} \times \{z_0, z_1\} - \{(x_1, y_1, z_1)\}$$

a *punctured cube*. It is easy to see that if a cylinder intersection contains P, then it must contain (x_1, y_1, z_1) as well. Indeed, this is (nearly) a defining feature of cylinder intersections. The analogous fact holds for rectangles on x, y using the analogous definition for 2 dimensional punctured cubes.

Suppose x, y, z are uniformly random disjoint sets, and S is an arbitrary set such that

$$\mathbb{P}[(x, y, z) \in S] \geq 2^{-n/100}.$$

Then I conjecture that S must contain a punctured cube P as above, and moreover there must be an element j for which $j \notin x_0, y_0, z_0$ but $j \in x_1, y_1, z_1$. I know how to use information based methods to prove the corresponding statement when $k = 2$. This leads to a new proof of the disjointness lower bound that seems much more amenable to generalization to the cases where $k > 2$. If this approach works, it would give a way to prove lower bounds on the randomized communication complexity of disjointness when $k > 2$. We would let S be the large cylinder intersections witnessing that the inputs are disjoint, and then the conjecture can be used to show that the error of the protocol must be too large.

Goal: Optimal lower bounds in the number-on-forehead model. Essentially the only known technique for proving lower bounds on the number-on-forehead model relies on using the Cauchy-Schwartz inequality k times [10]. This kind of proof can prove a lower bound of at most $n/2^k$. It is a tantalizing goal to adapt information based methods to this setting and so prove stronger lower bounds. A first step is to carry out the approach described above for proving lower bounds on disjointness. If that approach is successful, then I would consider it feasible to generalize the approach to proving stronger lower bounds for functions other than disjointness.

Data Structures

The concept of a data structures is one of the most basic concepts in computer science, with widespread applications from increasing the efficiency of algorithms to providing efficient databases. The *cell-probe* model, introduced by Yao [100], is the standard way to model data structures when proving lower bounds—it is simple, yet general enough that any reasonable implementation of a data structure can be described with it. The data structure is represented as an array of cells, each containing w bits. Each access to an element of the array takes one unit of time. In *static* data structure problems, an input x is preprocessed to give the array. Then, at run-time, an algorithm answers a query y by accessing this array to compute a function $f(x, y)$. The parameters of interest are the space (length of the array), and the time (number of accesses required to answer queries). In *dynamic* data structure problems, we maintain an object x. The algorithm of the data structure allows both updates to the data x as well as queries to it. For example, we may want to maintain a graph while retaining the ability to quickly add and remove edges, and ask whether two vertices are connected or not. The key parameter of interest for dynamic data structures is the time (number of cell reads) required to carry out all operations on the data.

Methods from communication complexity have found great success in proving lower bounds on the space and time complexity for many data structure problems. A key observation, first made by Miltersen, is that every data structure involves a communication protocol [62, 61]. The parties get the inputs x, y involved in the static problem. If the data structure makes t queries to an array of size s to compute $f(x, y)$, then we obtain a protocol with t rounds of communication. In each round, the party holding x sends $\log s$ bits to indicate the location of the cell being read, and the second party responds with the contents of the corresponding cell. This reduction does loses something: one can never prove that $t \log s > |x|$ using such an approach, because the trivial communication protocol where the first party sends x can always be used to compute $f(x, y)$ in the communication world. Consequently, if y is an n-bit input, and the number of choices for x is polynomial[4] in n, then the best lower bound on

[4]This is the most common setting of parameters for practically motivated data structure problems.

time that one can hope for using communication is at most $O(\log n)$. One can use this approach and lower bounds on the disjointness problem to give many tight lower bounds [68, 69].

A second way to leverage information theory based methods to prove lower bounds was found by Fredman and Saks [39], who invented the *chronogram* technique to prove lower bounds on dynamic data structures. These ideas were used extensively by others to prove lower bounds on a wide range of problems [70, 67, 71, 57, 101, 97]. A third technique to prove lower bounds, called the *cell-sampling* technique, was pioneered by Panigrahy, Talwar and Wieder [65]. In a breakthrough, Larsen [57] combined the chronogram technique with cell-sampling to prove lower bounds of the type $\Omega(\log^2 n/\log\log n)$ for dynamic data structure problems. These remain the strongest lower bounds that we know how to prove in the dynamic setting.

I find data structure lower bounds attractive because many basic questions about the complexity of data structures remain completely wide open, including problems that are very relevant to practice. This suggests that there is something fundamentally new that one must discover in order to prove lower bounds. For example, we do not know whether or not it is possible to have a deterministic dynamic data structure that maintains a set of numbers $S \subseteq \{1, 2, \ldots, n\}$, where all operations take $O(1)$ time, and the word size is $O(\log n)$. I suspect that the answer is that it is impossible to do this. Here the operations we would like to support include inserting new numbers into S, deleting numbers from S and querying whether or not a number belongs to S. There is a simple randomized data structure that can do this, but is there a deterministic one? An even harder task is to maintain such a set and allow for queries that compute the minimum element or median element of the set. Again, as far as we know, all of these operations can be supported in $O(1)$ time. The best known data structure for these problems is the van Emde Boas tree, which can handle all such queries in time $O(\log\log n)$ [94]. It is curious that such basic questions about simple algorithmic primitives remain unresolved.

Together with Ramamoorthy [73], we made some progress in addressing some of these questions under the assumption that the data structures are required to be *non-adaptive*. Prior to our work, no other lower bounds were known for these problems in the cell-probe model. An operation is non-adaptive if the identity of the cells being accessed is determined by the operation being performed and independent of the contents of the cells read during the operation. Non-adaptive data structures are useful in practice, because they require only one parallel time step to load the contents of cells into memory. If S is maintained using a binary search tree, then numbers can be inserted, deleted and queried non-adaptively in time $O(\log n)$, and one can compute the minimum and median of the numbers adaptively in time $O(\log n)$. We showed that any data structure that handles insertions non-adaptively cannot do much better—we prove that the time required for some operation must be at least $\Omega(\log n/\log\log n)$. We obtained a number of different bounds for other basic problems based on the same approach. Several past works have proved lower bounds on various computational models under the assumption of non-adaptivity (see for example [55]). In the context of data structures, Brody and Larsen [24] showed polynomial lower bounds for various dynamic problems in the non-adaptive setting. Among other results, they showed that any data structure for reachability in directed graphs that non-adaptively checks for reachability between pairs of vertices must take time $\Omega(n/w)$, where n is the size of the underlying graph. [4, 43] proved non-adaptive lower bounds on static data structures for the dictionary problem in the bit probe model.

Our results were obtained via an application of the famous sunflower lemma of Erdős and Rado [35]. The sunflower lemma was used in the past to prove lower bounds on dynamic data structures by Frandsen and Milterson [38] and then again for static data structures by Gal and Milterson [40]. The lemma proves that any large collection of sets must contain a *sunflower*—a collection of sets whose pairwise intersections are the same. The common intersection of all the sets in the sunflower is called the core of the sunflower. In our work, we combine the lemma with the chronogram technique of Fredman and Saks to prove our tight lower bounds. Given any data structure for maintaining a set of numbers S as above, we let $X_1, \ldots, X_n$ be sets of cells, where X_i is the set of cells accessed when the algorithm inserts the element i. By the sunflower lemma, this collection of sets must contain a sunflower. Our proof can be viewed as describing a communication process where each party only

reads and writes to the core of the sunflower. Thus, the size of the core is related to the communication in this process. Our lower bounds are proved by showing that the core must be large.

Research Directions

Goal: Understanding the cost of parallelizing data structures, and compressing data structures. If a particular static data structure problem can be solved with m memory, and q queries, and these parameters are best possible, how much memory does it take when the data structure gets k independent queries? Because of the connection to communication complexity listed above, this question is closely related to the direct sum for asymmetric communication protocols. In many respects it is more subtle. It would be interesting to understand if our work on compressing asymmetric communication has application to the data structures problem given above. We have not yet understood the connections clearly.

Goal: Lower bounds on basic data structures maintaining sets of numbers. Our results in [73] only apply to non-adaptive data structures. However, the intuition of the results does make sense to me even for adaptive data structures. Intuitively, if the data structure run in time $O(1)$, then there can be very little information exchanged between subsequent operations of the data structure. There is a communication process hiding in the data structure that one might be able to exploit to prove lower bounds. An important goal is to prove lower bounds for such data structures that can compute the minimum, median or predecessors of numbers.

Goal: Prove superlogarithmic lower bounds for static data structures. As we discussed above, any argument that uses communication complexity to prove a lower bound on static data structures cannot possibly prove superlogarithmic lower bounds. So, an interesting goal is to find a technique that circumvents this barrier. Although communication complexity itself cannot be used to achieve this goal, one can *open the box* of communication complexity, and attempt to use the various tools there to attack this problem. For example, can we directly used methods based on discrepancy or arguments analogous to those that were used to prove direct sum theorems to reason about data structures? These are questions that I find very interesting.

Boolean Circuits

A boolean circuit is a general model of computation for computing functions $f(x)$, where x is an n-bit string. It is a directed acyclic graph where every vertex is called a gate. Each gate has in-degree either 2 or 0. If the gate has in-degree 0, it is labeled by an input variable, otherwise it is labeled by a function mapping 2 bits to 1 bit. Sometimes, the gates are restricted to coming from the De Morgan basis, but I prefer to allow arbitrary gates. The circuit is evaluated by successively evaluating each gate that can be evaluated. The circuit computes a function f if some gate in the circuit evaluates to the value of the function. The size of the circuit is the number of gates, and the depth is the length of the longest path in the graph.

From the perspective of proving lower bounds, boolean circuits are very similar to algorithms—every algorithm with running time $t(n)$ can be simulated by a boolean circuit whose size is $O(t(n) \log t(n))$. The depth of the circuit can be viewed as a measure of how much time it takes to compute the output when the gates are evaluated in parallel. Shannon showed that random functions require depth n and size $\Omega(2^n/n)$ with high probability [84], and Lupanov showed that every boolean function of n bits has circuit depth at most n, and size at most $O(2^n/n)$ [60]. Is there a function that has a polynomial sized circuit yet still requires circuits of depth n? Is there an explicit example of a function that requires circuit size bigger than $O(n)$? These kinds of questions remain far out of reach of our current techniques. It is easy to prove that any function that depends on all n of its inputs requires a circuit of size at least n just to read all the inputs. We know of no argument that beats this trivial bound. In my opinion, these are two of the most important open problems in complexity theory.

There is a huge body of work in understanding bounded-depth circuits. Here the gates are restricted to computing the logical AND, OR, or NOT functions, and the number of alternations between AND and OR's are restricted to being a constant. Note that without restricting the number of alternations in the circuit, the first condition is not really a restriction—one can compute any function of 2 bits with a small number of AND, OR and NOT gates. This work on bounded depth circuits has led to such gems as Hastad's famous switching lemma [47], and Smolensky and Razborov's algebraic methods for proving lower bounds [91, 80]. One can show that even computing the parity function $(\sum_i x_i \mod 2)$ requires exponential sized circuits with such restrictions.

In recent years, I have been thinking about other kinds of reasonable circuit models that one might consider to guide the search for techniques that might lead to lower bounds on general boolean circuits. Pavel Hrubeš and I considered the model of boolean circuits where each gate can compute an arbitrary boolean function of k bits, where k is a parameter [50]. We called these circuits of *medium fan-in*. This model of computation is much stronger than the general model of boolean circuits described above, which corresponds to $k = 2$. Counting arguments show that most functions require exponential size even when $k = n/2$. In our work we managed to prove non-trivial lower bounds when $k = \Theta(n)$. The trivial lower bound says that $O(1)$ gates are sufficient for the circuit to read all of its inputs, whereas our work shows that there are simple functions that require $\Omega(\log^2 n)$ gates. This result is proved by appealing to lower bounds on the number-on-forehead model of communication. Indeed, optimal lower bounds in the number-on-forehead model would lead to lower bounds of the type $\sqrt{n}$ for our circuit model. We also consider circuits of depth 2 in this model, and showed that one can define simple functions that require $n^{\Omega(1)}$ gates.

Indeed, we showed that our model occupies a central place in understanding many other concrete models of computation. For example, our model can simulate oblivious branching programs, and the results we proved recover the best known lower bounds for oblivious branching programs. Moreover, we showed that any lower bound of the type $\Omega(n/\log\log n)$ would imply that the same function cannot be computed by a boolean circuit with $k = 2$, size $O(n)$ and depth $O(\log n)$. Thus, even a sublinear lower bound in our model would imply new lower bounds for the standard model of boolean circuits. This is interesting, because many of the techniques used to prove lower bounds on concrete models, like counting entropy, fail to work exactly because one cannot *fix* the values in the computation without fixing the input. However, if we are only aiming to prove sublinear lower bounds, then these techniques come back into play.

Another direction I have pursued recently is in response to a nice question posed by Kulikov and Podolskii [56]. They consider a model for boolean circuits where every gate can only compute the majority function of at most k bits. If the depth of the circuit is restricted to being 2, how large does k need to be for the circuit to compute the majority of all n of its input bits? The question is related to a long sequence of works in complexity theory. In their famous paper, Ajtai, Komlós and Szemerédi [2] constructed a *sorting network* of depth $O(\log n)$. This is a network that sorts n numbers by comparing pairs of numbers in each step. Each such sorting operation can be simulated by the majority of 2 inputs and a constant, and the majority of n bits is simply the middle number in the sorted order of all the bits. So, their construction shows that the majority of n bits can be expressed as a tree of majorities of depth $O(\log n)$, each taking only 3 bits as input. This was followed by a simple non-explicit construction of such a tree by Valiant [93]. More recently, Allender and Koucký [3] showed that the majority of n bits can be computed by a constant depth circuit with gates that compute the majority of n^ϵ bits, for any constant $\epsilon < 1$. All of this work is intimately connected to understanding the class TC_0 of constant depth circuits using threshhold gates (see [52] for details).

Kulikov and Podolskii found several non-trivial circuits of depth 2 that compute majority for small values of n, and proved that one requires $k \geq \Omega(n^{0.7+o(1)})$. Amano and Yoshida [6] showed that for every odd $n \geq 7$, one can compute the majority with $k = n - 2$. Engels, Garg, Makino and I showed that such a circuit must have $k \geq \Omega(n^{4/5})$. Our lower bound was proved using classical techniques from discrepancy theory. In very recent work, Hrubeš, Ramamoorthy, Yehudayoff and I gave essentially tight bounds [49], proving that $k = \Omega(n)$. Interestingly, our approach is completely different from all

of the results mentioned here. We use algebra and finite fields to reason about these circuits.

There is an important open problem in this domain that looks quite similar to direct sum questions. It was proposed by Karchmer, Raz and Wigderson [54]. Karchmer and Wigderson showed that the circuit depth of a boolean function is equivalent to the 2-party communication complexity of a game associated with the function. Given a function f mapping n bits to 1 bit, set f^n to be the function mapping n^2 bits to 1 bit as follows: first apply f n times to each block of n bits, and then apply f one more time to the n outputs obtained. The goal here is to show that if f requires circuits of large depth, then f^n must require circuits of even larger depth. If it takes c bits of communication to solve the corresponding game for f, how many bits must it take to do it for f^n? [54] showed that even a mild lower bound would separate logarithmic depth circuits from polynomial size circuits. In fact, [54] managed to use this approach to successfully separate circuits of depth $O(\log n)$ from polynomial sized circuits, when all circuits only use AND and OR gates.

Research Directions

Goal: Prove new lower bounds on circuits of medium fan-in. This is likely to be a challenging task, because any new lower bounds would imply new lower bounds on well studied models like oblivious branching programs. Moreover, this model appears to be stuck at bounds that require new lower bounds on the number-on-forehead model.

Goal: Study other models of depth-2 circuits. Kulikov and Podolskii's question inspires many similar questions about depth-2 circuits. For example, if our goal is to compute the majority of n bits, and each of our circuits can compute an arbitrary symmetric function, can we still prove strong lower bounds on k? What other classes of functions can be studied in this context? The methods developed in [50, 34, 49] are all quite different. They give us a family of techniques that can be applied to answer questions of this type.

Goal: Understand the Karchmer-Raz-Wigderson conjecture. I have returned to the topic of trying to prove bounds on boolean composition on and off over the years, but have never made substantial progress on it. I hope to find a variant of these questions that is both tractable and gives insights that I feel constitute progress on this approach.

Extension Complexity of Linear Programs

Polytopes are subsets of Euclidean space that can be defined by a finite number of linear inequalities. They are fundamental geometric objects that have been studied by mathematicians for centuries. Any $n \times d$ matrix A and $n \times 1$ vector b define the polytope $P = \{x \in \mathcal{R}^d : Ax \leq b\}$.

Besides being fundamental geometric objects, polytopes are useful from the perspective of algorithm design, because many interesting computational problems can be reduced to the problem of optimizing a linear function over some polytope—in other words, by writing a *linear program*. The complexity of solving optimization problems using linear programming is related to the number of facets of the polytope. Therefore, it is important to find polytopes that encode computational problems and have a small number of facets. A generic way to do this is via *extensions* of the polytope. A polytope $Q \subseteq \mathcal{R}^k$ is an extension of a polytope $P \subseteq \mathcal{R}^d$ if there is a linear map $L : \mathcal{R}^k \to \mathcal{R}^d$ such that $L(Q) = P$. The *extension complexity* of P is the minimum number of facets achieved by any extension of P. The point is that efficient linear programming can be carried out using the extension of P instead of P itself. This could give much more efficient linear programs.

Yannakakis was the first to attempt to prove lower bounds on the extension complexity of polytopes that encode algorithmic problems [98]. He used such a lower bound to refute the existence of efficient linear programs for the underlying problem. Starting with the breakthrough work of Fiorini, Massar, Pokutta, Tiwary and de Wolf [37], a sequence of papers has shown how to use techniques for proving lower bounds on communication complexity to prove lower bounds on the extension complexity of

natural polytopes [16, 17, 21]. This approach has led to exponential lower bounds for the extension complexity of the well known correlation polytope. The key observation is that proving lower bounds on extension complexity is equivalent to proving that a particular bivariate distribution associated with the polytope cannot be expressed as a convex combination of product distributions. This very much like showing that a boolean matrix cannot be partitioned into monochromatic rectangles, and the same techniques that are useful to show such results in communication complexity apply here.

One of the most celebrated lower bounds on extension complexity is an exponential lower bound on the extension complexity of the matching polytope, proved by Rothvoß [82]. The matching polytope is the convex hull of all matchings in a graph. Rothvoß's proof is quite complicated. The proof does not directly appear to use the connection to communication complexity described above. My student Sinha and I simplified this proof and brought it into a form that is very close to the information based lower bound on disjointness. This simplification will appear in our upcoming book with Yehudayoff, and is already available in the draft of our book. Sinha [90] went on to significantly strengthen Rothvoß's result, and answer an open question posed by Rothvoß. Sinha showed that if K is any polytope, and M is the matching polytope for graphs of size n, with $M \subseteq K \subseteq (1+\epsilon)M$, then the extension complexity of K must be at least $\binom{n}{O(1/\epsilon)}$. This bound is tight. This proves that not only can linear programming not be used to optimize over the matching polytope, it cannot be used to find *approximate* solutions any more efficiently than by using a known linear program. Sinha's proof opens the door to proving many lower bounds on extension complexity using techniques that I am very familiar with.

Research Directions

Goal: Prove new lower bounds on the extension complexity of polytopes approximating Knapsack. Avis and Tiwary, and Pokutta and Vyve [7, 72] have already shown that the extension complexity of the Max-Knapsack polytope is exponential. Just like for the matching polytope, there is a upper bound on approximate extension for this polytope, of size $\binom{n}{1/\epsilon}$. This suggests that similar techniques as used by Sinha for the matching polytope could be applicable to the Knapsack polytope.

Goal: Give communication complexity based lower bounds on the complexity of SDPs. Lee, Raghavendra and Steurer have managed to prove strong lower bounds on the complexity of semidefinite programs [58]. There is a strong analogy between their work and the work discussed above proving lower bounds on the extension complexity of polytopes. The SDP's correspond to quantum communication protocols. It is an important goal to simplify this connection more concretely and find simple proof for the extension complexity of SDPs.

Broad Impact

Our proposal aims to solve problems in a rich variety of areas. For a process to be considered a *computational process*, it must necessarily move information around in some way. Given this, it is not surprising how useful the methods used in communication complexity have been to proving lower bounds on other models of computation. I believe that I am well placed to further extend these connections. I hope that the work of my collaborators and I will have the effect of making researchers in complexity theory more familiar with the beautiful tools available for in communication complexity.

Educational: I have already developed a graduate level course about the uses of information theory in computer science, and a course on communication complexity. Similar courses have since been taught by my collaborators Mark Braverman at Princeton, and by Amir Yehudayoff at the Technion. I intend to encourage computer science students to learn techniques from information theory, and allow students well versed in information theory to hear about computer science problems that they are well placed to tackle.

Communication Complexity Book: In recent years, communication complexity has emerged as a master method for proving lower bounds in theoretical computer science. This progress has happened so quickly that most researchers who use communication complexity are not aware of all of the different ways it has been applied in different areas. Amir Yehudayoff and I are writing a textbook on communication complexity to give a single reference for all of the basic methods that have been discovered. An early draft of our text is available here: `https://homes.cs.washington.edu/~anuprao/pubs/book.pdf`. The book will be published by Cambridge University Press in 2019. We hope that it becomes a standard resource for communication complexity. Our goal is to keep an online version of the book abreast with new developments in communication complexity, as much as is possible.

Organizational Roles in Research Events: I plan to continue organizing research events to encourage young researchers to enter the target research areas of this project. I have a strong track record of helping to run such events in the past. I discuss these past contributions in the section on Results from Prior NSF support. In the coming years, I plan to continue running a workship at Banff on communication complexity, which has become a meeting point for researchers in areas related to communication complexity. In addition, I would like to run organize additional semesters at the Simons institute on lower bounds and communication complexity.

Relevant Results from Prior NSF Support

Project: Lower bounds on the Complexity of Parallelization This project investigated questions related to parallelizing computation. The goal was to prove lower bounds using information complexity.

Intellectual Merit: In this project, my collaborators and I proved several results relating information complexity and communication complexity. With his student Ramamoorthy, the PI showed how protocols with asymmetric information complexity can be compressed. With my (then) postdoc Hrubes, I proved new lower bounds for the boolean circuits of medium fan-in. My student Sinha proved a new lower bound on the extension complexity of approximations to the matching polytope supported by this project. With Yehudayoff, I proved nearly optimal lower bounds on the communication complexity of disjointness in the number-on-forehead model. Each of these results represents a fundamental advance in our state of knowledge in proving lower bounds. The results are discussed below in the products section.

Broad Impact: I have been involved in organizing several workshops and academic semesters built around the themes discussed in this research proposal. These include a weeklong workshops at Banff International Research Station in 2014 and 2017, a Trimester at the Institute Henri Poincare in 2016, and a semester long program at the Simons Institute in 2018. The research in this proposal has the potential for interaction with several different research communities, most notably information theory. The program we ran at the Institute Henri Poincare was meant to bring together researchers from information theory and computer science. Hundreds of researchers learnt about work happening in both areas. I have also organized sessions at the Information Theory and Applications (ITA) center to communicate research from computer science to information theorists.

Related Products:

- Anup Rao and Sivaramakrishnan Ramamoorthy. How to Compress Asymmetric Communication, CCC 2015. We study the relationship between communication and information in 2-party communication protocols when the information is asymmetric. If I^A denotes the number of bits of information revealed by the first party, I^B denotes the information revealed by the second party, and C is the number of bits of communication in the protocol, we show that one can simulate the protocol using order $I^A + I^B + \sqrt[4]{I^B C^3} + \sqrt{I^B C} \log C$

bits of communication, and one can simulate the protocol using order $I^A \cdot 2^{O(I^B)}$ bits of communication. The first result gives the best known bound on the complexity of a simulation when $I^A \gg I^B, C^{3/4}$. The second gives the best known bound when $I^B \ll \log C$. In addition we show that if a function is computed by a protocol with asymmetric information complexity, then the inputs must have a large, nearly monochromatic rectangle of the right dimensions, a fact that is useful for proving lower bounds on lopsided communication problems.

- Anup Rao and Amir Yehudayoff. Simplified Lower Bounds on the Multiparty Communication Complexity of Disjointness, CCC 2015. We show that the deterministic number-on-forehead communication complexity of set disjointness for k parties on a universe of size n is $\Omega(n/4^k)$. This gives the first lower bound that is linear in n, nearly matching Grolmusz's upper bound of $O(\log^2(n) + k^2 n/2^k)$. We also simplify Sherstov's proof showing an $\Omega(\sqrt{n}/(k2^k))$ lower bound for the randomized communication complexity of set disjointness.
- Boaz Barak, Mark Braverman, Xi Chen and Anup Rao. How to compress interactive communication, SICOMP 2013. Test of time award. We describe new ways to simulate 2-party communication protocols to get protocols with potentially smaller communication. We show that every communication protocol that communicates C bits and reveals I bits of information to the participating parties can be simulated by a new protocol involving at most $\tilde{O}(\sqrt{CI})$ bits of communication. In the case that the parties have inputs that are independent of each other, we get much better results, showing how to carry out the simulation with $\tilde{O}(I)$ bits of communication.
- Makrand Sinha. Lower bounds for Approximating the Matching Polytope. SODA 2018. We prove that any extended formulation that approximates the matching polytope on n-vertex graphs up to a factor of $(1+\epsilon)$ for any $2/n \leq \epsilon \leq 1$ must have at least $\binom{n}{\alpha n}$ defining inequalities, where $0 < \alpha < 1$ is an absolute constant. This is tight, as exhibited by the $(1+\epsilon)$ approximating linear program obtained by dropping the odd set constraints of size larger than $(1+\epsilon)/\epsilon$ from the description of the matching polytope. Previously, a tight lower bound of $2^{\Omega(n)}$ was only known for $\epsilon = O(1/n)$, whereas for $2/n \leq \epsilon \leq 1$, the best known lower bound was $2^{\Omega(1/\epsilon)}$. The key new ingredient in our proof is a close connection to the non-negative rank of a lopsided version of the unique disjointness matrix.
- Anup Rao. Parallel Repetition in Projection Games and a Concentration Bound. STOC 2008. Invited to the SICOMP special issue. If the players cannot win a cooperative two player game with probability better than $(1-\epsilon)$, what's the best they can do in the repeated game? We improve earlier results of [77] and [48], who showed that the players cannot win all copies in the repeated game with probability better than $(1-\epsilon/2)^{\Omega(n\epsilon^2/c)}$ (here c is the length of the answers in the game), to show that the probability of winning all copies is $(1-\epsilon/2)^{\Omega(\epsilon n)}$ as long as the game is a "projection game", the type of game most commonly used in hardness of approximation results.
- Pavel Hrubeš, Anup Rao. Circuits of Medium Fan-in, CCC 2015. We consider boolean circuits in which every gate may compute an arbitrary boolean function of k other gates, for a parameter k. We give an explicit function $f : \{0,1\}^n \to \{0,1\}$ that requires at least $\Omega(\log^2 n)$ non-input gates when $k = 2n/3$. When the circuit is restricted to being layered and depth 2, we prove a lower bound of $n^{\Omega(1)}$ on the number of non-input gates. When the circuit is a formula with gates of fan-in k, we give a lower bound $\Omega(n^2/k \log n)$ on the total number of gates.

  Our model is connected to some well known approaches to proving lower bounds in complexity theory. Optimal lower bounds for the Number-On-Forehead model in communication complexity, or for bounded depth circuits in AC_0, or extractors for varieties over small fields would imply strong lower bounds in our model. On the other hand, new lower bounds for

our model would prove new time-space tradeoffs for branching programs and impossibility results for (fan-in 2) circuits with linear size and logarithmic depth. In particular, our lower bound gives a different proof for a known time-space tradeoff for oblivious branching programs.

As a consequence, we prove that the internal information cost (namely the information revealed to the parties) involved in computing any relation or function using a two party interactive protocol is *exactly* equal to the amortized communication complexity of computing independent copies of the same relation or function. We also show that the only way to prove a strong direct sum theorem for randomized communication complexity is by solving a particular variant of the pointer jumping problem that we define. Our work implies that a strong direct sum theorem for communication complexity holds if and only if efficient compression of communication protocols is possible.

- Mark Braverman, Anup Rao, Omri Weinstein, Amir Yehudayoff. Direct Products in Communication Complexity. FOCS 2013. We give exponentially small upper bounds on the success probability for computing the direct product of any function over any distribution using a communication protocol. Let $\mathsf{suc}(\mu, f, C)$ denote the maximum success probability of a 2-party communication protocol for computing the boolean function $f(x, y)$ with C bits of communication, when the inputs (x, y) are drawn from the distribution μ. Let μ^n be the product distribution on n inputs and f^n denote the function that computes n copies of f on these inputs.

  We prove that if $T \log^{3/2} T \ll (C-1)\sqrt{n}$ and $\mathsf{suc}(\mu, f, C) < \frac{2}{3}$, then $\mathsf{suc}(\mu^n, f^n, T) \leq \exp(-\Omega(n))$. When μ is a product distribution, we prove a nearly optimal result: as long as $T \log^2 T \ll Cn$, we must have $\mathsf{suc}(\mu^n, f^n, T) \leq \exp(-\Omega(n))$.

References

[1] Scott Aaronson and Avi Wigderson. Algebrization: A new barrier in complexity theory. *TOCT*, 1(1), 2009.

[2] M. Ajtai, J. Komlós, and E. Szemerédi. Sorting in $c \log n$ parallel steps. *Combinatorica*, 3:1–19, 1983.

[3] Eric Allender and Michal Koucký. Amplifying lower bounds by means of self-reducibility. *J. ACM*, 57(3):14:1–14:36, 2010.

[4] Noga Alon and Uriel Feige. On the power of two, three and four probes. In *Proceedings of the Twentieth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2009, New York, NY, USA, January 4-6, 2009*. SIAM, 2009.

[5] Noga Alon and Alon Orlitsky. Repeated communication and ramsey graphs. *IEEE Transactions on Information Theory*, 41(5):1276–1289, 1995.

[6] Kazuyuki Amano and Masafumi Yoshida. Depth two (n-2)-majority circuits for n-majority. 2017. `https://www.cs.gunma-u.ac.jp/~amano/paper/maj.pdf`.

[7] David Avis and Hans Raj Tiwary. On the extension complexity of combinatorial polytopes. *Math. Program.*, 153(1):95–115, 2015.

[8] László Babai, Anna Gál, Peter G. Kimmel, and Satyanarayana V. Lokam. Communication complexity of simultaneous messages. *SIAM J. Comput*, 33(1):137–166, 2003.

[9] László Babai, Thomas P. Hayes, and Peter G. Kimmel. The cost of the missing bit: Communication complexity with help. *Combinatorica*, 21(4):455–488, 2001.

[10] László Babai, Noam Nisan, and Mario Szegedy. Multiparty protocols, pseudorandom generators for logspace, and time-space trade-offs. *J. Comput. Syst. Sci.*, 45(2):204–232, 1992.

[11] Ziv Bar-Yossef, T. S. Jayram, Ravi Kumar, and D. Sivakumar. An information statistics approach to data stream and communication complexity. *Journal of Computer and System Sciences*, 68(4):702–732, 2004.

[12] Boaz Barak, Mark Braverman, Xi Chen, and Anup Rao. How to compress interactive communication. *SIAM J. Comput*, 42(3):1327–1363, 2013.

[13] Paul Beame and Dang-Trinh Huynh-Ngoc. Multiparty communication complexity and threshold circuit size of AC^0. In *FOCS*, pages 53–62, 2009.

[14] Paul Beame, Toniann Pitassi, and Nathan Segerlind. Lower bounds for Lovász–Schrijver systems and beyond follow from multiparty communication complexity. *SIAM Journal on Computing*, 37(3):845–869, 2007.

[15] Paul Beame, Toniann Pitassi, Nathan Segerlind, and Avi Wigderson. A strong direct product theorem for corruption and the multiparty communication complexity of disjointness. *Computational Complexity*, 15(4):391–432, 2006.

[16] Gábor Braun, Samuel Fiorini, Sebastian Pokutta, and David Steurer. Approximation limits of linear programs (beyond hierarchies). *Math. Oper. Res.*, 40(3):756–772, 2015.

[17] Gábor Braun and Sebastian Pokutta. The matching polytope does not admit fully-polynomial size relaxation schemes. In *Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2015, San Diego, CA, USA, January 4-6, 2015*, pages 837–846, 2015.

[18] Mark Braverman. Interactive information complexity. In Howard J. Karloff and Toniann Pitassi, editors, *STOC*, pages 505–524. ACM, 2012.

[19] Mark Braverman and Ankit Garg. Small value parallel repetition for general games. In *Proceedings of the Forty-Seventh Annual ACM on Symposium on Theory of Computing, STOC 2015, Portland, OR, USA, June 14-17, 2015*, pages 335–340, 2015.

[20] Mark Braverman and Gillat Kol. Interactive compression to external information. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2018, Los Angeles, CA, USA, June 25-29, 2018*, pages 964–977, 2018.

[21] Mark Braverman and Ankur Moitra. An information complexity approach to extended formulations. In Dan Boneh, Tim Roughgarden, and Joan Feigenbaum, editors, *STOC*, pages 161–170. ACM, 2013.

[22] Mark Braverman and Anup Rao. Sampling with small communication, manuscript. *Manuscript*, 2009.

[23] Mark Braverman, Anup Rao, Omri Weinstein, and Amir Yehudayoff. Direct products in communication complexity. *Electronic Colloquium on Computational Complexity (ECCC)*, 19:143, 2012.

[24] Joshua Brody and Kasper Green Larsen. Adapt or die: Polynomial lower bounds for non-adaptive dynamic data structures. *Theory of Computing*, 11:471–489, 2015.

[25] Amit Chakrabarti, Yaoyun Shi, Anthony Wirth, and Andrew Yao. Informational complexity and the direct sum problem for simultaneous message complexity. In Bob Werner, editor, *Proceedings of the 42nd Annual IEEE Symposium on Foundations of Computer Science*, pages 270–278, Los Alamitos, CA, October 14–17 2001. IEEE Computer Society.

[26] Ashok K. Chandra, Merrick L. Furst, and Richard J. Lipton. Multi-party protocols. In *Proceedings of the 15th Annual ACM Symposium on Theory of Computing, 25-27 April, 1983, Boston, Massachusetts, USA*, pages 94–99, 1983.

[27] Arkadev Chattopadhyay. Discrepancy and the power of bottom fan-in in depth-three circuits. In *FOCS*, pages 449–458, 2007.

[28] Arkadev Chattopadhyay. *Circuits, Communication and Polynomials*. PhD thesis, University of Toronto, 2008.

[29] Arkadev Chattopadhyay and Toniann Pitassi. The story of set disjointness. *SIGACT News*, 41(3):59–85, 2010.

[30] Fan R. K. Chung and Prasad Tetali. Communication complexity and quasi randomness. *SIAM Journal on Discrete Mathematics*, 6(1):110–123, February 1993.

[31] Vincent Conitzer and Tuomas Sandholm. Communication complexity as a lower bound for learning in games. In Carla E. Brodley, editor, *ICML*, volume 69 of *ACM International Conference Proceeding Series*. ACM, 2004.

[32] Shahar Dobzinski and Noam Nisan. Limitations of VCG-based mechanisms. *Combinatorica*, 31(4):379–396, 2011.

[33] Andrew Drucker, Fabian Kuhn, and Rotem Oshman. On the power of the congested clique model. In Magnús M. Halldórsson and Shlomi Dolev, editors, *ACM Symposium on Principles of Distributed Computing, PODC '14, Paris, France, July 15-18, 2014*, pages 367–376. ACM, 2014.

[34] Christian Engels, Mohit Garg, Kazuhisa Makino, and Anup Rao. On expressing majority as a majority of majorities. *Electronic Colloquium on Computational Complexity (ECCC)*, 24:174, 2017.

[35] Paul Erdős and Richard Rado. Intersection theorems for systems of sets. *Journal of London Mathematical Society*, 35:85–90, 1960.

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

[37] Samuel Fiorini, Serge Massar, Sebastian Pokutta, Hans Raj Tiwary, and Ronald de Wolf. Exponential lower bounds for polytopes in combinatorial optimization. *J. ACM*, 62(2):17:1–17:23, 2015.

[38] Gudmund Skovbjerg Frandsen, Peter Bro Miltersen, and Sven Skyum. Dynamic word problems. *J. ACM*, 44(2):257–271, 1997.

[39] Michael L. Fredman and Michael E. Saks. The cell probe complexity of dynamic data structures. In David S. Johnson, editor, *Proceedings of the 21st Annual ACM Symposium on Theory of Computing, May 14-17, 1989, Seattle, Washigton, USA*, pages 345–354. ACM, 1989.

[40] Anna Gál and Peter Bro Miltersen. The cell probe complexity of succinct data structures. *Theor. Comput. Sci*, 379(3):405–417, 2007.

[41] Anat Ganor, Gillat Kol, and Ran Raz. Exponential separation of information and communication. In *FOCS*, pages 176–185. IEEE Computer Society, 2014.

[42] Anat Ganor, Gillat Kol, and Ran Raz. Exponential separation of information and communication for boolean functions. *Electronic Colloquium on Computational Complexity (ECCC)*, 21:113, 2014.

[43] Mohit Garg and Jaikumar Radhakrishnan. Set membership with non-adaptive bit probes. In *34th Symposium on Theoretical Aspects of Computer Science, STACS 2017, March 8-11, 2017, Hannover, Germany*, pages 38:1–38:13, 2017.

[44] Vince Grolmusz. The BNS lower bound for multi-party protocols in nearly optimal. *Information and Computation*, 112(1):51–54, July 1994.

[45] Sergiu Hart and Yishay Mansour. The communication complexity of uncoupled nash equilibrium procedures. In *STOC*, pages 345–353, 2007.

[46] Johan Håstad and Mikael Goldmann. On the power of small-depth threshold circuits. *Computational Complexity*, 1:113–129, 1991.

[47] John Hastad. Almost optimal lower bounds for small depth circuits. *Advances in Computing Research*, 5:143–170, 1989.

[48] Thomas Holenstein. Parallel repetition: Simplifications and the no-signaling case. In *Proceedings of the 39th Annual ACM Symposium on Theory of Computing*, 2007.

[49] Pavel Hrubeš, Sivaramakrishna Natarajan Ramamoorthy, Anup Rao, and Amir Yehudayoff. Personal communication.

[50] Pavel Hrubeš and Anup Rao. Circuits with medium fan-in. In *30th Conference on Computational Complexity, CCC 2015, June 17-19, 2015, Portland, Oregon, USA*, pages 381–391, 2015.

[51] Rahul Jain, Attila Pereszlényi, and Penghui Yao. A direct product theorem for bounded-round public-coin randomized communication complexity. *CoRR*, abs/1201.1666, 2012.

[52] Stasys Jukna. *Boolean Function Complexity - Advances and Frontiers*, volume 27 of *Algorithms and combinatorics*. Springer, 2012.

[53] Bala Kalyanasundaram and Georg Schnitger. The probabilistic communication complexity of set intersection. *SIAM J. Discrete Math*, 5(4):545–557, 1992.

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

[55] Jonathan Katz and Luca Trevisan. On the efficiency of local decoding procedures for error-correcting codes. In *STOC: ACM Symposium on Theory of Computing (STOC)*, 2000.

[56] Alexander S. Kulikov and Vladimir V. Podolskii. Computing majority by constant depth majority circuits with low fan-in gates. In *34th Symposium on Theoretical Aspects of Computer Science*, volume 66 of *LIPIcs*, pages 49:1–49:14. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2017.

[57] Kasper Green Larsen. The cell probe complexity of dynamic range counting. In *Proceedings of the 44th Symposium on Theory of Computing Conference, STOC 2012, New York, NY, USA, May 19 - 22, 2012*, pages 85–94, 2012.

[58] James R. Lee, Prasad Raghavendra, and David Steurer. Lower bounds on the size of semidefinite programming relaxations. In *Proceedings of the Forty-Seventh Annual ACM on Symposium on Theory of Computing, STOC 2015, Portland, OR, USA, June 14-17, 2015*, pages 567–576, 2015.

[59] Troy Lee and Adi Shraibman. Disjointness is hard in the multiparty number-on-the-forehead model. *Computational Complexity*, 18(2):309–336, 2009.

[60] O. B. Lupanov. The synthesis of contact circuits. *Dokl. Akad. Nauk SSSR (N.S.)*, 119:2326, 1958.

[61] Peter Miltersen, Noam Nisan, Shmuel Safra, and Avi Wigderson. On data structures and asymmetric communication complexity. *Journal of Computer and System Sciences*, 57:37–49, 1 1998.

[62] Peter Bro Miltersen. Lower bounds for union-split-find related problems on random access machines. In *Proceedings of the Twenty-sixth Annual ACM Symposium on Theory of Computing*, STOC '94, pages 625–634, New York, NY, USA, 1994. ACM.

[63] Noam Nisan. The communication complexity of approximate set packing and covering. *Lecture Notes in Computer Science*, 2380:868–875, 2002.

[64] Noam Nisan and Ilya Segal. The communication requirements of efficient allocations and supporting prices. *J. Economic Theory*, 129(1):192–224, 2006.

[65] Rina Panigrahy, Kunal Talwar, and Udi Wieder. Lower bounds on near neighbor search via metric expansion. In *51th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2010, October 23-26, 2010, Las Vegas, Nevada, USA*, pages 805–814, 2010.

[66] Christos H. Papadimitriou, Michael Schapira, and Yaron Singer. On the hardness of being truthful. In *FOCS*, pages 250–259. IEEE Computer Society, 2008.

[67] Mihai Pătraşcu. Lower bounds for 2-dimensional range counting. In *Proc. 39th ACM Symposium on Theory of Computing (STOC)*, pages 40–46, 2007.

[68] Mihai Pătraşcu. *Lower bound techniques for data structures*. PhD thesis, Massachusetts Institute of Technology, Dept. of Electrical Engineering and Computer Science, 2008.

[69] Mihai Pătraşcu. Unifying the landscape of cell-probe lower bounds. *SIAM Journal on Computing*, 40(3):827–847, ???? 2011.

[70] Mihai Pătraşcu and Erik D. Demaine. Logarithmic lower bounds in the cell-probe model. *SIAM Journal on Computing*, 35(4):932–963, 2006. See also STOC'04, SODA'04.

[71] Mihai Pătraşcu and Mikkel Thorup. Don't rush into a union: Take time to find your roots. In *Proc. 43rd ACM Symposium on Theory of Computing (STOC)*, pages 559–568, 2011. See also arXiv:1102.1783.

[72] Sebastian Pokutta and Mathieu Van Vyve. A note on the extension complexity of the knapsack polytope. *Oper. Res. Lett.*, 41(4):347–350, 2013.

[73] Sivaramakrishnan Natarajan Ramamoorthy and Anup Rao. Lower bounds on non-adaptive data structures maintaining sets of numbers, from sunflowers. In *33rd Computational Complexity Conference, CCC 2018, June 22-24, 2018, San Diego, CA, USA*, pages 27:1–27:16, 2018.

[74] Anup Rao. Parallel repetition in projection games and a concentration bound. In *Proceedings of the 40th Annual ACM Symposium on Theory of Computing*, 2008.

[75] Anup Rao and Makrand Sinha. Simplified separation of information and communication. *Theory of Computing (to appear)*, 2018.

[76] Anup Rao and Amir Yehudayoff. Simplified lower bounds on the multiparty communication complexity of disjointness. In *30th Conference on Computational Complexity, CCC 2015, June 17-19, 2015, Portland, Oregon, USA*, pages 88–101, 2015.

[77] Ran Raz. A parallel repetition theorem. *SIAM Journal on Computing*, 27(3):763–803, June 1998.

[78] Ran Raz. The BNS-chung criterion for multi-party communication complexity. *Computational Complexity*, 9(2):113–122, 2000.

[79] A. A. Razborov. On the distributed complexity of disjointness. *Theoretical Computer Science*, 106(2):385–390, 14 December 1992. Note.

[80] Alexander A. Razborov. On the method of approximations. In *Proceedings of the 21st Annual ACM Symposium on Theory of Computing, May 14-17, 1989, Seattle, Washigton, USA*, pages 167–176, 1989.

[81] Alexander A. Razborov and Avi Wigderson. $n^{\omega}(logn)$ lower bounds on the size of depth-3 threshold circuits with AND gates at the bottom. *Information Processing Letters*, 45(6):303–307, 1993.

[82] Thomas Rothvoß. A direct proof for lovett's bound on the communication complexity of low rank matrices. *CoRR*, abs/1409.6366, 2014.

[83] Claude E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27, 1948. Monograph B-1598.

[84] Claude E. Shannon. The synthesis of two-terminal switching circuits. *Bell Systems Technical Journal*, 28:59–98, 1949.

[85] Alexander A. Sherstov. Separating AC^0 from depth-2 majority circuits. *SIAM Journal of Computing*, 38(6):2113–2129, 2009.

[86] Alexander A. Sherstov. The pattern matrix method. *SIAM Journal of Computing*, 40(6):1969–2000, 2011.

[87] Alexander A. Sherstov. The multiparty communication complexity of set disjointness. In Howard J. Karloff and Toniann Pitassi, editors, *STOC*, pages 525–548. ACM, 2012.

[88] Alexander A. Sherstov. Communication lower bounds using directional derivatives. In *STOC*, pages 921–930, 2013.

[89] Alexander A. Sherstov. Compressing interactive communication under product distributions. *SIAM J. Comput.*, 47(2):367–419, 2018.

[90] Makrand Sinha. Lower bounds for approximating the matching polytope. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2018, New Orleans, LA, USA, January 7-10, 2018*, pages 1585–1604, 2018.

[91] Roman Smolensky. Algebraic methods in the theory of lower bounds for boolean circuit complexity. In *Proceedings of the 19th Annual ACM Symposium on Theory of Computing, 1987, New York, New York, USA*, pages 77–82, 1987.

[92] Pascal Tesson. Computational complexity questions related to finite monoids and semigroups, 2003.

[93] Leslie G. Valiant. Short monotone formulae for the majority function. *Journal of Algorithms*, 5(3):363–366, September 1984.

[94] Peter van Emde Boas. Preserving order in a forest in less than logarithmic time and linear space. *Information Processing Letters*, 6(3):80–82, June 1977.

[95] Emanuele Viola. Pseudorandom bits for constant-depth circuits with few arbitrary symmetric gates. *SIAM Journal on Computing*, 36(5):1387–1403, 2007.

[96] Emanuele Viola and Avi Wigderson. Norms, XOR lemmas, and lower bounds for polynomials and protocols. *Theory of Computing*, 4(1):137–168, 2008.

[97] Omri Weinstein and Huacheng Yu. Amortized dynamic cell-probe lower bounds from four-party communication. In *IEEE 57th Annual Symposium on Foundations of Computer Science, FOCS 2016, 9-11 October 2016, Hyatt Regency, New Brunswick, New Jersey, USA*, pages 305–314, 2016.

[98] M. Yannakakis. Expressing combinatorial optimization problems by linear programs. *JCSS: Journal of Computer and System Sciences*, 43, 1991.

[99] Andrew Chi-Chih Yao. Some complexity questions related to distributive computing (preliminary report). In *STOC*, pages 209–213. ACM, 1979.

[100] Andrew Chi-Chih Yao. Should tables be sorted? *J. ACM*, 28(3):615–628, 1981.

[101] Huacheng Yu. Cell-probe lower bounds for dynamic problems via a new communication model. In *Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2016, Cambridge, MA, USA, June 18-21, 2016*, pages 362–374, 2016.