

Research statement

Amir Yehudayoff

My main area of research is theoretical computer science, with a focus on algebraic complexity theory. I am also interested in other fields of mathematics, such as probability theory and combinatorics.

1 Past research

1.1 Algebraic complexity

Introduction. A typical question computational complexity theory asks is: what is the most efficient way to perform a given computational task? This question is interesting both from a practical and a theoretical point of view. To answer this question we need (1) to find some way to perform the task, and (2) to show that no other way, clever as it may be, can do better. In other words, we need to upper bound and lower bound the complexity of the task. Upper bounds and lower bounds are intimately related, and the main focus of my work was proving lower bounds.

In many cases, the computational tasks we wish to perform are of algebraic nature. One example of such a task is to solve a set of linear equations, and a well-known solution for it is Gaussian elimination. A more efficient solution was presented by Strassen [28]. His algorithm was later improved a few times, but we still do not know if the best known algorithm is indeed the best one.

Algebraic complexity is the study of such algebraic computational tasks. It is the study of the computation of polynomials, which are a natural family of algebraic functions. The standard model for computing polynomials is called *arithmetic circuits*. Roughly, an arithmetic circuit gets as input either variables or field elements, and is allowed to add and multiply any two polynomials that are already computed. The two common complexity measures for circuits are size and depth. The size corresponds to the time it takes to compute the polynomial, and the depth corresponds to the parallel time it takes. A special type of circuits are *formulas*, which are circuits whose computation graph is a directed tree.

Lower bounds for restricted circuits. When trying to prove lower bounds, we need to argue about all possible ways for computing a given polynomial. Such arguments are hard to find: proving even a lower bound that is asymptotically better than $n \log n$ for the size of arithmetic circuits is an outstanding open problem, where n is the number of variables. As part of our attempts to prove lower bounds, we first try to understand restricted families of circuits.

In a sequence of works with Hrubeš, Raz and Shpilka we studied three types of restrictions. I discuss these restrictions below; each discussion consists of a brief introduction, and a short survey of my research.

Multilinear circuits. Multilinear polynomials are a common and useful family of polynomials (a polynomial is *multilinear* if the degree of each variable in it is at most one). Trying to understand the computation of multilinear polynomials, Nisan and Wigderson [17] defined the notion of multilinear circuits, which are restricted circuits for computing multilinear polynomials.

In [22] we proved a roughly $n^{4/3}$ lower bound for the size of multilinear circuits. No lower bounds of the form $n^{1+\varepsilon}$ were known before. In [23] we introduced a method for proving lower bounds for a certain type of multilinear formulas. This method implies tight exponential lower bounds for the size of *monotone* circuits (a monotone circuit is allowed to use only positive numbers). This method also implies a tight exponential lower bound for the size of *orthogonal multilinear* formulas, that were defined by Aaronson [1] as a model for quantum computation. The ideas of [23] are related to randomness extraction and to communication complexity as well.

Constant depth circuits. A different type of restricted circuits are constant depth circuits. Constant depth arithmetic circuits have been studied extensively [8, 7, 26, 27, 21]. For depth greater than three, for example, only super-linear lower bounds are known [27, 21].

In [24] we proved exponential lower bounds for the size of *constant depth multilinear* circuits, improving over what is known for each restriction – multilinear and constant depth – separately. We also showed a superpolynomial separation between multilinear depth d and depth $d + 1$ circuits, for any constant d .

Homogeneous formulas. A third type of restriction I worked on is homogeneity¹. Strassen showed that any circuit can be efficiently simulated by a homogeneous circuit [29]. We do not know whether the same is true for formulas.

In [10] we showed that multilinear formulas can not be efficiently simulated by homogeneous multilinear formulas. In particular, any efficient simulation of general formulas by homogeneous formulas must violate the multilinearity condition, that is, must transform multilinear formulas into non-multilinear formulas. All known simulations of formulas by homogeneous formulas do not violate the multilinearity condition. We also showed how to compute the symmetric polynomials by smaller formulas than previously known (solving a question of Nisan and Wigderson [17]), and we presented a super-polynomial separation between monotone and non-monotone formulas in the non-commutative world (answering a question of Nisan [16]).

Complexity in algebraic extensions. The algebraic computations discussed so far assumed that the underlying structure is a field \mathbb{F} . It is natural to ask whether circuits with access to some extension of \mathbb{F} (a richer algebraic structure) are stronger than circuits with access only

¹A polynomial is *homogeneous* if all the monomials that occur in it have the same total degree, and a circuit is *homogeneous* if every step of the computation in it yields a homogeneous polynomial.

to \mathbb{F} . In a joint work with Hrubeš [11] we explored this question. Roughly, we showed that non-commutative extensions can give a lot of power, whereas commutative extensions do not: On one hand, for every polynomial f over a field \mathbb{F} , there is a non-commutative extension N of \mathbb{F} so that f is easy over N , that is, there is a polynomial size formula for f over N . On the other hand, for every field \mathbb{F} , there exist polynomials with zero-one coefficients that are hard over any commutative extension C of \mathbb{F} . We also showed that computations over low dimensional extensions of \mathbb{F} can be efficiently simulated by computations over \mathbb{F} . As a corollary, we obtained that division can be efficiently simulated by sums and products over any field, this was known before only for large enough fields [29].

1.2 Probability theory

For the sake of brevity, only two selected projects related to probability theory are presented.

Loop-erased random walks. Random walks are useful in physics, computer science, economics and more. A well-known example is a simple random walk on the integer lattice \mathbb{Z} , where at each step the walk moves by either $+1$ or -1 with probability one half. One of the fundamental questions about a random walk is what is its scaling limit; roughly, how does it behave after a long time with the appropriate scaling. The scaling limit of a simple random walk on \mathbb{Z} , for example, is Brownian motion.

A *loop-erased random walk* or LERW on a graph G is obtained by performing a random walk on G and then erasing the loops in chronological order. The resulting path is a self-avoiding path in the graph G . It was suggested by Lawler in [13], and has since been studied extensively on the graphs \mathbb{Z}^d [14, 12]. In order to study the case $d = 2$, Schramm [25] introduced a one-parameter family of random curves known as the Schramm-Loewner Evolution or SLE_κ . In [15] Lawler, Schramm and Werner proved that the scaling limit of LERW on \mathbb{Z}^2 is SLE_2 . Their result also holds for other two-dimensional lattices. Many other random processes have been shown to converge to SLE_κ for other values of κ .

In a joint work with Yadin [32], we focused on the scaling limit of LERW on planar graphs², not necessarily lattices. Our main result is a generalization of [15]. We showed that if the scaling limit of a random walk on a planar graph G is Brownian motion, then the scaling limit of LERW on G is SLE_2 . The main contribution of this work is a lemma that states that for planar graphs, certain discrete quantities converge to their continuous analogs.

One interesting example of such a graph is the infinite component³ of super-critical percolation on \mathbb{Z}^2 . Berger and Biskup proved that almost surely the scaling limit of a random walk on this infinite component is Brownian motion [5]. Together with our result, this implies that almost surely the scaling limit of LERW on the super-critical percolation cluster is SLE_2 .

Probability estimates for k -wise independence. A k -wise independent distribution on n bits is a joint distribution of the bits so that each k of them are independent. Such distributions

²Here a planar graph is a graph embedded into the complex plane so that edges do not cross each other.

³For bond percolation on \mathbb{Z}^2 (each bond open with probability $p > 1/2$, all bonds independent), there almost surely exists a unique infinite connected component.

are extremely useful in derandomization of algorithms. They were also shown to be pseudo-random for constant depth Boolean circuits with k poly-logarithmic [6].

Together with Peled and Yadin [18], I considered k -wise independent distributions with identical marginals, each bit has probability p to be 1. We addressed the following question “how high can the probability that all the bits are 1 be, for such distributions?” Such probability estimates are useful, for example, when we wish to estimate the probability of the intersection of the events E_1, \dots, E_n , and we know that the probability of the event $\bigcap_{i \in S} E_i$ is $p^{|S|}$, for every set $S \subset \{1, \dots, n\}$ of size k . This question can also be seen as a relaxation of a major open problem in coding theory⁴, namely, “how large can a linear error correcting code with given parameters be?”

For a wide range of the parameters n, k and p , we found an explicit lower bound for this probability, that matches an upper bound given by Benjamini et al. [4], up to multiplicative factors of lower order. This question could be seen as a discrete moment problem, and our approach was to show that bounds from the theory of the classical moment problem provide good approximations for it.

2 Research directions

Algebraic complexity: restricted theories. In [30] Valiant defined an algebraic analog for the P vs. NP problem, which is now known as the VP vs. VNP problem. Valiant not only defined this algebraic problem, but he also described the theory behind it, for example, he showed that the permanent is VNP-complete. The VP vs. VNP problem is the most important open problem in algebraic complexity.

In an ongoing project with Hrubeš and Wigderson we study restricted versions of the VP vs. VNP problem. The restrictions we study are of algebraic nature: we assume that the computation is done in a non-commutative or a non-associative world. Both of these restrictions were studied before, but a complete theory of the type Valiant defined is missing. In [9] we define the non-commutative version of the VP vs. VNP problem, and describe the theory behind it. We show that this problem is related to well-studied notions in mathematics, such as the Radon-Hurwitz numbers and Clifford algebras. The Radon-Hurwitz numbers, for example, were defined in the beginning of the twentieth century, and are known to be related to matrix theory. In particular, we show that an $n^{1+\epsilon}$ lower bound on a generalization of the Radon-Hurwitz numbers implies that $\text{VP} \neq \text{VNP}$ in the non-commutative world (proving a linear lower bound on this generalization of the Radon-Hurwitz numbers is trivial). We also prove $n^{1+\epsilon}$ lower bounds on some restricted versions of these numbers. These lower bounds may be of independent interest.

The questions that rise from studying the non-commutative world seem to require new ideas and approaches. I believe that such ideas will use tools from other areas of mathematics, such as geometry and invariant theory, and that an understanding of these tools will be a real step towards a better understanding of more general computational models. I plan to pursue this path.

⁴Any linear error correcting code with minimal distance $k + 1$ yields a k -wise independent distribution.

Non-black-box learning. Learning theory is a mathematical study of a concept we all know well, learning. Valiant introduced the first formal definition of learning in [31], where he defined *probably approximately correct* (PAC) learning. In PAC learning, the learner wishes to learn a Boolean function f , by seeing labelled examples of the form $(e, f(e))$, where $e = (e_1, \dots, e_n)$ is a Boolean vector drawn according to some distribution. Learning theory has since been a highly active research area, with many applications and links to other areas.

In an ongoing project with Dvir, Hrubeš and Rao we address the issue of *black-box* access in learning theory. In most learning models, the access of the learner to the function f is a black-box access, namely, the learner sees only the *value* of f on the input e . We propose to allow the learner more information, in a way that suits a computational world. Assume that f is actually computed by a computational device D . Instead of black-box access, the learner will have access to some sub-computations of D .

We are studying applications of this concept to learning theory and other areas, such as proof complexity and cryptography. In learning theory, we study, for example, the problem of learning CNF-formulas using this type of access. Learning CNF-formulas in polynomial time with black-box access is a major open problem in learning theory. In proof complexity, we study the following question: Assume that we wish to refute a CNF-formula F . That is, we wish to use, say, the resolution proof system⁵ to prove that F is unsatisfiable. If F has a refutation R , then for every partial assignment ρ , the formula F_ρ (which is F after applying the assignment ρ) has a refutation R_ρ . Can we find R by seeing a small number of refutations R_ρ ? Intuitively, this question asks whether we can combine the proofs of some special cases to the proof of the whole statement.

3 The future

I find the evolution of mathematical and physical structures fascinating, and wish to continue studying this process. Here are some questions/concepts that guide my research.

What properties of a structure make it complex? Finding explicit properties that are ‘evidence of complexity’ is a key ingredient of any study in complexity theory. An example of such a property is the discrepancy of a matrix in communication complexity. Another example from my work is correlation with product polynomials, which is a generalization of discrepancy, in multilinear formula complexity [23]. I plan to try and understand such known properties better, and search for new ones.

Simplifying ideas. The algorithmic way of thinking, that is common in the computational world, often simplifies our comprehension of complex ideas. One reason is that algorithms are naturally broken into simpler sub-tasks. The combination of these sub-tasks is usually reasonable as well. Here is an example from my research: There were several lower bound proofs for various classes of arithmetic circuits, and I found a simple high-level approach that unifies them [33]. Indeed, this approach consists of a few simple sub-tasks, that can be easily combined into a proof. This approach can be applied to other classes of circuits; our study [9] of non-commutative circuits follows this approach.

⁵Resolution uses simple logical rules, such as $(x \vee a) \wedge (\bar{x} \vee b) \rightarrow (a \vee b)$ where \bar{x} is the negation of x , to derive a contradiction from an unsatisfiable CNF.

Finding more connections between different areas of research. Many useful ideas emerge from the interaction between different concepts. One example from my research is the connection between non-commutative complexity and the Radon-Hurwitz numbers [9]: The Radon-Hurwitz numbers were studied using tools from algebraic topology. We thus have a connection between non-commutative complexity and well-studied deep mathematical theories. I am interested in exploring new connections between complexity theory and other areas of mathematics, such as probability theory and geometry.

References

- [1] S. Aaronson. Multilinear Formulas and Skepticism of Quantum Computing. In Proceedings of STOC 2004: 118–127.
- [2] M. Agrawal and V. Vinay. Arithmetic Circuits: a Chasm at Depth Four. In Proceeding of FOCS 2008: 67–75.
- [3] W. Baur and V. Strassen. The Complexity of Partial Derivatives. Theoretical Computer Science, 22, 1983: 317–330.
- [4] I. Benjamini, O. Gurel-Gurevich and R. Peled. Independence Sensitivity of Boolean Functions. In preparation.
- [5] N. Berger and M. Biskup. Quenched invariance principle for simple random walk on percolation clusters. Prob. Theory Rel. Fields 137, 2007: 83–120.
- [6] M. Braverman. Poly-logarithmic independence fools AC0 circuit. SIAM journal on Computing, to appear.
- [7] D. Grigoriev and A. A. Razborov. Exponential Lower Bounds for Depth 3 Arithmetic Circuits in Algebras of Functions over Finite Fields. Appl. Algebra Eng. Commun. Comput. 10(6), 2000: 465–487.
- [8] D. Grigoriev and M. Karpinski. An Exponential Lower Bound for Depth 3 Arithmetic Circuits. In the proceeding of STOC 1998: 577–582.
- [9] P. Hrubeš, A. Wigderson and A. Yehudayoff. On noncommutative VP vs. noncommutative VNP. In preparation.
- [10] P. Hrubeš and A. Yehudayoff. Homogeneous Formulas and Symmetric Polynomials. arXiv:0907.2621.
- [11] P. Hrubeš and A. Yehudayoff. Arithmetic Complexity in Algebraic Extensions. Manuscript.
- [12] G. Kozma. The scaling limit of loop-erased random walk in three dimensions. Acta Math. 199, 2007: 29–152.
- [13] G. Lawler. A self avoiding walk. Duke Math. J., 1980: 655–694.

- [14] G. Lawler. Loop-erased random walk. *Perplexing Probability Problems: Papers in Honor of Harry Kesten, M. Bramson and R. Durrett*, Eds. Birkhäuser, Boston, MA, 1999: 197–217.
- [15] G. F. Lawler, O. Schramm. and Wendelin Werner. One arm exponent for critical 2D percolation. *Electr. J. Probab.* 7, 2002: paper no. 2.
- [16] N. Nisan. Lower bounds for non-commutative computation. *Proceeding of the 23th STOC*, 1991: 410-418.
- [17] N. Nisan and A. Wigderson. Lower Bound on Arithmetic Circuits via Partial Derivatives. *Computational Complexity* 6, 1996: 217–234.
- [18] R. Peled, A. Yadin and A. Yehudayoff. The Maximal Probability that k -wise Independent Bits are All 1. [arXiv:0801.0059](https://arxiv.org/abs/0801.0059).
- [19] R. Raz. Multi-Linear Formulas for Permanent and Determinant are of Super-Polynomial Size. In the proceeding of *STOC 2004*: 633–641.
- [20] R. Raz. Separation of Multilinear Circuit and Formula Size. *Theory of Computing* 2, 2006: article 6.
- [21] R. Raz. Elusive Functions and Lower Bounds for Arithmetic Circuits. In the proceeding of *STOC 2008*: 711–720.
- [22] R. Raz, A. Shpilka and A. Yehudayoff. A Lower Bound for the Size of Syntactically Multilinear Arithmetic Circuits. *Proceedings of FOCS 2007*: 438 – 448.
- [23] R. Raz and A. Yehudayoff. Multilinear Formulas, Maximal-Partition Discrepancy and Mixed-Sources Extractors. In the proceeding of *FOCS 2008*.
- [24] R. Raz and A. Yehudayoff. Lower Bounds and Separations for Constant Depth Multilinear Formulas. In the proceedings of *Computational Complexity 2008*: 128–139.
- [25] O. Schramm. Scaling limits of loop-erased random walks and uniform spanning trees. *Israel J. Math.* 118, 2000: 221–288.
- [26] A. Shpilka and A. Wigderson. Depth-3 Arithmetic Formulae over Fields of Characteristic Zero. *Computational Complexity* 10(1), 2001: 1–27.
- [27] V. Shoup and R. Smolensky. Lower Bounds for Polynomial Evaluation and Interpolation Problems. In the proceedings of *FOCS 1991*: 378–383.
- [28] V. Strassen Gaussian elimination is not optimal. *Numer. Math.* 13, 1969: 354–356.
- [29] V. Strassen. Vermeidung von Divisionen. *J. Reine Angew. Math.* 264, 1973: 182–202.
- [30] L. G. Valiant. Completeness Classes in Algebra. In the proceedings of *STOC 1979*: 249–261.
- [31] L. G. Valiant. A Theory of the Learnable. *C. ACM* 27(11), 1984: 1134–1142.
- [32] A. Yadin and A. Yehudayoff. Loop-erased random walk and Poisson kernel on planar graphs. Submitted.

- [33] A. Yehudayoff. On the structure of arithmetic circuits, with applications. Barriers in Computational Complexity workshop, a talk.