Mathematisches Forschungsinstitut Oberwolfach

Report No. 51/2018

DOI: 10.4171/OWR/2018/51

Complexity Theory

Organised by
Peter Bürgisser, Berlin
Irit Dinur, Rehovot
Oded Goldreich, Rehovot
Salil Vadhan, Cambridge MA

11 November – 17 November 2018

Abstract. Computational Complexity Theory is the mathematical study of the intrinsic power and limitations of computational resources like time, space, or randomness. The current workshop focused on recent developments in various sub-areas including arithmetic complexity, Boolean complexity, communication complexity, cryptography, probabilistic proof systems, pseudorandomness, and quantum computation. Many of the developments are related to diverse mathematical fields such as algebraic geometry, combinatorial number theory, probability theory, representation theory, and the theory of error-correcting codes.

Mathematics Subject Classification (2010): 68Q01 (Theory of Computing / General), 68Q17 (Computational Difficulty of Problems), 68Q15 (Complexity Classes).

Introduction by the Organisers

The workshop *Complexity Theory* was organized by Peter Bürgisser (TU Berlin), Irit Dinur (Weizmann Institute), Oded Goldreich (Weizmann Institute), and Salil Vadhan (Harvard). The workshop was held on November 11th–17th 2018, and attended by approximately fifty participants spanning a wide range of interests within the field of Computational Complexity. The plenary program, attended by all participants, featured fifteen long lectures and nine short (5-minute) reports by students and postdocs. In addition, intensive interaction took place in smaller groups.

The Oberwolfach Meeting on Complexity Theory is marked by a long tradition and a continuous transformation. Originally starting with a focus on algebraic and

Boolean complexity, the meeting has continuously evolved to cover a wide variety of areas, most of which were not even in existence at the time of the first meeting (in 1972). While inviting many of the most prominent researchers in the field, the organizers try to identify and invite a fair number of promising young researchers.

Computational complexity (a.k.a. complexity theory) is a central field of computer science with a remarkable list of celebrated achievements as well as a vibrant research activity. The field is concerned with the study of the *intrinsic complexity* of computational tasks, and this study tends to *aim at generality*: it focuses on natural computational resources, and considers the effect of limiting these resources on the class of problems that can be solved. Computational complexity is related to and has substantial interaction with other areas of mathematics such as algebra, analysis, combinatorics, geometry, number theory, optimization, probability theory, and representation theory.

The workshop focused on several sub-areas of complexity theory and its nature may be best illustrated by a brief survey of some of the meeting's highlights.

Boolean Circuit Lower Bounds. Ryan Williams presented his joint work with Cody Murray that obtains *circuit lower bounds for Nondeterministic Quasi-Polynomial-time*. This works follows a framework presented by Ryan in 2010, translating algorithmic improvements over exhaustive search into circuit lower bound for functions computable in $\mathcal{NEXP}$ (or smaller amount of NTIME). The framework has two version, one relating to the problem of deciding *circuit satisfiability* for circuits of a target class $\mathcal{C}$, and the other relating to the problem of *distinguishing* circuits (in $\mathcal{C}$) that are satisfied by at least half of the possible inputs and circuits that are unsatisfiable, a problem which is easily solved in probabilistic polynomial-time.

The framework was originally applied to algorithms that run in time $2^n/n^c$, for some sufficiently large constant c, where n denotes the length of the input to the circuit, and has yielded size lower bounds for functions in $\mathcal{NEXP}$. The current work employs this framework to algorithms that present a more significant improvement in the running time; specifically, to algorithms running in time $2^{n-n^{c'}}$, for some constant $c' > 0$. In doing so this work obtains size lower bounds for functions in quasi-polynomial NTIME.

A point worthy highlighting is that algorithms solving the satisfiability problem in time $2^{(1-\Omega(1))\cdot n}$ are unlikely to exist even for the class of CNFs of bounded clause length, provided that the strong ETH holds (i.e., solving the satisfiability problem for $O(1)$-CNFs requires time $2^{(1-o(1))\cdot n}$). On the other hand, such (deterministic) algorithms (i.e., solving the distinguishing problem in time $2^{(1-\Omega(1))n}$) are widely believed to exist for the derandomization problem (e.g., $\mathcal{BPP} = \mathcal{P}$ (in promise versions) yields much stronger results).

The new results are based on an analogous scaled-down versions of the *easy witness lemma*, which seem harder to prove. This lemma refer to a generic verifier of proofs, denoted V, where these proofs are called witnesses. It asserts that if the set S of valid assertions accepted by V (along with an adequate proof) can be recognized by circuits of small size, then each input in S has proofs (or

witnesses) accepted by V that can be described by circuits of relatively small size. The original lemma referred to sets in $\mathcal{NEXP}$ and poly-size circuits, whereas the current lemma refers to a smaller gap between the NTIME bound and the circuit size bound. For example, it can refer to any super-polynomial NTIME bound and the circuit class $\mathcal{P}/\text{poly}$. Alternatively, it refers to $\mathcal{NP}$ and translates a bound of n^c on the size of circuits for the decision problem to a bound of $n^{O(c^3)}$ on the size of circuits describing witnesses.

Probabilistic Checkable Proofs carried to an extreme. Dor Minzer provided an overview of a sequence of works culminating in demonstrating the expressive power of so called 2-to-2 games, which are closely related to 1-to-1 games (a.k.a unique games). In contrast, standard Probabilistic Checkable Proofs with soundness error $\epsilon > 0$ correspond to M-by-M games with $M = \text{poly}(1/\epsilon)$.

The *Unique Game Conjecture* refers to distinguishing instances of a two-variable *Constraint Satisfaction Problem* (CSP) for which almost all constraints can be simultaneously satisfied from instances for which only few constraints can be simultaneously satisfied. Specifically, the constraints refer to two variables assigned values in a fixed alphabet, and the uniqueness condition postulates that for each value assigned to one variable in a constraint there exists at most one satisfying value for the other variable. The conjecture asserts that, for any $\epsilon > 0$, it is hard to *distinguish instances in which at least a* $1 - \epsilon$ *fraction of constraints can be simultaneously satisfied from instances in which at most an* ϵ *fraction of constraints can be simultaneously satisfied*.

The Unique Game Conjecture (UGC) yields optimal inapproximation thresholds for many natural problems (of the CSPs type), and the most acute challenge to it has arose from sub-exponential algorithms that solve this problem; specifically, their running time is 2^{n^β}, where $\beta = \beta(\epsilon)$ is a constant that depends on ϵ such that $\beta(\epsilon)$ vanishes with $\epsilon > 0$. The point is that an NP-hard problem is unlikely to have such algorithms (certainly, if the ETH holds), unless the reduction has a polynomial blow-up in which the degree of polynomial is $1/\beta(\epsilon)$.

Interestingly, the foregoing algorithms apply also for the problem of distinguishing instances in which at least half of constraints can be simultaneously satisfied from instances in which at most an ϵ fraction of constraints can be simultaneously satisfied, where half can be replace by any fixed constant (that is independent of ϵ). This is interesting since the current work shows that the latter problem is NP-hard, while indeed using a reduction with a polynomial blow-up where the degree of the polynomial grows with $1/\epsilon$. Hence, it seems that the doubts regarding UGC cast by these algorithms are eliminated.

The foregoing NP-hardness result follows by establishing the NP-hardness of an analogous 2-by-2 game in which the instances are two-variables constraints such that for each value assigned to one variable there exist at most two satisfying values for the other variable. The main result is that, for any positive ϵ, it is NP-hard to *distinguish 2-by-2 game instances in which at least a* $1 - \epsilon$ *fraction of constraints can be simultaneously satisfied from instances in which at most an* ϵ *fraction of constraints can be simultaneously satisfied*. Furthermore, assuming

ETH, the time complexity of this problem, which is upper bounded by $2^{n^{\beta(\epsilon)}}$, is lower bounded by $2^{n^{\alpha(\epsilon)}}$, where $\alpha(\epsilon) \in (0,1)$ for every $\epsilon > 0$.

High Dimensional Expanders. Tali Kaufman provided a brief introduction to High Dimensional Expanders (HDE), which are $O(1)$-dimensional simplicial complexes that possess some expansion properties, and their potential application in complexity theory. The local definition of HDE refers to the expansion of the "links" of the various lower dimensional "faces" (e.g., vertices), where the link of a face is the collection of all subsets that are disjoint of the face and form a higher dimensional face when added to the said face. (Indeed, the link of a vertex in a graph is the set of all its neighbors, and in a random simplicial complex (almost all) the links consist of isolated vertices.)

Amazingly, HDE exist and can be explicitly constructed; in fact, the current proofs of their existence are inherently constructive. An indication to their usefulness to complexity theory is provided by the fact that they can be used to construct an agreement test of parameters not achieved before. The global-vs-local behavior that they exhibit raises hope that they can be used towards constructing better locally testable codes and proofs (i.e., PCPs).

On the foundation of program obfuscation. While a strong notion of program obfuscation (knows as virtual black-box) is known to be unimplementable, a weak notion that only requires that obfuscated versions of functionally-identical programs be indistinguishable is not ruled out and (in contrast to initial beliefs) found many applications in cryptography. In particular, such *indistinguishability obfuscators* (iO) exist (in an uninteresting sense) if $\mathcal{P} = \mathcal{NP}$. Of course, believing that $\mathcal{P} \neq \mathcal{NP}$, the latter assertion only indicates that the impossibility of iO is hard to establish. But can we show that iO exists in a world in which $\mathcal{P} \neq \mathcal{NP}$, let alone assuming various reasonable complexity assumptions?

Rachel Lin surveyed work aimed at this direction. In particular, these works rely on the difficulty of the so called DDH problem w.r.t "multilinear maps". Specifically, this works attempt to base iO on various incarnations of the "multilinear assumption", while seeking to minimize the "level" of "linearity" with the hope of reaching the level two. Currently the basing of iO on level three seems most promising, while first attempts to base iO on level two have failed (in ways that leave the door open to more sophisticated attempts).

Doubly-efficient interactive proof systems. The invention of interactive proof systems and the exploration of their power are among the greatest success stories of computational complexity. While research in the 1980s referred to polynomial-time verification aided by a computationally unbounded prover, the term *doubly-efficient* refers to almost linear-time verification aided by a polynomial-time prover. Clearly, only polynomial-time solvable problems can have such a proof system. Furthermore, such problems have almost linear space complexity.

Ron Rothblum presented a joint work with Omer Reingold and Guy Rothblum that provides a (constant round) doubly-efficient interactive proof systems for any

set that can be decided by a machine that runs in polynomial-time and uses small space (e.g., space $n^{0.499}$, where n is the length of the input). The gap towards the aforementioned upper bound is both quantitative (i.e., $n^{0.499}$ versus $n^{1.001}$) and qualitative (i.e., whether the same algorithm or two different algorithms should satisfy the two complexity bounds).

On Classical verification of quantum computations. Thomas Vidick presented a work by Urmila Mahadev that refers to an incarnation of the doubly-efficient paradigm. Specifically, it refers to proof systems in which an efficient *quantum-computation* prover can convince an efficient *classical-computation* verifier of the correctness of an assertion that can be decided by an efficient quantum algorithm such that the soundness holds (only) computationally (i.e., w.r.t efficient quantum cheaters). The main result is that, under a rather standard assumption (i.e., quantum-hardness of LWE with exponential modulus and subexponential noise ratio), every set in QBP has a (four-round) proof system of the foregoing type.

The fine-grained complexity of approximation. A relatively recent direction of research in complexity theory refers to the study of problems that are known to have polynomial-time algorithms, where the aim is to provide evidence that the known algorithms are actually the best possible. Aviad Rubinstein presented the first result that refers to the fine-grained complexity of *approximation versions* of optimization problems that were previously studied in the fine-grained context. The focus of his presentation was on closest pair problems, which can be solved exactly in quadratic time. Assuming the strong ETH, it was shown that these problems cannot be approximated to within a factor of $\exp((\log n)^{1-o(1)})$ in time $n^{2-\epsilon}$, for any constant $\epsilon > 0$.

Other plenary presentations. In addition to the aforementioned presentations, the plenary session featured a few additional talks, some providing high-level overview to novel research directions and some reporting of a single result. These included:

- *Recent developments related to RL versus L.* Omer Reingold briefly outlined three directions of research that are aimed at obtaining better derandomization of randomized computation of (logarithmically) bounded space.
- *Invariant theory and complexity theory.* Avi Wigderson presented a brief introduction to invariant theory, highlighting the fact that many central questions in complexity theory can be formulated or rather cast as questions that refer to the most basic concepts of invariant theory.
- *Matrix Multiplication and the Tensor/Waring Rank.* Joseph Landsberg presented the discovery that the exponent of matrix multiplication is also determined by the Waring rank of the matrix multiplication polynomial, and its potential to lead to improvements in the upper bound on that exponent.

- *Pseudodeterministic Algorithms and Proofs.* The presentation of Shafi Goldwasser focused on the notion of pseudodeterministic algorithms, which are randomized algorithms for search problems that (whp) return the same "canonical" solution, whenever the instance does have a solution.
- *Interlacing Polynomials, Free Probability, and Random Matrices.* Nikhil Srivastava presentation focused on the connections between the "interlacing polynomials" method, free probability, and random matrix theory.
- *Near-optimal constructions of epsilon-biased sets.* Amnon Ta-Shma presented a construction of a sample space of size $O(n/\epsilon^{2+o(1)})$ of ϵ-biased sequences of length n, improving over prior bounds of $O(n^2/\epsilon^2)$ and $O(n/\epsilon^3)$.
- *Algebraic CSP dichotomy theorem.* Venkat Guruswami outlined the proof of this theorem, which is pivoted at the notion of a polymorphism (i.e., a local mapping of several valid solutions to a CSP (wrt a fixed predicate) to a valid assignment).

In addition, the following graduate students and post-doctoral fellows presented brief reports of their research agenda: Ankit Garg, Mika Göös, Rohit Gurjar, Pravesh Kothari, Inbal Livni, Cody Murray, Tselil Schramm, Roei Tell, and Jeroen Zuiddam.

Informal specialized sessions. Apart from the formal plenary program, intense interaction between the participants took place in smaller groups. Some of these took place in the form of specialized sessions, featuring the following presentations.

- *Circuit lower bounds for Nondeterministic Quasi-Polynomial-time.* As a follow-up on the overview provided by Ryan Williams in the plenary session, Ryan and Cody Murray provided more details in a specialized session.
- *On the hardness of 2-to-2 games.* As a follow-up on the overview provided by Dor Minzer in the plenary session, Dor and Muli Safra provided more details in a specialized session.
- *On Classical verification of quantum computations.* As a follow-up on the overview provided by him in the plenary session, Thomas Vidick provided more details on Mahadev's scheme, focusing on the security requirement for the underlying cryptographic primitive (a trapdoor claw-free function pair), and how this security requirement translates into soundness of the verification procedure.
- *Pseudorandom Generators for Obfuscation.* As a follow-up on Rachel Lin's plenary talk, Zvika Brakerski, Pravesh Kothari, and Rachel reviewed the different suggestions of using low-locality and low-degree pseudorandom generators in order to construct iO from multilinear maps with low level of multilinearity, and in particular from bilinear maps. They discussed the required properties for such generators, went over prior suggestions that were attacked, and finally discussed the most recent family of candidates and their properties.
- *Recent developments related to RL versus L.* Following-up on his plenary presentation, Omer Reingold discussed a notion of a "mild pseudorandom

restriction generator" and an approach pursued by a sequence of works that applies such generators to fool read-once constant-width branching programs with close to logarithmic seed. Raghu Meka discussed a recent result that provided a generator that fools read-once width three branching programs with close to logarithmic seed. David Zuckerman discussed a simple construction of a hitting-set generator for polynomial-width read-once branching programs in which the dependence of the seed-length on the error parameter is close to optimal. Lastly, Salil Vadhan discussed two recent works that give small memory algorithms (close to logarithmic) for evaluating various parameters of random walks in undirected graphs (e.g., hitting time).

- *Matrix Multiplication Upper Bounds.* Following-up on Joseph Landsberg's plenary presentation, Chris Umans reported on several paths to proving the exponent of matrix multiplication is two using group theoretic methods and their generalizations, leading to a few seemingly unrelated open questions that may each imply $\omega = 2$ (if resolved affirmatively). Jeroen Zuiddam gave an introduction to Strassen's theory of asymptotic spectra and briefly discussed the new construction of elements in the asymptotic spectrum of complex tensors via moment polytopes. Virginia Vassilevska-Williams reported on a recent work extending the limitations results of prior work to a much broader class of tensors.
- *Operator scaling and moment polytopes.* Following-up on Avi Wigderson's plenary presentation, Ankit Garg and Avi presented new algorithms developed for polyhedral optimization, which constitutes one of the most exciting developments in the interaction of Complexity and Optimization (on the CS side) and Invariant Theory (on the other side). This follows the "emergence of polytopes from group actions", Specifically, to every linear group action on a linear space (and indeed in much more general situations) one can associate a so-called "moment polytope". The new algorithms (both alternate minimization and geodesically convex) for the general "null cone problem" in invariant theory turn out to provide new, efficient separation oracles for this wide class of polytopes.
- *Computationally Sound Proof Systems.* This session featured three loosely related presentations. The first presentation, given by Yael Kalai, was about constructing highly efficient non-interactive computationally-sound proof systems in the common reference string (CRS) model, while allowing for public verifiability (i.e., anyone can verify the proof, and no secret information about the CRS is needed for verification). The second presentation, given by Guy Rothblum, showed that solving PPAD complete problems is no easier than breaking the soundness of the Fiat-Shamir transformation when it is applied to the sum-check protocol. The third presentation, given by Ron Rothblum, was about a recent sequence of papers that construct explicit hash functions that can be used to realize the Fiat-Shamir transformation, for any (statistically sound) interactive proof.

- *The Real Tau Conjecture and Variants.* Pascal Koiran surveyed various variants of the Real Tau conjecture, showing that they all imply Valiant's conjecture (which is an arithmetic version of the widely believed $\mathcal{P} \neq \mathcal{NP}$). The original variant claims that the number of real zeros of a univariate real polynomial given by a depth four circuit (sum of products of sparse polynomials) is polynomially upper-bounded in the size of circuit. Peter Bürgisser outlined a recent result saying that the Real Tau Conjecture is true on average for any depth four circuit with independent standard Gaussian coefficients.
- *A deterministic PTAS for the Algebraic Rank of Bounded Degree Polynomials.* Markus Bläser presented a deterministic polynomial-time approximation scheme (PTAS) for computing the algebraic rank of a set of bounded degree polynomials. More specifically, on input a set, $\mathbf{f}$, of n m-variate polynomials of degree at most d over $\mathbb{F}$, and a rational number $\epsilon > 0$, the algorithm runs in time $O((nmd)^{O(d^2/\epsilon)})$, and outputs a number in $[(1-\epsilon)r, r]$, where r is the algebraic rank of $\mathbf{f}$.
- *The status of geometric complexity theory.* Christian Ikenmeyer led a discussion of the recent developments in geometric complexity theory (GCT) and how they relate to each other, focusing on representation theoretic multiplicities in coordinate rings of orbits and orbit closures. The GCT approach is highly adjustable to a lot of situations and its flagship permanent vs "border determinant" conjecture can be equivalently phrased in many ways that each feature significantly different geometry and representation theory. The nonexistence of different types of occurrence obstructions in several of these computational models was discussed. Although the proofs of these "no-go" results about separating complexity classes using occurrences of irreducible representations are fairly robust, there seems to be no such no-go result yet about separating complexity classes using representation theoretic *multiplicities*.
- *Quantum Fully Homomorphic Encryption.* Quantum FHE is a scheme that allows a quantum server to apply a quantum function on encrypted data in a blindfolded manner, without learning anything about the contents of the encryption. The session, led by Zvika Brakerski, included a review of recent constructions.
- *Tensor decomposition and Sum of Squares.* Tensor decomposition is hard, even to approximate, in the worst case. Nonetheless, recently, there have been algorithmic developments for the tensor decomposition problem in certain special cases (which are relevant for machine learning and data science applications). Specifically, for tensors with certain low-rank structures, the sum-of-squares algorithm gives polynomial time algorithms. These developments were surveyed by Tselil Schramm, who provided a high-level outline of the algorithms and the proofs.
- *Lifting theorems for communication complexity and beyond.* Lifting theorems give a powerful general methodology for proving lower bounds in

various models of communication complexity by "lifting" the hardness of some Boolean function f in various models of computation (e.g., decision-trees) to lower bounds on the communication complexity of some computation associated with f (via some appropriately chosen "gadget" Two sessions focused on the recent resurgence of lifting theorems and their applications in areas beyond communication complexity. In the first session, Mika Göös, provided a brief survey on recent progress in lifting. The second session featured presentations of Linear Programming lower bounds for Max-CSPs via Lifting Theorems by Pravesh Kothari, a Lifting Theorem for Non-Negative Rank by Raghu Meka, and a a Lifting Theorem for Monotone Circuit Complexity by Mika Göös.

- *Lossless dimension expanders.* A dimension expander for a vector space $\mathbb{F}^n$ is a collection of d linear maps such that for every low-dimensional subspace U of $\mathbb{F}^n$, its image under all the maps has dimension at least $a \cdot \dim(U)$, where a is the "expansion factor". Dimension expanders are the linear-algebraic analog of vertex expanders. Over finite fields, w.h.p., a random collection of $d = O(1)$ linear maps offers excellent *lossless* expansion (with expansion factor almost equal to d). But explicit constructions (for growing n and fixed d) with even modest expansion factors are non-trivial to obtain. Venkat Guruswami described a recent work on explicit lossless dimension expanders over large fields (previously even constructions with expansion proportional to the degree were not known).
- *The KRW Conjecture: Results and Open Problems.* Proving super-logarithmic lower bounds on the depth of circuits is one of the main frontiers of circuit complexity. In the early 1990s, Karchmer, Raz and Wigderson observed that this question can be resolved by proving the following conjecture: For two (non-constant) Boolean functions, the depth complexity of their composition is about the sum of their individual depth complexities. While the conjecture is still open, there has been some exciting progress toward such a proof, some of it in the last few years. Or Meir surveyed the known results and discussed future directions for research on the KRW conjecture.
- *Algorithmic Fairness.* Omer Reingold and Guy Rothblum discussed a complexity-theoretic perspective on Algorithmic fairness. Often, definitions of fairness are based on socially identified groups, requiring that a given statistic be equal across a few demographic groups that are identified as deserving protection. Such broad-stroke statistical guarantees tend to be relatively easy to satisfy, but tend to be weak in the protections they provide. In contrast, recent research provides efficient learning algorithms that ensure protection (according to some fairness notion) to every subpopulation within some rich class of sets, where the classes are defined in complexity-theoretic terms. This research aims at obtaining the strongest fairness guarantees that can be obtained with the available computational resources.

Acknowledgments: The MFO and the workshop organizers would like to thank the National Science Foundation for supporting the participation of junior researchers in the workshop by the grant DMS-1049268, "US Junior Oberwolfach Fellows". Moreover, the MFO and the workshop organizers would like to thank the Simons Foundation for supporting Joseph Landsberg and Luca Trevisan in the "Simons Visiting Professors" program at the MFO.

Workshop: Complexity Theory

Table of Contents

Abstracts

Pseudo Deterministic Algorithms and Proofs

Shafi Goldwasser

Probabilistic algorithms for both decision and search problems can offer significant complexity improvements over deterministic algorithms. One major difference, however, is that they may output different solutions for different choices of randomness. This makes correctness amplification impossible for search algorithms and is less than desirable in setting where uniqueness of output is important such as generation of system wide cryptographic parameters or distributed setting where different sources of randomness are used.

Pseudo-deterministic algorithms are a class of randomized search algorithms, which output a unique answer with high probability. Intuitively, they are indistinguishable from deterministic algorithms by a polynomial time observer of their input/output behavior.

In this talk I described what is known about pseudo-deterministic algorithms in the sequential, sub-linear and parallel setting. For example, there exist pseudo-deterministic algorithms for number theory problems finding a generator for Z_p^* when the factorization of $p-1$ is known and contains a large prime divisor [GG11]; algebraic problems: finding a point x where $p(x) \neq 0$ for multivariate polynomial p over a finite field [GG11]; graph problems finding a perfect matching in a graph (bipartite and recently extended to general graphs) in parallel time(RNC [GG17]).

In the sub-linear search algorithms domain, we can show strict separations in the number of queries necessary for deterministic, randomized and pseudo-deterministic algorithms for some search problems [GGR12] when the problem exhibits high input sensitivity (informally, the set of solutions for x and x' (x with 1 bit flipped) have an empty intersection).

In interesting works of [OS17, OS18] they allow the pseudo-deterministic algorithm to run in sub-exponential time and relax the algorithm to work for infinitely many input lengths. They show how to obtain such relaxed pseudo-deterministic algorithms for "dense" search problems where it is possible to sample solutions efficiently for the underlying search problems – in particular they show as a special case that generating "canonical" primes is possible infinitely often in sub-exponential time. They also show to obtain pseudo-deterministic approximations for integer valued functions if randomized-approximation schemes exist, again infinitely often using sub-exponential time.

Another relaxation of pseudo-deterministic algorithms considered by [GL18] is algorithms which are guaranteed to output logarithmic number of solutions (rather than unique). The show such algorithms for any problem in randomized log space.

Finally, we describe an extension of pseudo-deterministic algorithms to pseudo-deterministic proofs: interactive proofs for search problems where the verifier on input x, is given a solution to some search problem and is guaranteed with high probability to output the same solution on different executions, regardless of

an(untrusted) prover strategies. We show how to do this for the graph isomorphism problem.

References

[GGH18] Shafi Goldwasser, Ofer Grossman, Dhiraj Holden, *Pseudo-Deterministic Proofs*, ITCS 2018: 17:1-17:18.

[GG17] Shafi Goldwasser, Ofer Grossman, *Bipartite Perfect Matching in Pseudo-Deterministic NC*, ICALP 2017: 87:1-87:13.

[GGR12] Oded Goldreich, Shafi Goldwasser, Dana Ron, *On the possibilities and limitations of pseudodeterministic algorithms*, Electronic Colloquium on Computational Complexity (ECCC) 19: 101 (2012).

[GG11] Eran Gat, Shafi Goldwasser, *Probabilistic Search Algorithms with Unique Answers and Their Cryptographic Applications*, Electronic Colloquium on Computational Complexity (ECCC) 18: 136 (2011).

[OS18] Igor Carboni Oliveira, Rahul Santhanam, *Pseudo-Derandomizing Learning and Approximation*, APPROX-RANDOM 2018: 55:1-55:19.

[OS17] Igor Carboni Oliveira, Rahul Santhanam, *Pseudodeterministic constructions in subexponential time*, STOC 2017: 665-677.

[GL18] Ofer Grossman, Yang P. Liu, *Reproducibility and Pseudo-Determinism in Log-Space*, CoRR abs/1803.04025 (2018).

Algebraic CSP dichotomy theorem: A polymorphic gateway between structure and algorithms

Venkatesan Guruswami

One of the major goals of the theory of computation is to classify broad classes of computational problems as easy (e.g., polytime solvable) or intractable (e.g., NP-hard), along with some structural understanding of what governs its complexity. Given the vast landscape of problems and the diverse algorithms to solve them, no single mathematical theory can hope to explain the underpinnings of the easiness/hardness of all problems. There is, however, one broad class of problems called constraint satisfaction problems (CSPs) for which an elegant mathematical theory has managed to establish a complexity dichotomy: every CSP is either polynomial time tractable or NP-complete. This dichotomy was first conjectured in an influential paper by Feder and Vardi [FV98].

A refinement of this conjecture, called the algebraic dichotomy conjecture [BJK05], further pinpoints the distinguishing feature between the easy and the hard cases: the existence of non-trivial operations called *polymorphisms* under which the solution space is closed. For instance, for linear equations, if v_1, v_2, v_3 are three solutions, then so is $v_1 - v_2 + v_3$, and the underlying polymorphism is $f(x, y, z) = x - y + z$. This algebraic formulation enabled tapping into the deep and rich methods of universal algebra to tackle the complexity of CSPs. After a long line of work resolving several special cases (eg., in [Sch78, HN90, Bul06, Bar11, BK14], etc.), the algebraic dichotomy conjecture was established in full generality in two recent independent breakthroughs [Bul17, Zhu17].

The talk was meant to be a broadly accessible survey into some of the backdrop and developments surrounding the algebraic CSP dichotomy conjecture (now theorem), focusing on the interplay between polymorphisms and complexity of CSPs via some illustrative special cases.

To state the algebraic dichotomy theorem precisely, let us first define CSPs and polymorphisms more formally. A CSP over domain D is specified by a finite collection Λ of relations over D (called a *template*), and is denoted as $\mathrm{CSP}(\Lambda)$. An instance of $\mathrm{CSP}(\Lambda)$ consists of a set of variables V and a collection of constraints $\{(\tau, P)\}$ where $P \in \Lambda$ and τ is a tuple of k distinct variables where k is the arity of P (i.e., $P \subseteq D^k$). The goal is find an assignment $\sigma : V \to D$ that satisfies all constraints, i.e., $(\sigma(\tau_1), \ldots, \sigma(\tau_k)) \in P$ for each constraint (τ, P). For example when $D = \{1, 2, 3\}$, and Λ consists of the single arity two relation $\{(a, b) \mid a \neq b\} \subseteq D^2$, we have the problem of telling if a graph is 3-colorable.

Polymorphisms. Polymorphisms are operations that preserve membership in a relation. Formally, $f : D^m \to D$ is a polymorphism of a relation $P \subseteq D^k$, denoted $f \in \mathrm{Pol}(P)$, if for every choice of m k-tuples from P, applying f component-wise to these tuples results in a tuple in P, i.e., for all $(a_1^{(i)}, \ldots, a_k^{(i)}) \in P$, $i = 1, 2, \ldots, m$, we have $\big(f(a_1^{(1)}, a_1^{(2)}, \ldots, a_1^{(m)}), \cdots, f(a_k^{(1)}, a_k^{(2)}, \ldots, a_k^{(m)})\big) \in P$. In other words, for every $m \times k$ matrix whose rows belong to P, applying f column-wise gives a new row also belonging to P.

For a template Λ of relations, one defines $\mathrm{Pol}(\Lambda) = \cap_{P \in \Lambda} \mathrm{Pol}(P)$ to be those operations that preserve all relations in Λ. For any instance of $\mathrm{CSP}(\Lambda)$, its set of satisfying assignments are closed under combinations by polymorphisms of Λ. Every Λ admits some trivial polymorphisms, namely dictator functions f where $f(x_1, \ldots, x_m) = x_i$ for some i (the polymorphism just "copies" the i'th row when applied to m tuples from the relation). Informally, a CSP is tractable if it has a "non-trivial" (roughly something non-dictatorial) polymorphism. The formal statement is below.

Theorem 1 (Algebraic dichotomy theorem)**.** *For a template Λ over domain D, the associated* $\mathrm{CSP}(\Lambda)$ *is polynomial tractable if* $\mathrm{Pol}(\Lambda)$ *includes a Weak-Near-Unanimity (WNU) function f, namely a function of some arity $m \geq 2$ such that $f(a, b, b, \ldots, b) = f(b, a, b, \ldots, b) = \cdots = f(b, b, \ldots, b, a)$ for all $a, b \in D$; otherwise it is NP-complete. (Note that the WNU condition rules out dictator functions. For the Boolean domain, the condition for tractability is equivalent to having a polymorphism that is not a dictator or its complement.)*

The hardness part of this link between (lack of) polymorphisms and computational complexity has been long known; the highly non-trivial algorithmic part took until recently to establish completely. The talk discussed the following results to shed light on the polymorphic principle behind the complexity of CSPs.

- The Galois connection which shows that polymorphisms capture the complexity of CSPs: $\mathrm{Pol}(\Lambda) \subseteq \mathrm{Pol}(\Gamma)$ if and only if $\mathrm{CSP}(\Gamma)$ pp-reduces to $\mathrm{CSP}(\Lambda)$. That is, richer the polymorphisms, easier the problem. Here

pp-reduction captures the simple local gadget reductions that express constraints of Γ as a CSP instance over Λ with auxiliary variables. The survey [Che09] can be referred to for a proof.

- One of the key properties of polymorphisms is that they form a *clone*: they contain all dictators and are closed under arbitrary composition. This gives them a rich structure. There are many equivalent conditions for tractability: the existence of WNU polymorphism is equivalent to the existence of a Taylor polymorphism (a weaker symmetry requirement) and a cyclic polymorphism (a stronger symmetry).
- The use of cyclic polymorphisms to prove the Hell-Nesetril theorem on the complexity dichotomy of graph homomorphism problems [HN90], following the survey [BKW17]. We note that every CSP is equivalent to a digraph homomorphism problem, so a dichotomy for the latter is equivalent to the dichotomy for CSPs.
- Polymorphic statement of Schaefer's classic dichotomy theorem for Boolean CSPs [Sch78], and hints at how one deduces the structured polymorphisms that correspond to the tractable cases (constant functions, AND, OR, 3-bit Majority, and 3-XOR).
- The algorithmic part of the dichotomy theorem was established via a program that gave algorithms under successively weaker assumptions about the polymorphisms, culminating with the WNU case that forms the boundary with the intractable. Toward illustrating how polymorphisms enable efficient algorithms, the talk presented algorithms in the case of symmetric polymorphisms of all large enough arities (using linear programming), 3-bit majority polymorphism (using a local propagation algorithm), and a Mal'stev polymorphism (using an iterative algorithm that maintains a compact respresentation of the solutions as constraints as added, giving an abstract generalization of Gaussian elimination).
- CSPs for which local propagation correctly decides satisfiability are called *bounded width*. An important milestone in the algebraic study of CSPs was the result of Barto and Kozik [BK14] that bounded width CSPs are precisely those that cannot express linear equations. The non-trivial aspect of completing the algorithmic part of CSP dichotomy was the many complicated ways in which CSPs could encode linear equations, and how to combine linear equation solving with local propagation algorithms. This was successfully achieved in the works of Bulatov and Zhuk [Bul17, Zhu17].

The excellent survey [BKW17] is recommended for further understanding of polymorphisms and their relation to CSP complexity. The resolution of the Feder-Vardi conjecture might seem like the end of the road, but the scope of polymorphic inquiries extends to many variants of CSPs, including optimization, counting, and approximation. Fascinatingly, *partial* polymorphisms govern the best exponential runtime for the case of NP-hard CSPs. A recent study has extended the polymorphic framework to the promise version of CSPs (that captures many interesting problems including approximate graph coloring), and has established several cases

where rich enough families of polymorphisms lead to algorithms, and limited polymorphisms leads to hardness [AGH17, BG18, BG19]. The dividing line between these cases is far from clear, and holds many exciting challenges for the future.

References

[AGH17] P. Austrin, V. Guruswami, and J. Håstad. $(2+\epsilon)$-SAT is NP-hard. *SIAM Journal on Computing*, 46(5):1554–1573, 2017.

[Bar11] L. Barto. The dichotomy for conservative constraint satisfaction problems revisited. In *Proceedings of the 26th Annual IEEE Symposium on Logic in Computer Science, LICS 2011, June 21-24, 2011, Toronto, Ontario, Canada*, pages 301–310, 2011.

[BG18] J. Brakensiek and V. Guruswami. Promise constraint satisfaction: Structure theory and a symmetric Boolean dichotomy. In *Proceedings of the 28th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1782–1801, 2018.

[BG19] J. Brakensiek and V. Guruswami. An algorithmic blend of LPs and ring equations for promise CSPs. In *Proceedings of the 30th Annual ACM-SIAM Symposium on Discrete Algorithms*, 2019.

[BJK05] A. A. Bulatov, P. Jeavons, and A. A. Krokhin. Classifying the complexity of constraints using finite algebras. *SIAM J. Comput.*, 34(3):720–742, 2005.

[BK14] L. Barto and M. Kozik. Constraint satisfaction problems solvable by local consistency methods. *J. ACM*, 61(1):3:1–3:19, 2014.

[BKW17] L. Barto, A. A. Krokhin, and R. Willard. Polymorphisms, and how to use them. In *The Constraint Satisfaction Problem: Complexity and Approximability*, volume 7 of *Dagstuhl Follow-Ups*. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, pages 1–44, 2017.

[Bul06] A. A. Bulatov. A dichotomy theorem for constraint satisfaction problems on a 3-element set. *Journal of the ACM*, 53(1):66–120, 2006.

[Bul17] A. Bulatov. A Dichotomy Theorem for Nonuniform CSPs. In *58th IEEE Annual Symposium on Foundations of Computer Science (FOCS)*, pages 319–330, 2017.

[Che09] H. Chen. A rendezvous of logic, complexity, and algebra. *ACM Comput. Surv.*, 42(1):2:1–2:32, Dec. 2009.

[FV98] T. Feder and M. Y. Vardi. The computational structure of monotone monadic SNP and constraint satisfaction: A study through datalog and group theory. *SIAM J. Comput.*, 28(1):57–104, 1998.

[HN90] P. Hell and J. Nesetril. On the complexity of H-coloring. *Journal of Combinatorial Theory, Series B*, 48(1):92 – 110, 1990.

[Sch78] T. J. Schaefer. The complexity of satisfiability problems. In *Proceedings of the Tenth Annual ACM Symposium on Theory of Computing*, STOC '78, pages 216–226, New York, NY, USA, 1978. ACM.

[Zhu17] D. Zhuk. A Proof of the CSP Dichotomy Conjecture. In *58th IEEE Annual Symposium on Foundations of Computer Science (FOCS)*, pages 331–342, 2017.

High Dimensional Expanders

Tali Kaufman

High dimensional expanders are high dimensional analogues of the well studied expander graphs. As opposed to expander graphs that are abundant (e.g. a random bounded degree graph is an expander with high probability), bounded degree high dimensional expanders are rare objects and their only known constructions are explicit and are based on sophisticated mathematics [LSV05, KO18].

Simplicial complexes. The object of study is a simplicial complex. A *d-dimensional simplicial complex* X is a collection of sets $(X(0), X(1), ..., X(d))$, $X(i)$ contains sets of size $i+1$, with a closure property: every set in the collection all its subsets are also there. A graph is a 1-dimensional simplicial complex. The underlying graph of a complex X is $(X(0), X(1))$. In order to define high dimensional expanders we should introduce links of simplicial complexes: For X, a d-dimensional simplicial complex; a *link* of $\sigma \in X(i)$ is a $d-i-1$ complex X_σ obtained by taking all d-cells containing σ and removing σ from them.
What is a High Dimensional Expander? Roughly speaking: A sparse complex whose links are dense or "similar" to the complete complex (A "sparsifier" of the complete complex).

A *spectral definition* of high dimensional expander is as follows. $X = (X(0), X(1), ..., X(d))$ is γ-*local spectral expander* if

- The graph $(X(0), X(1))$ is connected.
- For every $s \in X(i)$: $i < d-1$ the graph $(X_s(0), X_s(1))$ is a γ-expander (i.e., has second largest e.v. bounded by γ).

On global expansion from local expansion. Oppenheim [Opp18] has shown that local expansion of the links of a complex imply global expansion of its underlying graph! A family of (underlying) graphs of high dimensional expanders are expander graphs, whose expansion can be proven by *local* arguments! (compare to other known expanders whose expansion can not relay on local arguments).
High order random walks. Similar to graphs, it is possible to study high order random walks on high dimensional expanders. High order random walk refers to a walk from a k cell to a k cell via a $k+1$ cell that contains them. Expanding links imply fast mixing of high order Random walks! [KM17, DK17, KO18b, DDF+18].
High dimensional expanders imply local to global phenomena. On the complete complex various computational tasks are relatively understood. In high dimensional expanders links are similar to the complete complex. Thus, the ideology that we have demonstrated in this talk is that one can pull the behaviour of the links to the whole sparse complex. This is manifested in the following works:

- Global spectrum of the underlying graph of the complex from local spectrum of its links [Opp18].
- Use high dimensional expanders to construct codes with global list decoding property, by local list decoding of their restriction to links [DHK+19].
- Global agreement expansion (aka 'direct product testing) from local agreement expansion [DK17]
- Global co-systolic expansion (high dimensional "edge expansion") from local co-systolic expansion [KKL14, EK16].

Edge expansion in high dimensions. We have discussed only spectral definition of high dimensional expanders. For graphs: spectral definition and "edge expansion" definition are equivalent. In high dimensions "edge expansion" is called co-systolic, co-boundary expansion, and it has important topological implications. "Edge expansion" in high dimensions involves the notions of homology and cohomology, and it is beyond our scope. Different than graphs: spectral expansion and "edge

expansion" in high dimensions are not known to be equivalent; They are conjectured to be different. Interestingly, "Edge expansion" in high dimensions have natural interpretation as a local testability of some code [KL14].
Some Open Questions.

- Find better/different LTCs using high dimensional expanders.
- Find better/different PCPs using high dimensional expanders.
- More applications of high dimensional expanders.
- Random model of high dimensional expanders?

References

[DDF+18] Yotam Dikstein, Irit Dinur, Yuval Filmus, Prahladh Harsha, *Boolean Function Analysis on High-Dimensional Expanders*, APPROX-RANDOM 2018: 38:1–38:20.

[DHK+19] Irit Dinur, Prahladh Harsha, Tali Kaufman, Inbal Livni Navon, Amnon Ta-Shma, *List Decoding with Double Samplers*, SODA 2019, to appear.

[DK17] Irit Dinur and Tali Kaufman, *High Dimensional Expanders Imply Agreement Expanders*, FOCS 2017, 974–985.

[EK16] Shai Evra and Tali Kaufman, *Bounded degree cosystolic expanders of every dimension*, STOC 2016, 36–48.

[KKL14] Tali Kaufman, David Kazhdan and Alexander Lubotzky, *Ramanujan Complexes and bounded degree topological expanders*, FOCS 2014, 484–493.

[KL14] Tali Kaufman and Alexander Lubotzky, *High Dimensional Expanders and Property Testing*, ITCS 2014, 501–506.

[KM17] Tali Kaufman and David Mass, *High Dimensional Combinatorial Random Walks and Colorful Expansion*, Innovations in Theoretical Computer Science (ITCS), 2017: 4:1–4:27.

[KO18] Tali Kaufman and Izhar Oppenhaim, *New construction of bounded degree high dimensional expanders*, STOC 2018, 773–786.

[KO18b] Tali Kaufman and Izhar Oppenhaim, *High Order random walks: Beyond spectral gap*, APPROX-RANDOM 2018, 47:1–47:17.

[LSV05] Alex Lubotzky, Beth Samuels and Uzi Vishne, *Ramanujan complexes of type $\tilde{A}_d$*, Israel J. Math. Vol 149, 267–299, 2005.

[Opp18] Izhar Oppenheim, *Local Spectral Expansion Approach to High Dimensional Expanders Part I: Descent of Spectral Gaps*, Discrete and Computational Geometry 59(2): 293–330 (2018).

The complexity of matrix multiplication: developments since 2014

J.M. Landsberg

The complexity of all operations in linear algebra is governed by the complexity of matrix multiplication. In 1968 V. Strassen [Str69] discovered the way we usually multiply matrices is not the most efficient one and initiated the central problem of determining the complexity of matrix multiplication. He defined a fundamental constant ω, called the *exponent of matrix multiplication*, that governs its complexity. For a tensor $T \in \mathbb{C}^m \otimes \mathbb{C}^m \otimes \mathbb{C}^m$, let $\mathbf{R}(T)$ denote its tensor rank, the smallest r such that T may be written as a sum of r rank one tensors, and $\underline{\mathbf{R}}(T)$ its tensor border rank, the smallest r such that T may be written as a limit of a sequence

of rank r tensors. Bini [Bin80] proved the border rank of the matrix multiplication tensor $\underline{\mathbf{R}}(M_{\langle \mathbf{n}\rangle})$ asymptotically determines ω. More precisely, considering $\underline{\mathbf{R}}(M_{\langle \mathbf{n}\rangle})$ as a function of $\mathbf{n}$, $\omega = \inf_\tau\{\underline{\mathbf{R}}(M_{\langle \mathbf{n}\rangle}) = O(\mathbf{n}^\tau)\}$.

This talk has two goals: (i) report on progress in the last four years regarding upper and lower bounds for the complexity of matrix multiplication and tensors in general, and (ii) to explain the utility of algebraic geometry and representation theory for matrix multiplication and complexity theory in general.

Lower bounds. Strassen-Lickteig (1983, 1985) [Str83, Lic84] showed $\underline{\mathbf{R}}(M_{\langle \mathbf{n}\rangle}) \geq \frac{3\mathbf{n}^2}{2} + \frac{\mathbf{n}}{2} - 1$. Then, after 25 years without progress, Landsberg-Ottaviani (2013) [LO15] showed $\underline{\mathbf{R}}(M_{\langle \mathbf{n}\rangle}) \geq 2\mathbf{n}^2 - \mathbf{n}$. Around 2014 several authors [EGO+17, DM18, Gal17] independently proved that the existing lower bound methods would not go much further. In [LM17b] the *border substitution method* was developed, which led to the current best lower bound $\underline{\mathbf{R}}(M_{\langle \mathbf{n}\rangle}) \geq 2\mathbf{n}^2 - \log_2(\mathbf{n}) - 1$ [LM18].

The geometric approach to lower bounds is as follows: let $\sigma_r := \{T \in \mathbb{C}^m\otimes\mathbb{C}^m\otimes\mathbb{C}^m \mid \underline{\mathbf{R}}(T) \leq r\}$, the set of tensors in $\mathbb{C}^m\otimes\mathbb{C}^m\otimes\mathbb{C}^m$ of border rank at most r. This set is an *algebraic variety*, i.e., it is the zero set of a collection of (homogeneous) polynomials. Naïvely expressed, to prove $\underline{\mathbf{R}}(M_{\langle \mathbf{n}\rangle}) > r$, or to prove lower border rank bounds for any tensor, one simply looks for a polynomial in the ideal of σ_r (that is a polynomial P such that $P(T) = 0$ for all $T \in \sigma_r$) such that $P(M_{\langle \mathbf{n}\rangle}) \neq 0$ (here $m = \mathbf{n}^2$). But how can one find such polynomials? This is where *representation theory* comes in. The variety σ_r is *invariant* under changes of bases in the three spaces. That is, write $\mathbb{C}^m\otimes\mathbb{C}^m\otimes\mathbb{C}^m = A\otimes B\otimes C$, and let $GL(A)$ etc.. denote the invertible $m \times m$ matrices. There is a natural action of $G := GL(A) \times GL(B) \times GL(C)$ on $A\otimes B\otimes C$: on rank one tensors $(g_A, g_B, g_C)\cdot(a\otimes b\otimes c) := (g_Aa)\otimes(g_Bb)\otimes(g_Cc)$, and the action on $A\otimes B\otimes C$ is defined by extending this action linearly. Then for all $g \in G$ and $x \in \sigma_r$, one has $g \cdot x \in \sigma_r$. Whenever a variety is invariant under the action of a group, its ideal is invariant under the group as well via the induced action on polynomials. One can then attempt to use representation theory to decompose the space of all polynomials and systematically check which irreducible modules are in the ideal. This works well in small dimensions, e.g., to show $\underline{\mathbf{R}}(M_{\langle 2\rangle}) = 7$ [HIL13], but in general one must use additional methods. A classical approach is to try to embed $A\otimes B\otimes C$ into a space of matrices, and then take minors, which (in a slightly different context) dates back at least to Sylvester. The advance here was to look for *G-equivariant* (*G-homomorphic*) embeddings. This idea led to the 2013 advance, but the limits described in [EGO+17, DM18, Gal17] exactly apply to such embeddings, so to advance further one must find new techniques.

At this point I should mention the general *hay in a haystack* problem of finding explicit sequences of tensors $T_m \in \mathbb{C}^m\otimes\mathbb{C}^m\otimes\mathbb{C}^m$ of high rank and border rank. The maximum border rank is $\lceil \frac{m^3}{3m-2} \rceil$ for $m > 3$, and the maximum possible rank is *not known*. The current state of the art are explicit tensors with $\mathbf{R}(T_m) \geq 3m - o(m)$ [AFT11, Lan14] and others with $\underline{\mathbf{R}}(T_m) \geq 2m - 2$ [Lan15]. The border

substitution method promises to at least improve the state of the art on these problems.

A last remark on lower bounds: in the past two months there has been a very exciting breakthrough due to Buczynska-Buczynski (personal communication), that avoids the above-mentioned barriers in the polynomial situation. The method is a combination of the classical apolarity method with the border substitution method. Buczynski and I are currently working to extend these methods to the tensor situation.

Upper bounds. The *Kronecker product* of $T \in A \otimes B \otimes C$ and $T' \in A' \otimes B' \otimes C'$, denoted $T \boxtimes T'$, is the tensor $T \otimes T' \in (A \otimes A') \otimes (B \otimes B') \otimes (C \otimes C')$ regarded as 3-way tensor. The Kronecker powers of T, $T^{\boxtimes N} \in (A^{\otimes N}) \otimes (B^{\otimes N}) \otimes (C^{\otimes N})$ are defined similarly. Also let $T \oplus T' \in (A \oplus A') \otimes (B \oplus B') \otimes (C \oplus C')$ denote the direct sum. Let $M_{\langle \mathbf{l},\mathbf{m},\mathbf{n}\rangle}$ denote the rectangular matrix multiplication tensor. The matrix multiplication tensor has the remarkable property that $M_{\langle \mathbf{l},\mathbf{m},\mathbf{n}\rangle} \boxtimes M_{\langle \mathbf{l}',\mathbf{m}',\mathbf{n}'\rangle} = M_{\langle \mathbf{ll}',\mathbf{mm}',\mathbf{nn}'\rangle}$ which is the key to Strassen's *laser method*.

Following work of Strassen, Bini [Bin80] showed that for all $\mathbf{l}, \mathbf{m}, \mathbf{n}$, setting $q = (\mathbf{lmn})^{\frac{1}{3}}$, that

$$\omega \leq \frac{\log(\underline{\mathbf{R}}(M_{\langle \mathbf{l},\mathbf{m},\mathbf{n}\rangle}))}{\log(q)}.$$

This was generalized by Schönhage [Sch81]. A special case is as follows: say that for $1 \leq i \leq s$, $\mathbf{l}_i\mathbf{m}_i\mathbf{n}_i = q^3$, then

$$\omega \leq \frac{\log(\frac{1}{s}\underline{\mathbf{R}}(\bigoplus_{i=1}^{s} M_{\langle \mathbf{l}_i,\mathbf{m}_i,\mathbf{n}_i\rangle}))}{\log(q)}.$$

Schönhage also showed that border rank can be strictly sub-additive, so the result is nontrivial. This immediately implies that if T is a tensor such that $\bigoplus_{i=1}^{s} M_{\langle \mathbf{l}_i,\mathbf{m}_i,\mathbf{n}_i\rangle} \in \overline{G \cdot T}$, i.e., T *degenerate*s to $\bigoplus_{i=1}^{s} M_{\langle \mathbf{l}_i,\mathbf{m}_i,\mathbf{n}_i\rangle}$, then

$$\omega \leq \frac{\log(\frac{1}{s}\underline{\mathbf{R}}(T))}{\log(q)}.$$

This can be useful if the border rank of T is easier to estimate than that of the direct sum of matrix multiplication tensors. One should think of $\underline{\mathbf{R}}(T)$ as the *cost* of T and s and $\log(q)$ as determining the *value* of T. One gets a good upper bound if cost is low and value is high. Strassen [Str87] then showed that the same result holds if the matrix multiplication tensors are nearly disjoint (i.e., nearly direct sums) by taking Kronecker powers of T and degenerating the powers to disjoint matrix multiplication tensors. This method was used by Coppersmith and Winograd [CW90] with the now named "little Coppersmith-Winograd tensor":

$$T_{cw,q} := \sum_{j=1}^{q} a_0 \otimes b_j \otimes c_j + a_j \otimes b_0 \otimes c_j + a_j \otimes b_j \otimes c_0,$$

to show

$$\omega \leq \frac{\log(\frac{4}{27}\underline{\mathbf{R}}(T_{cw,q})^3)}{\log(q)}.$$

Since $\underline{\mathbf{R}}(T_{cw,q}) = q+2$, this implies $\omega < 2.41$ when $q = 8$. They improved this to $\omega \leq 2.3755$ using a slightly more complicated tensor (the Kronecker square of the big Coppersmith-Winograd tensor $T_{CW,q}$) which held the world record until 2012-3, when it was lowered to $\omega \leq 2.373$ [Sto10, Wil18, Le G14] using higher Kronecker powers of $T_{CW,q}$. Then in 2014, Ambainus, Filmus and LeGall [AFL15] proved that coordinate restrictions of Kronecker powers of the big Coppersmith-Winograd tensor could never be used to prove $\omega < 2.3$. This was generalized in [AW18b, AW18] to a larger class of tensors and degenerations, albeit with weaker bounds on the limitations. Regarding Kronecker powers, one also has

$$\omega \leq \frac{\log(\frac{4}{27}\underline{\mathbf{R}}(T_{cw,q}^{\boxtimes k})^{\frac{3}{k}})}{\log(q)}, \tag{1}$$

for any k.

I point out that all the above has little to do with practical matrix multiplication, the only better practical decomposition to arise since Strassen's 1968 work is due to V. Pan [Pan84].

Given the barriers from [AFL15, AW18b, AW18], it makes sense to ask what geometry can do for upper bounds.

First idea: study known rank decompositions of $M_{\langle \mathbf{n} \rangle}$ to obtain new ones. For a tensor T, let $G_T := \{g \in G \mid g \cdot T = T\}$, denote the *symmetry group* of T. For example $G_{M_{\langle \mathbf{n} \rangle}}$ is the image of $GL_{\mathbf{n}}^{\times 3}$ in $GL_{\mathbf{n}^2}^{\times 3}$. In [CIL+18, BIL+18] we studied decompositions and noticed that many of the decompositions had large *symmetry groups*, where if $\mathcal{S}_T$ is a rank decomposition of a tensor T, if one applies an element of G_T to the decomposition, it takes it to another rank decomposition of T, which sometimes is the same as the original. Let $\Gamma_{\mathcal{S}_T} := \{g \in G_T \mid g\mathcal{S}_T = \mathcal{S}_T\}$, denote the symmetry group of the decomposition. Then the decomposition may be expressed in terms of the orbit structure. For example, Strassen's 1968 rank 7 decomposition of $M_{\langle 2 \rangle}$ may be written

$$M_{\langle 2 \rangle} = \mathrm{Id}^{\otimes 3} + \Gamma \cdot \left[\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \otimes \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} \otimes \begin{pmatrix} 0 & 1 \\ 0 & -1 \end{pmatrix} \right]$$

where $\Gamma \simeq \mathfrak{S}_3 \rtimes \mathbb{Z}_2$, and $\Gamma\cdot$ denotes the sum of the terms in the Γ-orbit. Here the orbit consists of six terms. The identity is acted on trivially by Γ, so the decomposition is a union of two orbits. This is work in progress.

Second idea: expand the playing field. Given any tensor, one can symmetrize it to get a cubic polynomial. Let $sM_{\langle \mathbf{n} \rangle}$ denote the symmetrized matrix multiplication tensor, the polynomial $X \mapsto \mathrm{trace}(X^3)$. In [CHI+18] we showed that the Waring rank of $sM_{\langle \mathbf{n} \rangle}$ also governs the exponent of matrix multiplication, where the Waring rank of a cubic polynomial is the smallest r such that the polynomial may be written as a sum of r cubes.

Third idea: combine the first two. A Conner [Con17] found a remarkable Waring rank 18 decomposition of $sM_{\langle 3 \rangle}$, with symmetry group that of the Hasse diagram, namely $(\mathbb{Z}_3^{\times 2} \rtimes SL_2(\mathbb{F}_3)) \rtimes \mathbb{Z}_2$. He also found a Waring rank 40 decomposition of $sM_{\langle 4 \rangle}$ with symmetry group that of the cube. For comparison, the best known rank

decompositions of $M_{\langle 3\rangle}, M_{\langle 4\rangle}$ respectively are of ranks 23 and 49. This launched his program to find explicit sequences of finite groups $\Gamma_{\mathbf{n}} \subset G_{sM_{\langle \mathbf{n}\rangle}}$ such that the space of $\Gamma_{\mathbf{n}}$-invariants in the space of cubic polynomials in $\mathbf{n}^2$ variables only contains polynomials of low Waring rank, translating the study of upper bounds on ω to a study of properties of sequences of finite groups, in the spirit of (but very different from) the Cohn-Umans program [CU03].

Fourth idea: the Ambainus-Filmus-LeGall challenge: find new tensors useful for the laser method. Michalek and I [LM17] had the idea to isolate *geometric* properties of the Coppersmith-Winograd tensors and to find other tensors with similar geometric properties, in the hope that they might also be useful for the laser method. We succeeded in isolating many interesting geometric properties. Unfortunately, we then proved that the Coppersmith-Winograd tensors were the *unique* tensors with such properties.

Fourth idea, second try: In [CGL+18] we examine the symmetry groups of the Coppersmith-Winograd tensors. We found that the big Coppersmith-Winograd tensor has the largest dimensional symmetry group among 1-*generic tensors* in odd dimensions (1-genericity is a natural condition for implementing the laser method), but that in even dimensions, there is an even better tensor, which we call the *skew big Coppersmith-Winograd tensor*. We also found other tensors with large symmetry groups. Unfortunately, none of the new tensors with maximal or near maximal symmetry groups are better for the laser method than $T_{CW,q}$.

Fifth idea: Go back to the inequality (1) and upper bound Kronecker powers of $T_{cw,q}$ (this was posed as an open question for the square as early as [Bla13]), which could even (when $q = 2$) potentially show $\omega = 2$. Unfortunately, we show in [CGL+18] that $15 \leq \underline{\mathbf{R}}(T_{cw,2}^{\boxtimes 2}) \leq 16$, and we expect 16, and for $q > 2$, that $\underline{\mathbf{R}}(T_{cw,q}^{\boxtimes 2}) = (q+2)^2$.

Sixth idea: Combine the last two ideas. The skew cousin of the little Coppersmith-Winograd tensor also satisfies (1). It has the same value, but unfortunately, it has higher cost. For example, $\underline{\mathbf{R}}(T_{skew-cw,2}) = 5 > 4 = \underline{\mathbf{R}}(T_{cw,2})$. However, we show it satisfies $\underline{\mathbf{R}}(T_{skew-cw,2}^{\boxtimes 2}) = 17 \ll 25 = \underline{\mathbf{R}}(T_{skew-cw,2})^2$. This is one of the few explicit tensors known to have strictly submultiplicative border rank under Kronecker square (the first known being $M_{\langle 2\rangle}$), and if this drop continues, it would be very good indeed for the laser method.

Acknowledgments. Supported by NSF grants DMS-1405348 and AF-1814254. This work was partially supported by the grant 346300 for IMPAN from the Simons Foundation and the matching 2015-2019 Polish MNiSW fund as well as an Simons Visiting Professor grant supplied by the Simons Foundation and by the Mathematisches Forschungsinstitut Oberwolfach.

References

[AFT11] Boris Alexeev, Michael A. Forbes, and Jacob Tsimerman, *Tensor rank: some lower and upper bounds*, 26th Annual IEEE Conference on Computational Complexity, IEEE Computer Soc., Los Alamitos, CA, 2011, pp. 283–291. MR 3025382

[AW18] J. Alman and V. Vassilevska Williams, *Limits on All Known (and Some Unknown) Approaches to Matrix Multiplication*, ArXiv e-prints (2018).

[AW18b] Josh Alman and Virginia Vassilevska Williams, *Further limitations of the known approaches for matrix multiplication*, 9th Innovations in Theoretical Computer Science Conference, ITCS 2018, January 11-14, 2018, Cambridge, MA, USA, 2018, pp. 25:1–25:15.

[AFL15] Andris Ambainis, Yuval Filmus, and François Le Gall, *Fast matrix multiplication: limitations of the Coppersmith-Winograd method (extended abstract)*, STOC'15—Proceedings of the 2015 ACM Symposium on Theory of Computing, ACM, New York, 2015, pp. 585–593. MR 3388238

[BIL+18] Grey Ballard, Christian Ikenmeyer, J.M. Landsberg, and Nick Ryder, *The geometry of rank decompositions of matrix multiplication ii: 3x3 matrices*, to appear in JPAA.

[Bin80] D. Bini, *Relations between exact and approximate bilinear algorithms. Applications*, Calcolo **17** (1980), no. 1, 87–97. MR 605920 (83f:68043b)

[Bla13] Markus Bläser, *Fast matrix multiplication*, Graduate Surveys, no. 5, Theory of Computing Library, 2013.

[CHI+18] Luca Chiantini, Jonathan D. Hauenstein, Christian Ikenmeyer, Joseph M. Landsberg, and Giorgio Ottaviani, *Polynomials and the exponent of matrix multiplication*, Bull. Lond. Math. Soc. **50** (2018), no. 3, 369–389. MR 3829726

[CIL+18] Luca Chintini, Christian Ikenmeyer, J.M. Landsberg, and Giorgio Ottaviani, *The geometry of rank decompositions of matrix multiplication i: Strassen's algorithm*, to appear in Exper. Math.

[CU03] H Cohn and C. Umans, *A group theoretic approach to fast matrix multiplication*, Proceedings of the 44th annual Symposium on Foundations of Computer Science (2003), no. 2, 438–449.

[Con17] A. Conner, *A rank 18 Waring decomposition of $sM\_\langle 3\rangle$ with 432 symmetries*, ArXiv e-prints, to appear in Exper. Math. (2017).

[CGL+18] Austin Conner, Fulvio Gesmundo, J.M. Landsberg, and Emanuele Ventura, *Kronecker powers of tensors with symmetry*, preprint.

[CGL+18] ———, *Tensors with maximal symmetries*, preprint.

[CW90] Don Coppersmith and Shmuel Winograd, *Matrix multiplication via arithmetic progressions*, J. Symbolic Comput. **9** (1990), no. 3, 251–280. MR 91i:68058

[DM18] Harm Derksen and Visu Makam, *On non-commutative rank and tensor rank*, Linear Multilinear Algebra **66** (2018), no. 6, 1069–1084. MR 3781583

[EGO+17] K. Efremenko, A. Garg, R. Oliveira, and A. Wigderson, *Barriers for Rank Methods in Arithmetic Complexity*, ArXiv e-prints (2017).

[Gal17] Maciej Galcazka azka, *Vector bundles give equations of cactus varieties*, Linear Algebra Appl. **521** (2017), 254–262. MR 3611482

[HIL13] Jonathan D. Hauenstein, Christian Ikenmeyer, and J. M. Landsberg, *Equations for lower bounds on border rank*, Exp. Math. **22** (2013), no. 4, 372–383. MR 3171099

[Lan14] J. M. Landsberg, *New lower bounds for the rank of matrix multiplication*, SIAM J. Comput. **43** (2014), no. 1, 144–149. MR 3162411

[Lan15] ———, *Nontriviality of equations and explicit tensors in $\mathbb{C}^m \otimes \mathbb{C}^m \otimes \mathbb{C}^m$ of border rank at least $2m-2$*, J. Pure Appl. Algebra **219** (2015), no. 8, 3677–3684. MR 3320240

[LM17] J. M. Landsberg and Mateusz Michalek, *Abelian tensors*, J. Math. Pures Appl. (9) **108** (2017), no. 3, 333–371. MR 3682743

[LM17b] ———, *On the geometry of border rank decompositions for matrix multiplication and other tensors with symmetry*, SIAM J. Appl. Algebra Geom. **1** (2017), no. 1, 2–19. MR 3633766

[LM18] Joseph M. Landsberg and Mateusz Michał ek, *A* $2n^2 - \log_2(n) - 1$ *lower bound for the border rank of matrix multiplication*, Int. Math. Res. Not. IMRN (2018), no. 15, 4722–4733. MR 3842382

[LO15] Joseph M. Landsberg and Giorgio Ottaviani, *New lower bounds for the border rank of matrix multiplication*, Theory Comput. **11** (2015), 285–298. MR 3376667

[Le G14] François Le Gall, *Powers of tensors and fast matrix multiplication*, Proceedings of the 39th International Symposium on Symbolic and Algebraic Computation (New York, NY, USA), ISSAC '14, ACM, 2014, pp. 296–303.

[Lic84] Thomas Lickteig, *A note on border rank*, Inform. Process. Lett. **18** (1984), no. 3, 173–178. MR 86c:68040

[Pan84] Victor Pan, *How to multiply matrices faster*, Lecture Notes in Computer Science, vol. 179, Springer-Verlag, Berlin, 1984. MR MR765701 (86g:65006)

[Sch81] A. Schönhage, *Partial and total matrix multiplication*, SIAM J. Comput. **10** (1981), no. 3, 434–455. MR MR623057 (82h:68070)

[Sto10] A. Stothers, *On the complexity of matrix multiplication*, PhD thesis, University of Edinburgh, 2010.

[Str83] V. Strassen, *Rank and optimal computation of generic tensors*, Linear Algebra Appl. **52/53** (1983), 645–685. MR 85b:15039

[Str87] ———, *Relative bilinear complexity and matrix multiplication*, J. Reine Angew. Math. **375/376** (1987), 406–443. MR MR882307 (88h:11026)

[Str69] Volker Strassen, *Gaussian elimination is not optimal*, Numer. Math. **13** (1969), 354–356. MR 40 #2223

[Wil18] Virginia Williams, *Breaking the coppersimith-winograd barrier*, preprint.

On the Foundation of Program Obfuscation

Huijia (Rachel) Lin
(joint work with Christian Matt, Stefano Tessaro)

Indistinguishability obfuscation (IO), first defined in the seminal work of Barak et. al. [BGI+01], aims to obfuscate functionally equivalent programs into indistinguishable ones while preserving functionality. IO is an extraordinarily powerful object that has been shown to enable a large set of new cryptographic applications.

Starting from [GGH+13], the first-generation IO constructions rely on polynomial-degree *multilinear maps*. An L-linear map [BS03] essentially allows to evaluate degree-L polynomials on secret encoded values, and to test whether the output of such polynomials is zero or not. While bilinear maps (i.e., $L = 2$) can be efficiently instantiated from elliptic curves, instantiation of L-linear maps for $L \geq 3$ has remained elusive—so far, vulnerabilities were demonstrated against all known candidates based on Lattices.

Motivated by the state-of-affairs, several recent works [Lin16, AS17, Lin17, LT17, Agr18a, AJS18, LM18] focused on building IO from minimal degree multilinear maps, leading to new constructions from trilinear or even bilinear maps. These new constructions crucially rely on Pseudo Random Generators (PRGs) with special simple structures. In this talk, we give an overview of some of the new constructions.

In the first line of works [Lin16, AS17, Lin17, LT17], we show that PRGs with small output locality, or even a relaxed notion of block-locality, can be instrumental for constructing IO. A PRG has block locality L if every output bit depends on at most L *input blocks*, each consisting of up to $\log \lambda$ input bits. We show a construction of IO from block-locality L PRGs and L-linear maps, assuming additionally the Learning With Errors (LWE) assumption. PRGs with block-locality $L \geq 3$ can be instantiated by generalizing Goldreich's local functions [Gol00, MST03, OW14, AL16]. Unfortunately, PRGs with block-locality $L \geq 2$ have been shown vulnerable to attacks [LV17, BBKK18].

Towards the goal of constructing IO from bilinear maps, we [LM18] seek PRGs with alternative structures and weak security guarantees that are useful for constructing IO and can be evaluated using just bilinear maps. To this end, we introduce *Pseudo Flawed-smudging Generators* (PFGs). A PFG is a function $g : \mathbb{Z}_p^n \to \mathbb{Z}_p^m$ with a polynomial stretch $m = n^{1+\epsilon}$ for $\epsilon > 0$, and an input distribution $\mathcal{X}$. Its output distribution $g(\mathcal{X})$ is *i)* polynomially bounded (i.e., every output has polynomial infinity norm), and *ii)* can be used to partially hide a small noise vector with inverse polynomial probability. Assuming LWE and the existence of constant-locality PRGs, we show a construction of IO from PFGs computable by degree-d polynomials and d-linear maps. Candidate degree 2 PFGs that are special instances of multivariate *quadratic* polynomials over $\mathbb{Z}_p$ have been proposed, which gives the first candidate IO from bilinear maps.

References

[Agr18a] Shweta Agrawal. New methods for indistinguishability obfuscation: Bootstrapping and instantiation. Cryptology ePrint Archive, Report 2018/633, 2018. `https://eprint.iacr.org/2018/633`.

[Gol00] O. Goldreich, "Candidate one-way functions based on expander graphs," *Electronic Colloquium on Computational Complexity (ECCC)*, vol. 7, no. 90, 2000.

[MST03] E. Mossel, A. Shpilka, and L. Trevisan, "On e-biased generators in NC0," in *44th Symposium on Foundations of Computer Science (FOCS 2003), 11-14 October 2003, Cambridge, MA, USA, Proceedings*, pp. 136–145, 2003.

[OW14] R. O'Donnell and D. Witmer, "Goldreich's PRG: evidence for near-optimal polynomial stretch," in *IEEE 29th Conference on Computational Complexity, CCC 2014, Vancouver, BC, Canada, June 11-13, 2014*, pp. 1–12, 2014.

[AL16] B. Applebaum and S. Lovett, "Algebraic attacks against random local functions and their countermeasures," in *Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2016, Cambridge, MA, USA, June 18-21, 2016*, pp. 1087–1100, 2016.

[Lin16] H. Lin, "Indistinguishability obfuscation from constant-degree graded encoding schemes," 2016. To Appear in Eurocrypt'16.

[AS17] Prabhanjan Ananth and Amit Sahai. Projective arithmetic functional encryption and indistinguishability obfuscation from degree-5 multilinear maps. In Jean-Sébastien Coron and Jesper Buus Nielsen, editors, *Advances in Cryptology – EUROCRYPT 2017*, pages 152–181, Cham, 2017. Springer International Publishing.

[Lin17] Huijia Lin. Indistinguishability obfuscation from SXDH on 5-linear maps and locality-5 PRGs. In Jonathan Katz and Hovav Shacham, editors, *Advances in Cryptology – CRYPTO 2017*, pages 599–629, Cham, 2017. Springer International Publishing.

[LT17] Huijia Lin and Stefano Tessaro. Indistinguishability obfuscation from trilinear maps and block-wise local PRGs. In Jonathan Katz and Hovav Shacham, editors, *Advances in Cryptology – CRYPTO 2017*, pages 630–660, Cham, 2017. Springer International Publishing.

[BBKK18] Boaz Barak, Zvika Brakerski, Ilan Komargodski, and Pravesh K. Kothari. Limits on low-degree pseudorandom generators (or: Sum-of-squares meets program obfuscation). In Jesper Buus Nielsen and Vincent Rijmen, editors, *Advances in Cryptology – EUROCRYPT 2018*, pages 649–679, Cham, 2018. Springer International Publishing.

[AJS18] Prabhanjan Ananth, Aayush Jain, and Amit Sahai. Indistinguishability obfuscation without multilinear maps: iO from LWE, bilinear maps, and weak pseudorandomness. Cryptology ePrint Archive, Report 2018/615, 2018. `https://eprint.iacr.org/2018/615`.

[BHJ+18] Boaz Barak, Sam Hopkins, Aayush Jain, Pravesh Kothari, and Amit Sahai. Sum-of-squares meets program obfuscation, revisited. *Unpublished Work*, 2018.

[LM18] Huijia Lin, Christian Matt. Pseudo Flawed-Smudging Generators and Their Application to Indistinguishability Obfuscation. Cryptology ePrint Archive, Report 2018/646, 2018. `https://eprint.iacr.org/2018/646`.

[BGI+01] Boaz Barak, Oded Goldreich, Rusell Impagliazzo, Steven Rudich, Amit Sahai, Salil Vadhan, and Ke Yang. On the (im)possibility of obfuscating programs. In Joe Kilian, editor, *Advances in Cryptology — CRYPTO 2001*, pages 1–18, Berlin, Heidelberg, 2001. Springer.

[GGH+13] Sanjam Garg, Craig Gentry, Shai Halevi, Mariana Raykova, Amit Sahai, and Brent Waters. Candidate indistinguishability obfuscation and functional encryption for all circuits. In *2013 IEEE 54th Annual Symposium on Foundations of Computer Science*, pages 40–49, 10 2013.

[BS03] Dan Boneh and Alice Silverberg. Applications of multilinear forms to cryptography. *Contemporary Mathematics*, 324(1):71–90, 2003.

[LV17] Alex Lombardi and Vinod Vaikuntanathan. Limits on the locality of pseudorandom generators and applications to indistinguishability obfuscation. In Yael Kalai and Leonid Reyzin, editors, *Theory of Cryptography*, pages 119–137, Cham, 2017. Springer International Publishing.

On the Hardness of 2-to-2 Games

Dor Minzer

(joint work with Irit Dinur, Subhash Khot, Guy Kindler and Muli Safra)

Background. The PCP Theorem [FGL+96, AS98, AKM+98] is a major result in Theoretical Computer Science from the 90's, giving a new characterization of the computational class NP, using highly efficient probabilistic verification. This result has many implications throughout TCS, in particular in the field of Hardness of Approximation.

Definition 1. *An instance of the Label-Cover problem is* $\psi = (G = (U \cup V, E), \Sigma_U, \Sigma_V, \Phi)$ *where* G *is an undirected bipartite graph,* Σ_U *are finite set of labels, and* $\Phi = \{\phi_e \mid e \in E\}$ *is a set of constraints* $\phi_e \colon \Sigma_U \to \Sigma_V$*, one for each edge.*

The value of an instance, denoted by $\mathsf{val}(\psi)$*, is the maximum fraction of constraints an assignment of labels to* U *(by* Σ_U *symbols) and to* V *(by* Σ_V *symbols) satisfies.*

For $s < c$ between 0 and 1, the gap-LabelCover$_k[c, s]$ is the promise problem in which one is given a Label-Cover instance ψ with alphabet sizes at most k. The goal is to accept an instance if it has value at least c, and reject it if it has value at most s. One of the many formulations of the PCP Theorem states:

Theorem 2 (The PCP Theorem)**.** *There are $s < 1$ and $k \in \mathbb{N}$ such that* gap-LabelCover$_k[1, s]$ *is NP-hard.*

The hardness of Label-Cover is one of the most popular starting points for hardness of approximation results. The above formulation however, is often times too weak to yield strong inapproximability results, and one has to use stronger versions. The two most useful strengthenings of the PCP theorem address the following points: (1) decrease the soundness value, s, for which the problem remains NP-hard (preferably close to 0), (2) obtain strong structure on the constraints in the Label-Cover instance. The first of these points was largely achieved: first, by amplification of the soundness parameter s using Raz's Parallel Repetition Theorem [Raz98] [1], and secondly by a more direct construction of Moshkovitz and Raz [MR10, DH13]. Using low soundness PCPs, along with the Long-Code [BGS95] and Fourier analysis, one can prove numerous strong hardness of approximation results, see for instance [Has96, Has97, ABH+05, Fei96, ST06, DKS98, DS05]. The second point, of imposing strong structure from the constraints of the Label-Cover instance, is still largely an open question. In its most extreme form, this problem is known as Khot's Unique-Games Conjecture.

The d-to-1 Games Conjectures. A label cover instance is called d-to-1 Game if for every edge e and every $\sigma \in \Sigma_V$, the pre-image of σ under ϕ_e contains at most d elements. This notion was first studied (somewhat implicitly) in the work of Dinur and Safra [DS05], wherein they prove it is NP-hard to approximate the minimum Vertex-Cover within factor ≈ 1.36.

Let d-to-1 Games$_k$ be the maximization problem where on is given a d-to-1 Label-Cover instance with alphabet size k, and the goal is to find an assignment satisfying maximum fraction of constraints. The d-to-1 Games Conjecture of Khot [Kho02] states that for $d \geq 2$, distinguishing between satisfiable instances, and at most $o(1)$-value instances, is NP hard:

Conjecture 3. *For any $d \geq 2$ and $s > 0$, there exists $k > 0$ such that* d-to-1 Games$_k[1, s]$ *is NP-hard.*

A related, more well known problem, is the Unique-Games Conjecture, stating that for $d = 1$ (in which case the 1-to-1 Games problem is often referred to as the Unique-Games problem), it is NP-hard to distinguish $1 - o(1)$-value instances from $o(1)$-value instances.[2]

[1] By now there are multiple simplifications and strengthenings of Raz's result, for instnace [Hol07, Rao11, DS14, BG15].

[2] A simple propagation algorithm shows that given a satisfiable Unique-Games instance, one can efficiently find an assignment that satisfies all constraints. Thus, for the problem to have a chance at being NP-hard, one must settle for imperfect completeness.

Conjecture 4 (Khot's Unique-Games Conjecture [Kho02])**.** *For any $\varepsilon > 0$, there exists $k > 0$ such that* $\mathsf{Unique\text{-}Games}_k[1-\varepsilon, \varepsilon]$ *is NP-hard.*

The Unique-Games Conjecture is now a prominent open question, known to have vast implications. Research into the conjecture has mainly developed in three directions:

(1) Attempting to refute the conjecture by designing algorithms, for instance [Kho02, GT06, CMM06, Tre08, AKK+08, Kol10, ABS15, BBH+12]. All of these algorithms fall short of refuting the Unique-Games Conjecture.
(2) Proving consequences of the conjecture. Much progress has been achieved in this direction, see for instance [KKM+07, KR08, CKK+05, DMR09], and the surveys [Kho14, Tre12]. A striking result in this direction by Raghavendra [Rag08], states that assuming the Unique-Games Conjecture, the best efficient approximation algorithm for the class of Constraint Satisfaction Problems, is a simple Semi-Definite Programming algorithm.
(3) Partial results and candidate hard constructions. Among these are the work of O'Donnell and Wright [OW12], who proved it is NP-hard to distinguish between instances whose value is $\geq \frac{1}{2}$ and instances whose value is $\leq \frac{3}{8}$, and the work of Khot and Moshkovitz [KM16], who gave a candidate construction for a hard Unique-Games instance.

The works discussed in this talk [KMS17, DKK+18, DKK+18, KMS18, BKS18] belong to the third direction. The main result proved therein asserts that the 2-to-1 Games Conjecture holds, albeit with imperfect completeness.

Theorem 5. *For every $\varepsilon > 0$ there exists $k \in \mathbb{N}$ such that* $\mathsf{2\text{-}to\text{-}1\ Games}_k[1-\varepsilon, \varepsilon]$ *is NP-hard.*

This result has several consequences: for combinatorial optimization problems (such as Vertex-Cover, Max-Cut-Gain and Approximate-Coloring, see [DKK+18, KMS18]), and for Unique-Games (achieving, for the first time, a hard gap known with constant completeness and vanishing soundness).

Theorem 6. *For every $\varepsilon > 0$ there exists $k \in \mathbb{N}$ such that* $\mathsf{Unique\text{-}Games}_k[\frac{1}{2}, \varepsilon]$ *is NP-hard.*

This talk discussed some ideas that go into the proof of the 2-to-1 Games Conjecture with imperfect completeness, such as smooth parallel repetition, the Covering Property [KS11], the reduction and the Grassmann Test [KMS17, DKK+18], and its analysis via structure of non-optimally expanding sets on the Grassmann Graph [DKK+18, KMS18, BKS18].

References

[AKM+98] S. Arora, C. Lund, R. Motawani, M. Sudan, and M. Szegedy. Proof verification and the hardness of approximation problems. *Journal of the ACM*, 45(3):501–555, 1998.

[AS98] S. Arora and S. Safra. Probabilistic checking of proofs : A new characterization of NP. *Journal of the ACM*, 45(1):70–122, 1998.

[ABS15] Sanjeev Arora, Boaz Barak, and David Steurer. Subexponential algorithms for unique games and related problems. *J. ACM*, 62(5):42:1–42:25, 2015.

[ABH+05] Sanjeev Arora, Eli Berger, Elad Hazan, Guy Kindler, and Muli Safra. On non-approximability for quadratic programs. In *Proc. Annual IEEE Symposium on Foundations of Computer Science*, pages 206–215, 2005.

[AKK+08] Sanjeev Arora, Subhash Khot, Alexandra Kolla, David Steurer, Madhur Tulsiani, and Nisheeth K. Vishnoi. Unique games on expanding constraint graphs are easy. In *Proc. ACM Symposium on the Theory of Computing*, pages 21–28, 2008.

[BBH+12] Boaz Barak, Fernando G. S. L. Brandão, Aram Wettroth Harrow, Jonathan A. Kelner, David Steurer, and Yuan Zhou. Hypercontractivity, sum-of-squares proofs, and their applications. In *Proceedings of the 44th Symposium on Theory of Computing Conference, STOC 2012, New York, NY, USA, May 19 - 22, 2012*, pages 307–326, 2012.

[BKS18] Boaz Barak, Pravesh Kothari, and David Steurer. Small-set expansion in shortcode graph and the 2-to-2 conjecture. *Electronic Colloquium on Computational Complexity (ECCC)*, 25:77, 2018.

[BGS95] M. Bellare, O. Goldreich, and M. Sudan. Free bits, PCPs and non-approximability. *Electronic Colloquium on Computational Complexity, Technical Report TR95-024*, 1995.

[BG15] 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.

[CMM06] Moses Charikar, Konstantin Makarychev, and Yury Makarychev. Near-optimal algorithms for unique games. In *Proc. ACM Symposium on the Theory of Computing*, pages 205–214, 2006.

[CKK+05] S. Chawla, R. Krauthgamer, R. Kumar, Y. Rabani, and D. Sivakumar. On the hardness of approximating multicut and sparsest-cut. In *Proc. 20th IEEE Conference on Computational Complexity*, pages 144–153, 2005.

[DH13] Irit Dinur and Prahladh Harsha. Composition of low-error 2-query pcps using decodable pcps. *SIAM J. Comput.*, 42(6):2452–2486, 2013.

[DKK+18] Irit Dinur, Subhash Khot, Guy Kindler, Dor Minzer, and Muli Safra. On non-optimally expanding sets in grassmann graphs. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2018, Los Angeles, CA, USA, June 25-29, 2018*, pages 940–951, 2018.

[DKK+18] Irit Dinur, Subhash Khot, Guy Kindler, Dor Minzer, and Muli Safra. Towards a proof of the 2-to-1 games conjecture? In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2018, Los Angeles, CA, USA, June 25-29, 2018*, pages 376–389, 2018.

[DKS98] Irit Dinur, Guy Kindler, and Muli Safra. Approximating CVP to within almost-polynomial factors is NP-hard. In *Proc. 39th IEEE Symposium on Foundations of Computer Science*, 1998.

[DMR09] Irit Dinur, Elchanan Mossel, and Oded Regev. Conditional hardness for approximate coloring. *SIAM J. Comput.*, 39(3):843–873, 2009.

[DS05] Irit Dinur and Samuel Safra. On the hardness of approximating minimum vertex cover. *Ann. of Math. (2)*, 162(1):439–485, 2005.

[DS14] Irit Dinur and David Steurer. Analytical approach to parallel repetition. In *Symposium on Theory of Computing, STOC 2014, New York, NY, USA, May 31 - June 03, 2014*, pages 624–633, 2014.

[Fei96] Uriel Feige. A threshold of $\ln n$ for approximating set cover. *Journal of the ACM*, 45(4):634–652, 1998.

[FGL+96] Uriel Feige, Shafi Goldwasser, László Lovász, Shmuel Safra, and Mario Szegedy. Interactive proofs and the hardness of approximating cliques. *J. ACM*, 43(2):268–292, 1996.

[GT06] A. Gupta and K. Talwar. Approximating unique games. In *Proc. ACM-SIAM Symposium on Discrete Algorithms*, pages 99–106, 2006.

[Has96] Johan Håstad. Clique is hard to approximate within $n^{1-\epsilon}$. *Acta Mathematica*, 182:105–142, 1999.

[Has97] Johan Håstad. Some optimal inapproximability results. *Journal of ACM*, 48:798–859, 2001.

[Hol07] Thomas Holenstein. Parallel repetition: simplifications and the no-signaling case. In *Proc. ACM Symposium on the Theory of Computing*, pages 411–419, 2007.

[Kho02] Subhash Khot. On the power of unique 2-prover 1-round games. In *Proc. 34th ACM Symposium on Theory of Computing*, 2002.

[Kho14] Subhash Khot. Hardness of approximation. *In Proc. of the International Congress of Mathematicians*, 2014.

[KKM+07] Subhash Khot, Guy Kindler, Elchanan Mossel, and Ryan O'Donnell. Optimal inapproximability results for MAX-CUT and other 2-variable CSPs? *SIAM J. Comput.*, 37(1):319–357, 2007.

[KMS17] Subhash Khot, Dor Minzer, and Muli Safra. On independent sets, 2-to-2 games, and grassmann graphs. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2017, Montreal, QC, Canada, June 19-23, 2017*, pages 576–589, 2017.

[KMS18] Subhash Khot, Dor Minzer, and Muli Safra. Pseudorandom sets in grassmann graph have near-perfect expansion. In *59th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2018, Paris, France, October 7-9, 2018*, pages 592–601, 2018.

[KM16] Subhash Khot and Dana Moshkovitz. Candidate hard unique game. In *Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2016, Cambridge, MA, USA, June 18-21, 2016*, pages 63–76, 2016.

[KR08] Subhash Khot and Oded Regev. Vertex cover might be hard to approximate to within $2-\epsilon$. *J. Comput. Syst. Sci.*, 74(3):335–349, 2008.

[KS11] Subhash Khot and Muli Safra. A two prover one round game with strong soundness. In *FOCS*, pages 648–657, 2011.

[Kol10] A. Kolla. Spectral algorithms for unique games. In *Proc. Annual IEEE Conference on Computational Complexity*, 2010.

[MR10] Dana Moshkovitz and Ran Raz. Two-query pcp with subconstant error. *Journal of the ACM*, 57(5), 2010.

[OW12] Ryan O'Donnell and John Wright. A new point of NP-hardness for unique games. In *Proceedings of the 44th Symposium on Theory of Computing Conference, STOC 2012, New York, NY, USA, May 19 - 22, 2012*, pages 289–306, 2012.

[Rag08] Prasad Raghavendra. Optimal algorithms and inapproximability results for every CSP? In *Proceedings of the 40th Annual ACM Symposium on Theory of Computing, Victoria, British Columbia, Canada, May 17-20, 2008*, pages 245–254, 2008.

[Rao11] Anup Rao. Parallel repetition in projection games and a concentration bound. *SIAM J. Comput.*, 40(6):1871–1891, 2011.

[Raz98] Ran Raz. A parallel repetition theorem. *SIAM J. of Computing*, 27(3):763–803, 1998.

[ST06] A. Samorodnitsky and L. Trevisan. Gowers uniformity, influence of variables, and pcps. In *Proc. 38th ACM Symposium on Theory of Computing*, 2006.

[Tre08] Luca Trevisan. Approximation algorithms for unique games. *Theory of Computing*, 4(1):111–128, 2008.

[Tre12] Luca Trevisan. On Khot's unique games conjecture. *Bull. Amer. Math. Soc. (N.S.)*, 49(1):91–111, 2012.

Recent Developments Related to RL vs. L

Omer Reingold

One of the most important complexity-theoretic challenges is showing that randomized algorithms are not much more powerful than deterministic algorithms. Specifically, the two main challenges are to show that randomness "cannot save time" and that randomness "cannot save memory". Our focus here is on the latter. More precisely, the ultimate goal is to show that every problem solvable by a randomized space-bounded algorithm is also solvable by a deterministic algorithm that only uses a constant factor more space (where space refers to memory). By standard padding arguments, this is equivalent to showing that randomized logspace equals deterministic logspace (i.e. RL = L or BPL = L, depending on whether we consider 1-sided or 2-sided error). This problem has drawn a considerable attention in recent decades, leading to exciting research, and here we discuss some fundamental progress from the last couple of years.

A major tool in the study of the RL vs. L problem is pseudorandom-generators that fool bounded space computations. For technical reasons, these generators need to fool a stronger (non-uniform) model of computation of a (layered) read-once branching-programs (ROBP). In this model, each of n layers contain w vertices (we refer to w as the width of the program and to n as its length). Each vertex in layer $i < n$ is connected to two vertices in layer $i+1$. One of these edges is labeled 0 and the other is labeled 1. One of the verices in layer one is the start vertex and one of the vertices in layer n is an accept vertex. Most of these generators apply to ordered ROBP. Such a program accepts the string $x = x_1x_2...x_n$ if following the path made of edges labeled x_1, x_2, $\ldots x_n$ from the start vertex leads to the accept vertex (and it rejects x otherwise). We will also consider unordered ROBP, where the order of reading the bits can be arbitrary (but fixed for all computation paths). The Nisan generator [Nis92, INW94] expands a short seed of $O(\log n \log(wn/\epsilon))$ bits to n bits that fool (ordered) ROBP up to error ϵ.

We discuss three recent threads of research:

- Hitting-set generators and pseudorandom pseudo-distributions with better dependence on the error parameter than Nisan's generator.
- Small memory algorithms for analyzing the behavior of random walks in undirected graphs.
- Pseudorandom generators that fool constant-width ROBPs ($w = O(1)$) with close to optimal seed.

Dependence on the error parameter. To prove that RL=L, using a pseudorandom generator, we would like the seed length to be $O(\log(wn/\epsilon))$. Nevertheless, significant improvements towards derandomizing RL can be obtained already if the dependence of rhe seed on the width and the error improves. In two recent works [BCG18, HZ18], objects that are somewhat weaker than pseudorandom generators with a close to optimal dependence on the error parameter ϵ. We focused on the result of [BCG18] that provided seed-$\tilde{O}(\log n \log(wn) + log(1/\epsilon))$ pseudorandom pseudo-distributions. Specifically, we considered the notion of pseudo-distributions that was introduced by this work into this area and is quite intriguing. Consider a collection $\tilde{D}$ of pairs (ρ_i, s_i), where each ρ_i is a real number (a weight) and each s_i an n-bit string. If the the weights are all non-negative and sum to one, then $\tilde{D}$ defines a distribution on n-bit strings. If the weights are allowed to be negative, then this is a pseudo-distribution. Generating a pseudorandom pseudo-distribution using a short seed could still derandomize (1-sided) random computations (thus it can prove RL=L). We hinted on the way pseudo-distributions come about in [BCG18] through an analogy to a recursive construction of averaging samplers.
Undirected random walks. Connectivity in undirected graphs have been shown to be computable in logarithmic space [Rei08]. The result also implies a particular kind of pseudorandom walks known as pseudo-converging walks [Rei08, RTV06] that behave like random-walks in the limit. In other words, these walks converge to the stationary distribution of the random walk (in number of steps that is polynomially related to the convergence time of the random walk). We discussed recent works [MRS+17] (as well as a more recent, unpublished work) that give (almost) logarithmic-space computations of various parameters of the random walks on undirected graphs such as hitting times, commute times, escape probabilities and t-step conductance. This is obtained through small space computations of the Laplacian of undirected graphs and t-step Laplacian of such graphs. The results combine techniques from small-space derandomization with techniques that arise in the literature on almost linear time Laplacian solvers.
Constant-width ROBP. The last thread we discussed has to do with pseudorandom generators that fool constant-width ROBP. Such pseudorandom generators generalize many fundamental pseudorandom objects (such as epsilon-bias distributions). We discussed the more complete history, leading to several recent milestones, in two directions:

- [CHR+18, FK18] give polylogarithmic seed (in fact, $\tilde{O}(\log^2 n)$-long seed) pseudorandom generators that fool *unordered* ROBP (major progress over previous results).
- [MRT18] give seed-$\tilde{O}(\log n)$ pseudorandom generators that fool width-3 (ordered) ROBP (previously, no improvement over Nisans generator was known even for general width 3 ROBP).

The results rely on a the mild pseudorandom restriction technique of [GMR+12], which is significantly different than the recursive approach of [Nis92, INW94].

References

[BCG18] M. Braverman, G. Cohen, and S. Garg, *Hitting sets with near-optimal error for read-once branching programs*, STOC, (2018), 353–362.

[CHR+18] E. Chattopadhyay, P. Hatami, O. Reingold, and A. Tal, *Improved pseudorandomness for unordered branching programs through local monotonicity*, STOC, (2018), 363–375

[FK18] M. A. Forbes, and Z. Kelley, *Pseudorandom Generators for Read-Once Branching Programs, in Any Order*, FOCS, (2018), 946-955

[GMR+12] P. Gopalan, R. Meka, O. Reingold, L. Trevisan, and S. P. Vadhan, *Better Pseudorandom Generators from Milder Pseudorandom Restrictions*, FOCS, (2012), 120-129.

[HZ18] W. Hoza, and D. Zuckerman, *Simple Optimal Hitting Sets for Small-Success RL*, FOCS, (2018), 59-64.

[INW94] R. Impagliazzo, N. Nisan, and A. Wigderson, *Pseudorandomness for network algorithms*, STOC, (1994), 356–364.

[MRT18] R. Meka, O. Reingold, and A. Tal, *Pseudorandom Generators for Width-3 Branching Programs*, manuscript, 2018.

[MRS+17] J. Murtagh, O. Reingold, A. Sidford, and S. P. Vadhan, *Derandomization Beyond Connectivity: Undirected Laplacian Systems in Nearly Logarithmic Space*, FOCS, (2017), 801-812

[Nis92] N. Nisan, *Pseudorandom generators for space-bounded computation*, Combinatorica, **12(4)** (1992), 449–461.

[Rei08] O. Reingold, *Undirected connectivity in log-space*, J. ACM **55(4)**, (2008), 17:1-17:24

[RTV06] O. Reingold, L. Trevisan, S. P. Vadhan, Pseudorandom walks on regular digraphs and the RL vs. L problem. STOC 2006: 457-466

Doubly Efficient Interactive Proofs

Ron Rothblum

(joint work with Omer Reingold, Guy Rothblum)

The celebrated $\mathsf{IP} = \mathsf{PSPACE}$ Theorem [LFKN92, Sha92] allows an all-powerful but untrusted prover to convince a polynomial-time verifier of the validity of extremely complicated statements (as long as they can be evaluated using polynomial space). The interactive proof system designed for this purpose requires a polynomial number of communication rounds and an exponential-time (polynomial-space complete) prover.

In this paper [RRR16], we study the power of more efficient interactive proof systems. In particular, we seek proof-systems in which the prescribed prover strategy can be computed in polynomial-time, whereas the verifier can be computed even faster, say in linear time. These proof-systems, called *doubly-efficient interactive proofs*, were introduced by Goldwasser et. al [GKR08] who also constructed doubly efficient interactive proofs for all of NC (albeit with a large number of rounds).

In the talk we presented a result showing that for every statement that can be evaluated in polynomial time and bounded-polynomial space there exists an interactive proof that satisfies the following strict efficiency requirements: (1) the honest prover runs in polynomial time, (2) the verifier is almost linear time (and under some conditions even sub linear), and (3) the interaction consists of only

a *constant number of communication rounds*. Prior to this work, very little was known about the power of efficient, constant-round interactive proofs (rather than arguments). This result represents significant progress on the round complexity of interactive proofs (even if we ignore the running time of the honest prover), and on the expressive power of interactive proofs with polynomial-time honest prover (even if we ignore the round complexity). This result has several applications, and in particular it can be used for verifiable delegation of computation.

The new construction of the doubly interactive proof for bounded space leverages several new notions of interactive proofs, which may be of independent interest. One of these notions is that of *unambiguous interactive proofs* where the prover has a unique successful strategy. Another notion is that of *probabilistically checkable interactive proofs*[1] (PCIPs) where the verifier only reads a few bits of the transcript in checking the proof (this could be viewed as an interactive extension of PCPs).

Beyond the aforementioned works, we also refer the interested reader to Goldreich's recent survey [Gol18].

References

[BCS16] Eli Ben-Sasson, Alessandro Chiesa, and Nicholas Spooner. Interactive oracle proofs. In *Theory of Cryptography - 14th International Conference, TCC 2016-B, Beijing, China, October 31 - November 3, 2016, Proceedings, Part II*, pages 31–60, 2016.

[GKR08] Shafi Goldwasser, Yael Tauman Kalai, and Guy N. Rothblum. Delegating computation: interactive proofs for muggles. In *STOC*, pages 113–122, 2008.

[Gol18] Oded Goldreich. On doubly-efficient interactive proof systems. *Foundations and Trends in Theoretical Computer Science*, 13(3):158–246, 2018.

[LFKN92] Carsten Lund, Lance Fortnow, Howard J. Karloff, and Noam Nisan. Algebraic methods for interactive proof systems. *J. ACM*, 39(4):859–868, 1992.

[RRR16] Omer Reingold, Guy N. Rothblum, and Ron D. Rothblum Constant-round Interactive Proofs for Delegating Computation In *STOC*, pages 49–62, 2016.

[Sha92] Adi Shamir. IP = PSPACE. *J. ACM*, 39(4):869–877, 1992.

Distributed PCPs and applications

Aviad Rubinstein

The talk presents the notion of **Distributed PCP** introduced by [ARW17]. It is a new *distributed* model of probabilistically checkable proofs (PCP). A satisfying assignment $x \in \{0,1\}^n$ to a CNF formula φ is shared between two parties, where Alice knows $x_1, \ldots, x_{n/2}$, Bob knows $x_{n/2+1}, \ldots, x_n$, and both parties know φ. The goal is to have Alice and Bob jointly write a PCP that x satisfies φ, while exchanging little or no information. Unfortunately, this model as-is does not allow for nontrivial query complexity. Instead, we focus on a *non-deterministic* variant, where the players are helped by Merlin, a third party who knows all of x.

[1] An equivalent notion to PCIPs, called *interactive oracle proofs*, was recently introduced in an independent work of Ben-Sasson et. al [BCS16].

The talk surveys recent fine-grained and parametrized hardness of approximation results proved in this framework [ARW17, AR18, Rub18, KLM18, Che18, CW19, CGL+19, CM19].

A flagship application of this framework is the Bichromatic Closest Pair Problem. This problem can be instantiated with various similarity measures, for example inner product (over the reals):

Definition 1 (The Bichromatic Closest Pair Problem with Inner Product similarity measure (Max-IP))**.** *Given two sets A, B, each of N binary vectors in $\{0,1\}^d$, return a pair $(a,b) \in A \times B$ that maximizes the inner product $a \cdot b$.*

Thinking of the vectors as subsets of $[d]$, this Max-IP problem asks to find the pair with largest overlap, a natural similarity measure. A naïve algorithm solves the problem in $O(N^2 d)$ time, and one of the most-cited fine-grained results is a SETH[1] lower bound for this problem. Assuming SETH, we cannot solve Max-IP exactly in $N^{2-\varepsilon} \cdot 2^{o(d)}$ time, for any $\varepsilon > 0$ [Wil05]. For approximation, we use the Distributed PCP framework, to show that it is the problem remains hard for nearly-$\mathsf{poly}(N)$ factors.

The talk also presents some of the speaker's favourite open problems in this area, desribed below (based on [ARW17, Rub18]).

LCS Problem **(with two strings).** The Distributed PCP framework gives hardness of approximation for the Bichromatic Closest Pair Problem with Longest Common Subsequence (LCS) similarity measure [ARW17]. Simple gadgets constructed in a similar fashion (even without the Distributed PCP), can be combined together (along with some additional gadgets) into two long strings A, B of length m, in a way that yields a reduction from SETH to computing the longest common subsequence (LCS) of (A, B), ruling out *exact* algorithms in $O(n^{2-\varepsilon})$ [AWW14, ABV15, BI15, BK15, AHWW16]. However, in the instances output by this reduction, approximating the value of the LCS reduces to approximating the *fraction* of assignments that satisfy the original formula; it is easy to obtain a good additive approximation by sampling random assignments. The recent works of [AR18, CGL+19] mentioned above, combines Distributed PCP with complexity assumptions on deterministic algorithms (the latter building on [AB17]) to tackle this issue, but their ideas do not seem to generalize to randomized algorithms.

Open Question 2. *Is there a* 1.1*-approximation for* LCS *running in $O(n^{2-\varepsilon})$ time, for some $\varepsilon > 0$? (Open for all alphabet sizes.)*

The triangle inequality barrier. Consider a naive gadget reduction for the Bichromatic Closest Pair Problem where we construct a vector for each half assignment. Let $\alpha_1, \alpha_2 \in \{0,1\}^{n/2}$ be partial assignments to the first half of the variables, and $\beta_1, \beta_2 \in \{0,1\}^{n/2}$ for the second half. Suppose that $(\alpha_1; \beta_1), (\alpha_2; \beta_1), (\alpha_2; \beta_2)$ satisfy the formula, but (α_1, β_2) does not. Let

[1]The Strong Exponential Time Hypothesis (SETH) postulates that for every ε there is a $k = k(\varepsilon)$ such that k-SAT over n variables requires $(2-\varepsilon)^n$ time.

$a^{\alpha_1}, a^{\alpha_2}, b^{\beta_1}, b^{\beta_2}$ be the corresponding vectors. Then, if our reduction has completeness c and soundness s, we would like to have

$$\left\|a^{\alpha_1} - b^{\beta_2}\right\| \geq s \geq 3c \geq \left\|a^{\alpha_1} - b^{\beta_1}\right\| + \left\|a^{\alpha_2} - b^{\beta_2}\right\| + \left\|a^{\alpha_2} - b^{\beta_1}\right\|.$$

But that would violate the triangle inequality. Note that this restricts our ability to prove stronger hardness of approximation even for more complicated metrics like edit distance. It is also important to remark that the reductions based on Distributed PCP (including the ones in this talk) do not exactly fall into this naive gadget reduction framework; nevertheless it is not at all clear that they can overcome this obstacle.

Open Question 3 (3-approximation)**.** *Prove that, assuming SETH and for some constant $\varepsilon > 0$, approximating* Bichromatic Closest Pair *with Euclidean metric to within factor* 3 *requires time* $\Omega\left(N^{1+\varepsilon}\right)$.

References

[AB17] Amir Abboud and Arturs Backurs. Towards hardness of approximation for polynomial time problems. In *8th Innovations in Theoretical Computer Science Conference, ITCS 2017, January 9-11, 2017, Berkeley, CA, USA*, pages 11:1–11:26, 2017.

[ABV15] Amir Abboud, Arturs Backurs, and Virginia Vassilevska Williams. Tight hardness results for LCS and other sequence similarity measures. In *Proc. of the 56th FOCS*, pages 59–78, 2015.

[AHWW16] Amir Abboud, Thomas Dueholm Hansen, Virginia Vassilevska Williams, and Ryan Williams. Simulating branching programs with edit distance and friends: or: a polylog shaved is a lower bound made. In *Proc. of the 48th STOC*, pages 375–388, 2016.

[AR18] Amir Abboud and Aviad Rubinstein. Fast and deterministic constant factor approximation algorithms for LCS imply new circuit lower bounds. In *9th Innovations in Theoretical Computer Science Conference, ITCS 2018, January 11-14, 2018, Cambridge, MA, USA*, pages 35:1–35:14, 2018.

[ARW17] Amir Abboud, Aviad Rubinstein, and R. Ryan Williams. Distributed PCP theorems for hardness of approximation in P. In *58th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2017, Berkeley, CA, USA, October 15-17, 2017*, pages 25–36, 2017.

[AWW14] Amir Abboud, Virginia Vassilevska Williams, and Oren Weimann. Consequences of faster alignment of sequences. In *Proc. of the 41st ICALP*, pages 39–51, 2014.

[BI15] Arturs Backurs and Piotr Indyk. Edit Distance Cannot Be Computed in Strongly Subquadratic Time (unless SETH is false). In *Proc. of the 47th Annual ACM SIGACT Symposium on Theory of Computing (STOC)*, pages 51–58, 2015.

[BK15] Karl Bringmann and Marvin Kunnemann. Quadratic conditional lower bounds for string problems and dynamic time warping. In *Proc. of the 56th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 79–97, 2015.

[CGL+19] Lijie Chen, Shafi Goldwasser, Kaifeng Lyu, Guy N. Rothblum, and Aviad Rubinstein. Fine-grained complexity meets ip = pspace. In *SODA*, 2019. To appear.

[Che18] Lijie Chen. On the hardness of approximate and exact (bichromatic) maximum inner product. *Electronic Colloquium on Computational Complexity (ECCC)*, 25:26, 2018.

[CM19] Karthik C.S. and Pasin Manurangsi. Closest pair in euclidean metric: Monochromatic is as hard as bichromatic. In *ITCS*, 2019. To appear.

[CW19] Lijie Chen and Ryan Williams. An equivalence class for orthogonal vectors. In *SODA*, 2019. To appear.

[KLM18] Karthik C. S., Bundit Laekhanukit, and Pasin Manurangsi. On the parameterized complexity of approximating dominating set. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2018, Los Angeles, CA, USA, June 25-29, 2018*, pages 1283–1296, 2018.

[Rub18] Aviad Rubinstein. Hardness of approximate nearest neighbor search. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2018, Los Angeles, CA, USA, June 25-29, 2018*, pages 1260–1268, 2018.

[Wil05] R. Ryan Williams. A new algorithm for optimal 2-constraint satisfaction and its implications. *Theoretical Computer Science*, 348(2–3):357–365, 2005.

Free Probability, Interlacing Polynomials, and Expander Graphs

Nikhil Srivastava

(joint work with Adam Marcus, Daniel Spielman)

1. Expanders and T_d

We give a gentle introduction to Free Probability theory in the context of a very concrete problem of interest in computer science, namely constructions of expander graphs. We are interested in the following extremal question: for a fixed degree d, what is the largest possible spectral gap attainable by an infinite sequence of $d-$regular graphs? The Alon-Boppana bound gives a limit to how large it can be, stating that if A is the adjacency matrix of a $d-$regular graph and $d = \lambda_1, \lambda_2, \ldots, \lambda_n$ are its eigenvalues, then

$$\lambda_2(A) \geq 2\sqrt{d-1} - o(1),$$

where the $o(1)$ term goes to zero as the diameter tends to infinity. The number $2\sqrt{d-1}$ has a conceptual meaning: it is the spectral radius of the (infinite) adjacency matrix of the infinite $d-$regular tree T_d. A finite graph with $\lambda_2(A) \leq 2\sqrt{d-1}$ is called a one-sided Ramanujan graph, and a graph with $|\lambda_i(A)| \leq 2\sqrt{d-1}$ for all $i \neq 1$ is called a Ramanujan graph.

There are at present two kinds of known constructions of Ramanujan graphs. The first, due to Lubotzky-Phillips-Sarnak and Margulis [LPS88], are Cayley graphs of certain nonabelian groups, analyzed using number theory. This proof technique is able to show that infinite sequences of Ramanujan graphs for $d = p^k+1$ where p is a prime. The second [MSS15b], is able to show the existence of one-sided Ramanujan graphs for every $d \geq 3$ and even integer n; see also [HPS18, Coh16].

We will explain the second construction, which begins by considering random graphs. Let G be a union of d i.i.d. uniformly random perfect matchings on n vertices, where n is even. Then, if M is the adjacency matrix of a single matching, the adjacency matrix of G is given by

$$A = \sum_{i \leq d} P_i M P_i^T,$$

where the P_i are i.i.d. uniform permutations. Then G is regular so $\lambda_1(A) = d$, which is also the sum of the top eigenvalues of the summands $P_i M P_i$, which are all equal to one. The relationship between the rest of the eigenvalues of A and those of $P_i M P_i^T$ is much more complicated; note that M itself has a very simple spectrum with distribution $\frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1$.

Surprisingly, it turns out that in the case of T_d itself, the relationship between the spectra can be understood precisely. To do this we need to introduce a new notion of convolution of probability measures, which we do in the next section.

2. Three Convolutions

Let X and Y be two independent random variables, each uniformly distributed on a set of n real numbers. Independence implies that the moments of $X + Y$ are functions of the moments of X and the moments of Y individually. Letting m_k denote the k^{th} moment of a random variable, and $\mathbf{m}_k = (m_1, \ldots, m_k)$ a tuple of moments up to k, we have for independent X and Y:

$$m_k(X+Y) = \sum_{j \le k} \binom{k}{j} m_j(X) m_{k-j}(Y) =: P_k(\mathbf{m}_k(X), \mathbf{m}_k(Y)),$$

where P is an explicit polynomial depending only on k. Thus, the polynomials P_k effectively *define* what convolution of discrete probability measures is.

We now write this as a statement about random matrices. Let A and B be diagonal $n \times n$ matrics containing the supports of X and Y respectively, and let A_c denote the $cn \times cn$ matrix containing c copies of A on the diagonal. Then it is an easy exercise to see that for every positive integer k:

$$\mathbb{E}_P \frac{1}{cn} \mathrm{Tr}(A_c + P B_c P^T)^k = P_k(\mathbf{m}_k(X), \mathbf{m}_k(Y)) + o_c(1), \tag{1}$$

where the expectation is over a uniformly random permutation on cn elements. Thus, in the limit as $c \to \infty$, the moments of these random matrices specify classical convolution.

It turns out that if one replaces the random permutations by random orthogonal matrices Q, then the limit still exists, and gives an alternate notion of convolution. In particular, we have for every positive integer k:

$$\mathbb{E}_Q \frac{1}{cn} \mathrm{Tr}(A_c + Q B_c Q^T)^k = Q_k(\mathbf{m}_k(X), \mathbf{m}_k(Y)) + o_c(1), \tag{2}$$

for some different polynomials Q_k which again depend only on k. Voiculescu [Voi00] defined a new convolution in terms of these Q_k.

Definition. The *free convolution* $\mu_X \boxplus \mu_Y$ of two discrete measures on n atoms is the unique measure with moments given by $Q_k(\mathbf{m}_k(X), \mathbf{m}_k(Y))$ for all k.

The operation $\boxplus$ is commutative and associative. The reason we care about this is that the free convolution allows us to understand the spectrum of the T_d. The adjacency matrix A_∞ of this tree may be written as a sum of d adjacency matrices of matchings, $A_\infty = M_1 + \ldots + M_d$. Defining what spectral measure means in the infinite case would take us too far afield, but let us accept that

the tree has a spectral measure μ_d and the matchings all have the same spectral measure $\mu_1 := \frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1$, which is the same as in the finite case. The punch line is that $\mu_d = \mu_1 \boxplus \mu_1 \boxplus \ldots \boxplus \mu_1$ (d times), so the free convolution expresses the (complicated) spectrum of the tree in terms of the (simple) spectra of the matchings. In fact, this is essentially why it is called the free convolution — for even d the T_d is the Cayley graph of the free group on $d/2$ generators. The relation of moments prescribed by the polynomials Q_k is referred to as *free independence*.

Unfortunately, free independence is not realizable for finite matrices. To get around this, we consider a third convolution which works at the finite level. Namely, given X, Y, A, B as above, consider the degree n polynomial $p_{A,B} := \mathbb{E}_Q \chi(A + QBQ^T)$, where Q is random orthogonal and $\chi(M) = \det(zI - M)$ is the characteristic polynomial. It can be shown that the roots of $p_{A,B}$ are all real whenever A and B are real.

Definition The *finite free convolution* of μ_X and μ_Y, denoted $\mu_X \boxplus_n \mu_Y$, is the uniform measure on the roots of $p_{A,B}$.

Moreover, one can show [MSS15] that $\boxplus_n$ is dominated by $\boxplus$ in the following sense.

Theorem. for every μ_X and μ_Y supported on n points, it holds that

$$\lambda_{max}(\mu_X \boxplus_n \mu_Y) < \lambda_{max}(\mu_X \boxplus \mu_Y).$$

where λ_{max} is the largest point in the support of a measure, and this generalizes to a convolution of any number of measures.

In particular, we have for our case of interest:

$$\lambda_{max}(\mu_1^{\boxplus_n d}) < \lambda_{max}(\mu_1^{\boxplus d}) = \lambda_{max}(\mu_d) = 2\sqrt{d-1}.$$

3. Ramanujan Graphs

We now use finite free convolutions to show that Ramanujan graphs exist. The first step is to show that with nonzero probability,

$$\lambda_2(A) < \lambda_2(\mathbb{E}\chi(A)),$$

where λ_2 refers to the second largest root on the right. This is done by an "interlacing families" argument [MSS15b], which is not the focus of this note. Thus, to show the existence of one-sided Ramanujan graphs, we merely need to show that the polynomial $p_d := \mathbb{E}\chi\left(\sum_{i\le d} P_i M P_i^T\right)$ has second root bounded by $2\sqrt{d-1}$. Now a miracle happens: it turns out that it is possible to replace the random permutations $P_1, \ldots, P_d$ by random orthogonal matrices $Q_1, \ldots, Q_d$, conditioned to fix the all ones vector (i.e., $Q_i 1 = 1$ for all i, since this is what permutations do) *without changing the average*. This may be seen as a sort of invariance principle and is a consequence of the fact that the determinant is multilinear so its restrictions are low degree polynomials. After isolating the trivial root at d, we therefore have: $p_d = (x-d)\mathbb{E}\chi\left(\sum_{i\le d} Q_i M' Q_i^T\right)$, where M' is a matching with the trivial eigenspace at 1 removed and the Q_i are now fully random orthogonal

matrices. A short argument shows that we may as well assume M' has spectral measure μ_1, and we conclude that:

$$\lambda_2(p_d) \leq \lambda_{max}(\mu_1^{\boxplus_n d}) < \lambda_{max}(\mu_1^{\boxplus d}) = 2\sqrt{d-1},$$

as desired.

References

[Coh16] Michael B Cohen, *Ramanujan graphs in polynomial time*, Foundations of Computer Science (FOCS), 2016 IEEE 57th Annual Symposium on, IEEE, 2016, pp. 276–281.

[HPS18] Chris Hall, Doron Puder, and William F Sawin, *Ramanujan coverings of graphs*, Advances in Mathematics **323** (2018), 367–410.

[LPS88] Alexander Lubotzky, Ralph Phillips, and Peter Sarnak, *Ramanujan graphs*, Combinatorica **8** (1988), no. 3, 261–277.

[MSS15] Adam Marcus, Daniel A Spielman, and Nikhil Srivastava, *Finite free convolutions of polynomials*, arXiv preprint arXiv:1504.00350 (2015).

[MSS15b] Adam W Marcus, Daniel A Spielman, and Nikhil Srivastava, *Interlacing families iv: Bipartite ramanujan graphs of all sizes*, Foundations of Computer Science (FOCS), 2015 IEEE 56th Annual Symposium on, IEEE, 2015, pp. 1358–1377.

[Voi00] Dan Voiculescu, *Lectures on free probability*, Lectures Notes in Mathematics **1738** (2000).

Parity Samplers and Explicit, Epsilon-Balanced Codes Close to the GV Bound

Amnon Ta-Shma

The question of finding an epsilon-biased set with close to optimal support size, or, equivalently, finding an explicit binary code with distance $\frac{1-\epsilon}{2}$ and rate close to the Gilbert-Varshamov bound, attracted a lot of attention in recent decades. In this talk we present a solution of an explicit ϵ-biased set over k bits with support size $O(\frac{k}{\epsilon^{2+o(1)}})$. This improves upon all previous explicit constructions. The result is close to the Gilbert-Varshamov bound which is $O(\frac{k}{\epsilon^2})$ and the lower bound which is $\Omega(\frac{k}{\epsilon^2 \log \frac{1}{\epsilon}})$.

The main technical tool we use is bias amplification with the s-wide replacement product. The sum of two independent samples from an ϵ-biased set is ϵ^2 biased. Previous work showed how to amplify the bias more economically by choosing two samples from a walk over an expander. We show that amplification with a random walk over the s-wide replacement product reduces the bias almost optimally. We abstract this approach in two ways:

- First we define parity samplers and show how they can be used to amplify the code distance, and,
- Second, we define memory augmented random walks and show how they can give better explicit parity samplers.

An $[n, k, d]_2$ code is a subspace $C \subseteq \mathbb{F}_2^n$ of dimension k such that the distance between any two codewords is at least d. The Gilbert-Varshamov (GV) bound states that for every $\delta > 0$, there exists a family $\{C_n\}$ of linear codes

such that C_n has length n, relative distance δ and relative rate $1 - H(\delta) - o(1)$, where H is Shannon's entropy function. The GV bound proves the existence of a good family of codes, but does not show any specific such family. We construct a fully explicit binary code approaching the GV bound for the important parameter regime where δ is close to half. For this parameter regime, the Gilbert-Varshamov bound shows that random linear codes of length $n = O(k/\varepsilon^2)$ are w.h.p. $[n, k, \frac{1-\varepsilon}{2}n]_2$ codes. The LP bound gives an almost matching lower bound showing that $n = \Omega(\min\left\{\frac{k}{\varepsilon^2 \log \frac{1}{\varepsilon}}, 2^k\right\})$ (see, e.g., [AGH+92][Section 7]). There were many attempts trying to achieve the GV bound explicitly, and, roughly speaking, they can be divided into two main approaches:

- Using concatenation: Concatenating Reed Solomon with Hadamard gives one of the constructions in AGHP with $n = O(\frac{k^2}{\varepsilon^2})$[AGH+92]. Concatenating Reed Solomon with the Wozencraft ensemble gives the Justensen code [J72] with constant relative distance and constant relative rate. Starting with an AG code and concatenating with Hadamard over a field of size $O(\frac{1}{\varepsilon^2})$ gives a construction with about $O(\frac{k}{\varepsilon^3})$ support size. Taking the Hermitian code below the genus over a smaller field and concatenating with Hadamard gives the construction in [BT13] with support size $O((\frac{k}{\varepsilon^2})^{5/4})$. All of the above bounds fall short of achieving the GV bound which is $O(\frac{k}{\varepsilon^2})$, and this is due, in part, to the expensive concatenation step. The attainable parameters when concatenating with Reed Solomon as the outer code are captured by the Zyablov bound, which for distance $\frac{1-\varepsilon}{2}$ gives code length $\Theta(\frac{k}{\varepsilon^3})$ (see [ABN+92, Section 1]). Similarly, concatenating with any high genus AG code cannot attain the GV bound for distance close to half [BT11, Section 4].
- Distance amplification: A second approach was first suggested by Naor and Naor [NN93], and later Alon et al. [ABN+92]. The idea is to start with a binary error-correcting code that has moderate distance (say, some constant relative distance and constant relative rate) and *amplify* it to a binary error correcting code with a higher relative distance (say, close to half). This was done by Naor and Naor [NN93], and later Alon et al. [ABN+92], using expanders, or more generally dispersers. This approach also gives binary error correcting codes of length n, dimension k and distance $\frac{1-\varepsilon}{2}$ with $n = O(\frac{k}{\varepsilon^3})$. Nevertheless, Alon et al. show that for a certain non-binary field size their construction lies above the Zyablov bound.

We show:

Theorem 1. *For every $k = k(n)$ and $\varepsilon > 0$ there exists a fully explicit family of linear codes $\{C_n\}$ such that C_n is an $[n, k, d = \frac{1-\varepsilon}{2}n]_2$ code and $n = O(\frac{k}{\varepsilon^{2+\phi(\varepsilon)}})$ where $\phi(\varepsilon) = O((\frac{\log\log\frac{1}{\varepsilon}}{\log\frac{1}{\varepsilon}})^{1/3})$. Notice that $\lim_{\varepsilon\to 0}\phi(\varepsilon) = 0$.*

We remark the simple corollary that for every constant $\alpha > 0$ there exists a fully explicit family of $[n = O(\frac{k}{\varepsilon^{2+\alpha}}), k, \frac{1-\varepsilon}{2}]_2$ codes, where the constant in the

bigO notation may depend on α. This is because either $\alpha \geq \phi(\varepsilon)$ and then we can use the construction of Theorem 1, or $\alpha \leq \phi(\varepsilon)$ in which case ε is a constant and then we can use previous constructions (e.g., [NN93]) and the dependence on ε may be absorbed into the BigO notation.

We now strengthen the notion of an error correcting code to that of a *balanced* code:

Definition 2. *Let $c \in \mathbb{F}_2^n$. $Bias(c) = \frac{1}{n}|\sum_{i=1}^{n}(-1)^{c_i}|$, i.e., the difference between the fraction of 0 symbols and 1 symbols in c (in absolute value). We say c is ε balanced if $Bias(c) \leq \varepsilon$. We say an $[n,k]_2$ code is ε–balanced if every non-zero codeword of C is ε balanced. We also say C is a $[n,k,\frac{1\pm\varepsilon}{2}n]_2$ code.*

We now abstract a general way of amplifying the distance of a linear error correcting code in a black-box way:

Definition 3. *(Sampler) A bipartite multi-graph $\Phi = (L,R,F)$ is a* sampler*, where each element $a \in L$ corresponds to the multi-set of all its neighbours in R (with multiplicities). The* degree *of the sampler is the maximal left-degree of the bipartite graph. The* size *of the sampler is $|L|$.*

We now define *parity samplers*. Given a bipartite graph $\Phi = (L,R,F)$ we think of Φ as inducing a linear map $\Phi : \mathbb{F}_2^R \to \mathbb{F}_2^L$ as follows. Given $f : R \to \mathbb{F}_2$, we think of f as attaching the label $f_2(b) \in \{0,1\}$ to the vertex $b \in R$. We then define a function $g = \Phi(f) : L \to \mathbb{F}_2$, by letting $g(a)$ be the parity (or sum over $\mathbb{F}_2$) of all a's neighbours. Formally,

Definition 4. *A bipartite multi-graph $\Phi = (L,R,F)$ induces a mapping $\Phi : \mathbb{F}_2^R \to \mathbb{F}_2^L$ in the following way. For every $f : R \to \mathbb{F}_2$ and $a \in L$,*

$$\Phi(f)(a) \quad = \quad \sum_{b:(a,b)\in F} f(b).$$

We say Φ is a $(n,\varepsilon_0) \to (n',\varepsilon)$ parity sampler if $|L| = n'$, $|R| = n$ and Φ maps any ε_0-balanced string $f \in \mathbb{F}_2^n$ to an ε-balanced string $\Phi(f) \in \mathbb{F}_2^{n'}$.

Next we notice that we can combine an ε_0 balanced code, and an $(n,\varepsilon_0) \to (n',\varepsilon)$ parity sampler, to get a new more balanced code. Specifically, suppose: C is an $[n,k,d=\frac{1-\varepsilon_0}{2}]_2$ code, where the i'th coordinate of the code (for $i \in [n]$) computes the linear function $\ell_i : \mathbb{F}_2^k \to \mathbb{F}_2$ and $\Phi = (L,R,F)$ is a $(n,\varepsilon_0) \to (n',\varepsilon)$ parity sampler. Define a new $[n',k]_2$ code $C' = \phi(C)$ where the a'th coordinate of C' (for $a \in [n']$) is the linear function $\ell'_a : \mathbb{F}_2^k \to \mathbb{F}_2$ defined by $\ell'_a(x) = \sum_{i\in[n]:(a,i)\in F} \ell_i(x)$, where addition is over $\mathbb{F}_2$. We claim:

Lemma 5. *If C is an $[n,k,d=\frac{1-\varepsilon_0}{2}]_2$ code and $\Phi = (L,R,F)$ is a $(n,\varepsilon_0) \to (n',\varepsilon)$ parity sampler, then $C' = \Phi(C)$ is an $[n',k,\frac{1\pm\varepsilon}{2}n']_2$ code. Furthermore, if C and Φ are explicit (resp. fully explicit) then so does C'.*

Our base code is the Justensen code which is asymptotically good and has $\varepsilon_0 \leq 0.8$. Our construction follows by constructing a good explicit parity sampler. Our starting point is the remarkable fact that a random walk over an expander

graph is a good parity sampler. This was observed by Noga Alon (but was not published) and was found independently by us. We prove:

Theorem 6. *Suppose $G = (V, E)$ is an (n, D, λ) expander. Let $\varepsilon_0 > 0$. Φ_G is a $(n, \varepsilon_0) \to (nD^t, \varepsilon)$ parity sampler for $\varepsilon = (\varepsilon_0 + 2\lambda)^{\lfloor t/2 \rfloor}$.*

This gives an ε-balanced code of rate about $O(\varepsilon^4)$. We improve upon this by taking a memory augmented random walk:

Definition 7. *(Memory augmented random walk). Let $G = (V, E)$ be an undirected D regular graph, X an auxiliary state, $f : X \to [D]$ and $g : X \times [D_2] \to X$ functions. A walk on G starting at vertex $v_0 \in V$, auxiliary state $x_0 \in X$ and instructions $\sigma_1, \ldots, \sigma_t \in [D_2]$ does the following: At step i, for $i = 1, \ldots, t$:*

- *$v_i = v_{i-1}[f(x_{i-1})]$, i.e., we walk on G from v_{i-1} according to $f(x_{i-1}) \in [D]$, and,*
- *$x_i = g(x_{i-1}, \sigma_i)$, i.e., we update the auxiliary state using the instruction σ_i.*

A memory augmented random walk walks on a degree D graph, but at each time step uses only $D_2 \ll D$ edges going out of the current vertex, where the choice of which D_2 edges are used depends on the current auxilary state.

This specific memory augmented random walk that we take is based on a related construction from [BT11]. It uses an auxilary expander graph $H = (X = [D^w], E_2)$ of degree $D_2 = \sqrt{D}$ and the functions

- $f : X \to [D]$ that views $x \in X$ as a vector of dimension w over $[D]$ and returns the projection on the first coordinate, and,
- The function $g : X \times [D_2] \to X$ that takes a step on H according to the instruction σ and then cyclically rotates the vector.

This may be seen as extending the replacement product of [RVW00] to a memory augmented random walk. Using this memory augmented random walk we get our explicit parity sampler:

Theorem 8. *Let $\varepsilon_0 > 0$. Define the bipartite graph that for every length t path in the above memory augmented random walk with width $w = w(1)$, connects the path with all the vertices on the path. Then this bipartite graph is a $(n, \varepsilon_0) \to (nD^t, \varepsilon)$ parity sampler for $\varepsilon = (\varepsilon_0 + 2\lambda)^{(1-o(1)t}$.*

Finally, this explicit parity sampler immediately gives an explicit binary code with distance $\frac{1}{2} - \varepsilon$ and rate about ε^2 as desired.

References

[ABN+92] N. Alon, J. Bruck, J. Naor, M. Naor, and R. Roth. Construction of asymptotically good, low-rate error-correcting codes through pseudo-random graphs. *IEEE Transactions on Information Theory*, 38:509–516, 1992.

[AGH+92] N. Alon, O. Goldreich, J. Hastad, and R. Peralta. Simple constructions of almost k–wise independent random variables. *Random Structures and Algorithms*, 3(3):289–303, 1992.

[BT11] A. Ben-Aroya and A. Ta-Shma. A combinatorial construction of almost-ramanujan graphs using the zig-zag product. *SIAM Journal on Computing*, 40(2):267–290, 2011.

[BT13] A. Ben-Aroya and A. Ta-Shma. Constructing small-bias sets from algebraic-geometric codes. *Theory of Computing*, 9(5):253–272, 2013.

[J72] J. Justesen. Class of constructive asymptotically good algebraic codes. *IEEE Transactions on Information Theory*, 18(5):652–656, 1972.

[NN93] J. Naor and M. Naor. Small–bias probability spaces: Efficient constructions and applications. *SIAM Journal on Computing*, 22(4):838–856, 1993.

[RVW00] O. Reingold, S. Vadhan, and A. Wigderson. Entropy waves, the zig-zag graph product, and new constant-degree expanders. *Annals of Mathematics*, 155(1):157–187, 2002.

Classical Verification of Quantum Computations

Thomas Vidick

Quantum computing enthusiasts hope to soon reach a stage where engineered devices based on the laws of quantum mechanics are able to implement computations that can no longer be emulated on a classical computer. Once that stage is reached, will it be possible to verify the results of the quantum device?

Recently Mahadev [Mah18] introduced a solution to the following informally stated problem: given black-box access to a quantum device, i.e. given only the ability to generate classical instructions and obtain classical readout information in return, is it possible to delegate a quantum computation to the device in a way that the outcome obtained can be verified on a classical computer — even when the quantum device may be faulty or even adversarially designed to fool the verification procedure?

The question has a long history, that prior to Mahadev's work had resulted in partial answers for different models of verification. Some of the most important results include the concurrent works of Aharonov et al. [ABE08] and Broadbent et al. [BFK08] showing how to achieve verification in a model where the verification procedure itself has access to a small, trusted quantum computer, and the work of Reichardt et al. [RUV13] in a model where the verification procedure is entirely classical, but has access to two spatially isolated quantum computers, sharing entanglement, whose implementation of a quantum computation it aims to verify. In contrast to Mahadev's result, these works achieve information-theoretic (instead of computational) soundness guarantees in their respective models.

The most novel element of Mahadev's solution is an ingenious use of classical cryptographic techniques to tie a "cryptographic leash" around the quantum device. In this talk I present Mahadev's result, focusing on the connection with complexity theory and the use of techniques from classical cryptography. The main result can be stated as follows.

Theorem 1. *For any language L in BQP, there is a four-message argument system for L with a classical polynomial-time verifier that has the following properties:*

- *There is a quantum polynomial-time prover P such that for any $x \in L$, on input x the verifier accepts its interaction with P with probability $1 - O(2^{-|x|})$.*
- *Assuming the hardness of the learning with errors problem (LWE) against quantum polynomial-time attacks, for any $x \notin L$ no quantum polynomial-time prover can convince the verifier to accept on input x with probability larger than $1 - \Omega(1/poly(|x|))$.*

We note that the theorem requires hardness of LWE for sub-exponential noise ratio. The soundness parameter can be amplified by straightforward sequential repetition.

The proof of the theorem has two main steps. The first step, essentially due to Kitaev, with further refinements by Cubitt and Montanaro [CM16], reduces any quantum polynomial-time computation to a decision problem about the smallest eigenvalue of a well-structured local Hamiltonian that can be efficiently computed from the circuit. Informally, this step is a quantum analogue of a formulation of the Cook-Levin Theorem that would reduce the verification of correctness of the tableau of a classical computation to the verification that a certain instance of the MAX-CUT constraint satisfaction problem computed from the tableau has a large enough cut.

The second step is the key contribution on Mahadev, and I will focus on it in the talk. It is based on the introduction of a new cryptographic primitive, a "trap-door claw-free function pair", and using that primitive to devise a quantum qubit commitment scheme that is computationally binding with respect to measurement outcomes on the qubit in the computational or Hadamard bases.

References

[ABE08] Dorit Aharonov, Micahel Ben-Or, and Elad Eban, *Interactive Proofs For Quantum Computations*, Arxiv preprint arXiv:0810.5375 (2008).

[BFK08] Anne Broadbent, Joseph F. Fitzsimons, and Elham Kashefi, *Universal blind quantum computation*, Arxiv preprint arXiv:0807.4154 (2008).

[CM16] Toby Cubitt and Ashley Montanaro, *Complexity classification of local hamiltonian problems*, SIAM Journal on Computing **45** (2016), no. 2, 268–316.

[Mah18] Urmila Mahadev, *Classical verification of quantum computations*, arXiv preprint arXiv:1804.01082 (2018).

[RUV13] Ben W Reichardt, Falk Unger, and Umesh Vazirani, *Classical command of quantum systems*, Nature **496** (2013), no. 7446, 456.

Invariant theory - a gentle introduction for computer scientists (optimization and complexity)

Avi Wigderson
(joint work with Ankit Garg)

Invariant theory deals with understanding the symmetries of mathematical objects, namely transformations of the underlying space which leave an object unchanged or invariant. Here we only touch on some of its main objects, problems and results. This study, of what *does not* change in certain dynamic processes (namely group actions), interacts with many mathematical fields including group theory, commutative algebra and algebraic geometry, and is a central to modern physics. We will stress some of the many facets which interact with computational complexity and optimization.

Let G be a group which acts *linearly* on a vector space V. [1] That is $g\cdot(v_1+v_2) = g\cdot v_1 + g\cdot v_2$ and $g\cdot(cv) = cg\cdot v$. We assume $V=\mathbb{F}^d$ (and the field $\mathbb{F}$ is usually $\mathbb{C}$). Hence G acts naturally on $\mathbb{F}[x_1,\ldots,x_d]$.

Invariant Polynomials. Invariant polynomials are polynomial functions on V left invariant by the action of G. One simple example is

- Group $G = SL_n(\mathbb{C})\times SL_n(\mathbb{C})$ acts on $V = M_n(\mathbb{C}) = \mathbb{C}^{n\times n}$ (dimension $d=n^2$ here) by a change of bases of the rows and columns, namely left-right multiplication: that is, (A,B) maps X to AXB. Here, $\det(X)$ is an invariant polynomial and in fact every invariant polynomial must be a univariate polynomial in $\det(X)$.

The above phenomenon that the ring of invariant of polynomials is generated by a finite number of invariant polynomials is not a coincidence. The *finite generation theorem* due to Hilbert [Hil90, Hil93] states that, for a large class of groups (including the groups mentioned above), the invariant ring must be finitely generated. These two papers of Hilbert are highly influential and laid the foundations of commutative algebra. In particular, "finite basis theorem" and "Nullstellansatz" were proved as "lemmas" on the way towards proving the finite generation theorem!

Orbits and Orbit-Closures. The *orbit* of a vector $v\in V$ is the set of all vectors obtained by the action of G on v. The *orbit-closure* of v is the closure (under the Euclidean topology) of the orbit of v. Many fundamental problems in theoretical computer science (and many more across mathematics) can be phrased as questions about such orbits and orbit-closures. Here are some familiar examples:

- Graph isomorphism problem can be phrased as checking if the orbits of two graphs are the same or not, under the action of the symmetric group permuting the vertices.

[1] Invariant theory is nicest when the underlying field is $\mathbb{C}$ and the group G is either finite, the general linear group $GL_n(\mathbb{C})$ (or its Abelian diagonal subgroup), the special linear group $SL_n(\mathbb{C})$ (or its Abelian diagonal subgroup), or a direct product of these groups (more generally a complex algebraic reductive group).

- Geometric complexity theory (GCT) [Bür12] formulates a variant of $\mathcal{VP}$ vs. $\mathcal{VNP}$ question as checking if the (padded) permanent lies in the orbit-closure of the determinant (of an appropriate size), under the action of the general linear group on polynomials induced by its natural linear action on the variables.

Furthermore, it turns out that even the simplest concept involving orbit-closures, namely the *null cone*, already captures a lot of interesting problems. A vector v is in the null cone if 0 lies in the orbit-closure of v. For different group actions, the null cone captures concepts like nilpotency, singularity, bipartite matching, linear programming and non-commutative rational identity testing. So even the most basic question (from an invariant theoretic perspective) of testing if a vector is in the null cone is already extremely interesting from a computational point of view. And in fact, a sub-field of invariant theory, geometric invariant theory, already provides non-trivial computational perspectives on the null cone.

Geometric invariant theory provides a geometric and analytic viewpoint to invariant theory. It morally puts the null cone membership problem in NP $\cap$ coNP. The Hilbert-Mumford criterion [Hil93, Mum65] provides a way to certify if a vector is in the null cone and the Kempf-Ness theorem [KN79] provides a way to certify if a vector is not in the null cone. We elaborate a bit on the Kempf-Ness theorem below.

Given a vector $v \in V$, consider the optimization problem which finds *a vector of minimum ℓ_2-norm* in the orbit-closure of v:

$$N(v) = \inf_{g \in G} \|g \cdot v\|_2^2 \tag{1}$$

It is easy to see that v is in the null cone iff $N(v) = 0$. For non-commutative group actions (think of $G = GL_n(\mathbb{C})$ for concreteness), the function $f_v(g) = \|g \cdot v\|_2^2$ is not convex in the Euclidean geometry but is geodescially convex (e.g. see [Woo11]). A consequence of geodesic convexity is the Kempf-Ness theorem [KN79], that states that any critical point (i.e., point with zero gradient) of $f_v(g)$ must be a global minimum. This brings us to *moment maps*.

Informally, the moment map $\mu_G(v)$ is the gradient of $f_v(g)$ at $g = id$, the identity element of G. The Kempf-Ness theorem draws the following beautiful connection between the moment map and $N(v)$. It is a duality theorem which, along with the Hilbert-Mumford criterion, greatly generalizes linear programming duality to a "non-commutative" setting.

Theorem 1 ([KN79])**.** *Fix an action of group G on a vector space V and let $v \in V$. v is not in the null cone iff there exists a non-zero w in the orbit-closure of v s.t. $\mu_G(w) = 0$.*

The Kempf-Ness theorem provides an optimization approach towards the null cone membership problem through norm-squared minimization. For special group actions, this optimization problem has been well studied, e.g. matrix balancing, matrix scaling [Sin64], geometric programming, to mention a few. Recently there has been a surge of activity on designing scaling algorithms for minimizing the

norm-squared function for an even more general class of group actions and this has resulted in algorithms for diverse problems such as non-commutative rational identity testing [GGOW16], testing feasibility of Brascamp-Lieb inequalities [GGOW17], tensor scaling [BGOWW18], orbit-closure intersection for the left-right action [AGLOW18] and the one-body quantum marginal problem [BFGOWW18]. There is an interesting interplay of analysis and algebra here: these are analytic algorithms for algebraic problems but at the same time, the algebraic nature of the problem plays an important role in the analysis of the algorithms. The potential functions are typically invariants under the action of certain subgroups of the acting group.

Some of these analytic algorithms yield determinsitic algorithms for subclasses of the polynomial identity testing (PIT) problem which are of a different nature than those typically studied in the literature (namely restricted circuit classes). One example is from [GGOW16], where one can deterministically test (in polynomial time) if the polynomial $\det\left(\sum_i A_i \otimes X_i\right)$ is identically zero. Here A_i's are $n \times n$ matrices and X_i's are $d \times d$ matrices with entries distinct formal commuting variables with $d = n$. Note that the PIT problem corresponds to $d = 1$.

Several excellent resources (lecture notes, slides and videos) are available online for the reader who wants to learn about these topics. These include Avi's lectures at CCC 2017, workshop at IAS, workshop at FOCS 2018 and the recent survey [GO18]. The links for the above are below:

http://www.computationalcomplexity.org/Archive/2017/tutorial.php
https://www.math.ias.edu/ocit2018
https://staff.fnwi.uva.nl/m.walter/focs2018scaling/

References

[AGLOW18] Zeyuan Allen-Zhu, Ankit Garg, Yuanzhi Li, Rafael Oliveira, and Avi Wigderson. Operator scaling via geodesically convex optimization, invariant theory and polynomial identity testing. *STOC*, 2018.

[BFGOWW18] Peter Bürgisser, Cole Franks, Ankit Garg, Rafael Oliveira, Michael Walter, and Avi Wigderson. Efficient algorithms for tensor scaling, quantum marginals and moment polytopes. *arXiv preprint arXiv:1804.04739*, 2018.

[BGOWW18] Peter Bürgisser, Ankit Garg, Rafael Oliveira, Michael Walter, and Avi Wigderson. Alternating minimization, scaling algorithms, and the null-cone problem from invariant theory. *Proceedings of Innovations in Theoretical Computer Science (ITCS)*, 2018.

[Bür12] Peter Bürgisser. Prospects for geometric complexity theory. In *2012 IEEE 27th Conference on Computational Complexity*, pages 235–235, June 2012.

[GGOW16] Ankit Garg, Leonid Gurvits, Rafael Oliveira, and Avi Wigderson. A deterministic polynomial time algorithm for non-commutative rational identity testing. In *Proceedings of the Symposium on Foundations of Computer Science (FOCS 2016)*, pages 109–117. IEEE, 2016.

[GGOW17] Ankit Garg, Leonid Gurvits, Rafael Oliveira, and Avi Wigderson. Algorithmic and optimization aspects of brascamp-lieb inequalities, via operator scaling. In *Proceedings of the Symposium on the Theory of Computing (STOC 2017)*, pages 397–409. ACM, 2017.

[GO18] Ankit Garg and Rafael Oliveira. Recent progress on scaling algorithms and applications. *arXiv preprint arXiv:1808.09669*, 2018.

[Hil90] David Hilbert. Ueber die Theorie der algebraischen Formen. *Mathematische Annalen*, 36(4):473–534, 1890.

[Hil93] David Hilbert. Uber die vollen Invariantensysteme. *Math. Ann.*, 42:313–370, 1893.

[KN79] George Kempf and Linda Ness. The length of vectors in representation spaces. In *Algebraic Geometry, Lecture Notes in Math.*, pages 233–243. 1979.

[Mum65] David Mumford. *Geometric invariant theory*. Springer-Verlag, Berlin-New York, 1965.

[Sin64] R. Sinkhorn. A relationship between arbitrary positive matrices and doubly stochastic matrices. *The Annals of Mathematical Statistics*, 35:876–879, 1964.

[Woo11] Chris Woodward. Moment maps and geometric invariant theory. *Lecture notes, arXiv:0912.1132*, 2011.

Circuit Lower Bounds for Quasi-NP

Ryan Williams

(joint work with Cody Murray)

Recall ACC^0 is the class of decision problems that have polynomial-size circuit families of constant depth, with unbounded fan-in AND, OR, and MODm gates for a fixed integer $m > 2$, where a MODm is a Boolean function outputting 1 if and only if the sum of its inputs is divisible by m [Bar89]. The class ACC^0 has been widely conjectured for over 30 years to be an extraordinarily small class of functions (it is believed that even the simple MAJORITY function is not contained in ACC^0 [Bar89]).

However, it was only recently that any non-trivial lower bound was shown against ACC^0. In 2011, I showed the first non-trivial lower bound against polynomial-size ACC^0: there are functions in the (gigantic) complexity class $NEXP$ that do not have (quasi) polynomial-size ACC^0 circuits [Wil14]. This lower bound was proved by exploiting a long-known representation of ACC^0 in a new way. In particular, ACC^0 circuits can be efficiently translated into depth-two circuits of symmetric gates; this translation can be used to devise a Satisfiability (SAT) algorithm for ACC^0 circuits that is a notable improvement over exhaustive search, running in 2^{n-n^ϵ} time on 2^{n^ϵ}-size circuits for sufficiently small $\epsilon > 0$ (depending on the depth and modulus m of the circuit). Furthermore, such a SAT algorithm can be used to derive lower bounds against $NEXP$-complete functions.

Within the last year, the complexity of the hard-for-ACC^0 function has been significantly improved. In this talk, I gave an overview of the recent result [MW18] (with Cody Murray) that Quasi-NP does not have ACC^0 circuits of polynomial size. This new lower bound was achieved by improving the known generic connections between SAT algorithms and circuit lower bounds (reducing the complexity of the hard function obtained), and applying the aforementioned ACC^0 SAT algorithm [Wil14].

The main new ingredient in our work is an *easy witness lemma* for Quasi-NP and NP, extending the easy witness lemma for $NEXP$ [IKW02]. This lemma

effectively says that if Quasi-NP (or NP) problems can be decided by very small circuit families, then every verifier for every Quasi-NP (respectively, NP) problem has *easy witnesses*: every yes-instance of the problem admits a witness that can be represented by a very small circuit. (Such a structural lemma is needed in the connection between SAT algorithms and circuit lower bounds.) Within the proof of the new easy witness lemma, the main ingredient is a new circuit lower bound for Merlin-Arthur games, which shows there are Merlin-Arthur games (with mild non-uniformity) which do not have fixed-polynomial circuits, even for relatively "sparse" sequences of input lengths.

References

[Bar89] David A.Barrington. *Bounded-width polynomial-size branching programs recognize exactly those languages in NC1.* J. Comput. Syst. Sci. 38(1) (1989), 150–164.

[IKW02] Russell Impagliazzo, Valentine Kabanets, and Avi Wigderson. *In search of an easy witness: exponential time vs. probabilistic polynomial time.* J. Comput. Syst. Sci. 65(4) (2002), 672–694.

[MW18] Cody Murray and R. Ryan Williams. *Circuit lower bounds for nondeterministic quasi-polytime: an easy witness lemma for NP and NQP.* Proceedings of *STOC* (2018), 890–901.

[Wil14] Ryan Williams. *Nonuniform ACC Circuit Lower Bounds.* J. ACM 61(1) (2014), 2:1–2:32.

Reporter: Roei Tell