Publication List and Description of Work

Shuo Pang

ORCID
`https://orcid.org/0000-0001-5652-9737`

Google Scholar
`https://scholar.google.com/citations?hl=en&user=kCQFJi8AAAAJ`

1. *Graph colouring is hard on average for Polynomial Calculus and Nullstellensatz*

   With Jonas Conneryd, Susanna de Rezende, Jakob Norström, Kilian Risse

   FOCS 2023

   Description. This work proves that Polynomial Calculus, hence also Nullstellensatz, requires linear degree to refute the 3-colorability of sparse random graphs and random regular graphs. This gives optimal, exponential size lower bounds via the known size-degree relation.

   The proof constructs an Alekhnovich-Razborov pseudo-reduction operator, based on the technique in [Romero-Tunçel 2022] for dealing with large-girth graphs. We extend this technique to rid of the large-girth assumption, which enables us to deal with general sparse expanders in an elegant and simple way.

2. *SoS lower bound for exact planted clique*

   CCC 2021

   Description. This work proves Sum-of-Squares (SoS) degree lower bounds for the Exact Planted Clique problem on random graphs $G(n, \frac{1}{2})$. The algorithms' task is to refute the existence of a size ω clique in a sample of random graphs, where ω is so large that with probability say $> 99.99\%$ the sample cannot contain such a large clique. Here, by 'exact' we mean the SoS algorithms are allowed to reason with polynomial identities generated by axioms including $\sum_{i=1}^{n} x_i = \omega$, the one about the size of the (nonexistent) clique. This is a canonical and feasible modeling of the SDP algorithms, while previous lower bound method have to weaken this axiom into an approximate one. We overcome this shortcoming by proving a SoS degree lower bound in this setting in the form $d = \Omega(\frac{\epsilon^2 \log n}{\log\log n})$ if $\omega = O(n^{\frac{1}{2}-\epsilon})$, which is almost optimal in d and ω.

One motivation of the work is to further develop the average-case SoS lower bound techniques. We define the pseudo-expectation operator $\widetilde{E}(\cdot)$ differently than the popular pseudo-calibration method; the cost is, the resulting moment matrix is complicated and is quite *rigid* in a sense, so that the factorization method breaks down. We use a two-step decomposition to deal with it: 1. An Hadamard product with Johnson schemes, inspired by [Feige-Krauthgamer03]; and 2. A relativized factorization [Barak-Hopkins-Kelner-Kothari-Moitra-Potechin 2016]. The task then reduces to studying a special family of matrices, which we term the factorial Hankel matrices. The decomposition argument relies on somewhat intricate combinatorial transforms, whose applicability depends on the choice of $\widetilde{E}(\cdot)$ in, again, a very rigid way. A conceptual question remains: is there a clear a priori rationale, in either analytic or combinatorial terms, that justifies the apparently contingent choice we made here?

3. *On CDCL-based proof systems with the ordered decision strategy*

   With Nathan Mull and Alexander Razborov

   SAT 2020, SICOMP 2022

   Description. In this work, we prove that conflict-driven clause learning (CDCL) SAT-solvers with the *ordered decision strategy* and *DECISION learning scheme*, are equivalent to ordered resolution. We also prove that if replacing this learning scheme with its opposite, which stops backtracking right after the first non-conflict clause, then the solvers become equivalent to general resolution. This is among the first theoretical studies of the interplay between specific decision strategies and clause learning.

   For both results, we allow nondeterminism in the solver's ability to perform unit propagation, conflict analysis, and restarts, in a way similar to previous works in the literature. To aid the presentation of our results, and possibly future research, we define a model and language for discussing CDCL-based proof systems that allow for succinct and precise theorem statements.

4. *Large clique is hard on average for resolution*

   CSR 2021

   Description. The main result is a $2^{\Omega\left(k^{1-o(1)}\right)}$ resolution size lower bounds for the k-Clique problem on suitable random graphs, for $k < n^{1/3}$. This complements the result in [Beame-Impagliazzo-Subharwal 2007] which is for $k > n^{5/6}$. The proof uses the classical bottleneck counting/random restriction framework plus a variant of clause width, which is defined using neighborhood density that appeared in [Beyesdorff-Galesi-Lauria 2013, Atserias-Bonacina-De Rezende-Lauria-Nordström-Razborov 2018].

Work in Preparation

5. *Supercritical and Robust Trade-offs for Resolution Depth Versus Width and Weisfeiler–Leman*

   With Duri Janett and Jakob Norström

   Description. We present the first robust resolution trade-offs for which low width implies depth superlinear in the formula size. We give analogous results for the Weisfeiler–Leman algorithm, which also translate into trade-offs between number of variables and quantifier depth in first-order logic.

   Our main technical contribution is a new compression scheme, together with its analysis, of the so-called compressed Cop-Robber game introduced by [Grohe-Lichter-Neuen-Schweitzer 2023].

6. *SoS lower bounds for Non-Gaussian Component Analysis*

   With Ilias Diakonikolas, Sushrut Karmalkar, and Aaron Potechin

   Description. Non-Gaussian Component Analysis (NGCA) is a fundamental question in statistics and machine learning. Given i.i.d. samples from a distribution P_v on $\mathbb{R}^n$ that behaves like a known distribution A in a hidden direction v and like a standard Gaussian in $v^{\perp}$, the goal is to approximate the hidden direction. When A matches low-order moments with Gaussian and satisfies some mild conditions, it is known that 'traditional' learning algorithms, like Statistical Query and low-degree polynomial tests, require a large number of samples.

   Here we study the complexity of NGCA in the Sum-of-Squares (SoS) algorithmic framework. Our main contribution is the first super-constant degree SoS lower bound, which translates to super-polynomial time for the corresponding Semi-Definite Programming algorithms. Specifically, we show that if the distribution A matches the first $(k-1)$ moments of $\mathcal{N}(0,1)$ and satisfies mild conditions, then when fed with $< n^{\frac{(1-\epsilon)k}{2}}$ many Gaussian samples, with high probability degree $o(\sqrt{\log n})$ SoS fails to refute the existence of direction v. This significantly strengthens prior work by establishing a lower bound against a broader family of algorithms.

The lower bound proof introduces a new technique along with a few improvements of some previous ones. As in previous works, we use the framework of [Barak et al. 2016] where for the moment matrix M, we find an approximate factorization $M \approx LQL^T$ using minimum vertex separators, and show that with high probability Q is PSD while error terms are small. What's new is the following. First, instead of the minimum weight vertex separator, we use the minimum square separator. Second, proving that the matrix Q is positive is challenging due to an intrinsic reason. In all prior works, the matrix Q modulo a negligible error was a constant term, meaning its entries are constant functions of the input. Here, however, it is (quite) a nontrivial linear combination of a class of non-constant, equally dominating terms. To address this difficulty, we introduce an algebraic method that we believe is of more general interest. Specifically, we model the multiplications between the "important" pseudo-random matrices by an $\mathbb{R}$-algebra, construct all irreducible representations of this algebra, and use the resulting Wedderburn-Artin decomposition to analyze Q. Via this approach, we show that the PSDness of Q boils down to the multiplicative identities of Hermite polynomials.