Description of Work

Shuo Pang

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

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

6. *Truly Supercritical Trade-offs for Resolution, Cutting Planes, Monotone Circuits, and Weisfeiler-Leman*

   To write

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

   With Ilias Diakonikolas, Sushrut Karmalkar, and Aaron Potechin

   To appear at FOCS 2024

   Description. Non-Gaussian Component Analysis (NGCA) is the task of finding a non-Gaussian direction in a high-dimensional distribution—in short, detecting a signal among noise. 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 $v^\perp$, we want to find v. Of particular interest is when the distribution A matches the low-order moments with the Gaussian, in which case it is known (under mild conditions) that the current learning algorithms such as statistical-query and low-degree polynomial tests require a large number of samples to complete the task.

   This work studies the complexity of NGCA in the Sum-of-Squares (SoS) algorithmic framework, a state-of-the-art class of algorithms that broadly strengthens the aforementioned ones. We prove a super-constant degree SoS lower bound, which translates to super-polynomial time for the corresponding algorithms. Specifically, we show that if distribution A

matches the first $(k-1)$ moments of $\mathcal{N}(0,1)$ and satisfies mild conditions, then with high probability, even with $< n^{\frac{(1-\epsilon)k}{2}}$ many Gaussian samples the degree $\widetilde{o}(\sqrt{\log n})$ SoS still fails to refute the existence of direction v. This sample lower bound is nearly tight since there's an upper bound $n^{k/2}$ [Dudeja-Hsu 22]. As a corollary, we obtain SoS lower bounds for several problems in robust statistics and learning of mixture models.

The proof introduces a new technique, along with a few improvements of the previous ones. We begin with the pseudo-calibration framework [Barak-Hopkins-Kelner-Kothari-Moitra-Potechin 2016] where, given 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, we make a shift in approach for analyzing Q for a rather intrinsic-looking reason as follows. In all prior works, Q, modulo a negligible error, is a "real matrix" whose entries are numbers. Here, however, it is (quite) a nontrivial linear combination of a class of equally dominating pseudo-random matrices (i.e., *graph matrices*, whose entries are polynomials). We introduce an algebraic method that may be of more general interest. Namely, we model the multiplications between the "important" pseudo-random matrices by an $\mathbb{R}$-algebra, construct all irreducible representations of this algebra, and then use them to study the special element Q. Via this approach, we see that the PSDness of Q boils down to the multiplicative identities of Hermite polynomials.

4. *Graph Colouring Is Hard on Average for Polynomial Calculus and Nullstellensatz*
With Jonas Conneryd, Susanna de Rezende, Jakob Norström, Kilian Risse
FOCS 2023

Description. In this work, we prove that Polynomial Calculus, hence also Nullstellensatz, requires linear degree to refute the 3-colorability of sparse random graphs and random regular graphs. An optimal, exponential size lower bound follows via the known size-degree relation.

The proof goes by constructing an Alekhnovich-Razborov pseudo-reduction operator, where the core task is about constructing the closure of each small vertex set. Here, a closed vertex set S can be thought of as a (small) vertex subset that already provides sufficient information for any low-degree prover to reason about the colours of vertices in it—

that is, if one wants to use of the information of vertices outside of S *to deduce more about the possible colourings of* S, one would inevitably speak of high-degree polynomials. To construct the closed sets, we use a technique introduced by [Romero-Tunçel 2021] in the context graphs with large girth. We extend this technique to get rid of the large-girth assumption, where the key is to upper bound the size of a closed set using the everywhere-sparseness of the graph. In the case of random regular graph with vertex degree d, every-where sparseness amounts to being an $(d-2-\epsilon)$-expander; for random graphs, we handle large-degree vertices by always including them in the construction of a closed set.

3. *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 $G(n, 1/2)$. The SoS algorithm aims to refute the existence of cliques of size ω in a sampled graph, where ω is so large that with probability $> 99.99\%$ the graph has no such cliques. Our task is to show the algorithm has to fail miserably unless its degree is so high that a brute force is possible.

   The word *exact* in the title means that the algorithms have access to the full set of axioms, including the one on clique size $\sum_{i=1}^{n} x_i = \omega$, often called the "global" axiom, which the previous tight lower bound technique [Barak-Hopkins-Kelner-Kothari-Moitra-Potechin 2016] needs to weaken in some way (into an objective function). This work deals with this problem by showing a degree lower bound $d = \Omega(\frac{\epsilon^2 \log n}{\log \log n})$ for $\omega = O(n^{\frac{1}{2}-\epsilon})$, which is almost optimal in both d and ω.

   Another motivation is to further the average-case SoS lower bound techniques. To deal with the global axiom, we design the *pseudo-expectation* $\widetilde{E}(\cdot)$ differently from the pseudo-calibration in [Barak et al. 2016]. Cost is, the moment matrix no longer has entries in a product-like form, making it harder to analyze. To address this, we simplify the target matrix using the Hadamard product (one factor being a Johnson scheme, inspired by [Feige-Krauthgamer 2003]) and then use a relativized matrix factorization to the major factor (where 'relativized' means conditioning on each subset being in the assumed clique). The final positive semidefiniteness (PSD) proof relies on some combinatorial transforms and the analytical properties of a matrix family we call *factorial Hankel* matrices. In retrospect, the

success of this approach relies on the design of $\widetilde{E}(\cdot)$ being "correct", but is there an a priori explanation for this design, which admittedly appears contingent?

2. *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, by replacing this learning scheme with its opposite—which stops backtracking right after the first non-conflict clause—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. 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.

1. *Large Clique Is Hard on Average for Resolution*
   CSR 2021

   Description. The main result of the paper is a $2^{\Omega\left(k^{1-o(1)}\right)}$ resolution size lower bound for the k-Clique problem on suitable random graph models, where $k < n^{1/3}$. This complements the result in [Beame-Impagliazzo-Subharwal 2007] which holds for $k > n^{5/6}$.

   Our proof is based on the bottleneck counting framework, where we use a variant of clause width. This width variant is defined by thresholding the 'density' of the vertex sets associated with a clause, using the same notion of density as in previous works [Beyesdorff-Galesi-Lauria 2013, Atserias-Bonacina-De Rezende-Lauria-Nordström-Razborov 2018]. We also extend the $n^{\Omega(k)}$ regular resolution lower bound [ABDLNR 2018] to a slightly stronger system that permits a certain degree of irregularity.